School Improvement Grants in Ohio: Effects on Student Achievement and School Administration

January 2018

Deven Carlson
University of Oklahoma
Department of Political Science
205 Dale Hall Tower
455 W. Lindsey St.
Norman, OK 73019
decarlson@ou.edu

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

Author Bios
DEVEN CARLSON is Presidential Research Professor, Associate Professor of Political Science, and Associate Director for Education at the National Institute for Risk and Resilience at the University of Oklahoma. His research analyzes the operations of school accountability and school choice policies, and explores their effects on social, economic, and political outcomes.

STÉPHANE LAVERTU is an associate professor at The Ohio State University's John Glenn College of Public Affairs. His research focuses on K-12 education governance, including school boards, charter schools, school finance, accountability systems, and interventions such as school closure and turnaround.

Acknowledgments:
We thank the Ohio Department of Education—particularly Sherry Panizo, Eben Dowell, Jo Hannah Ward, Matt Cohen, and Paul Conaway—for their help in tracking down the information we needed to complete the analysis. We also thank the staff at the Ohio Education Research Center, particularly Lisa Neilson and Josh Hawley, for providing data. Finally, we thank Andrew McEachin, Randy Olsen, and two anonymous reviewers for providing comments and suggestions that improved the quality of this article.

Abstract:
The federal School Improvement Grant (SIG) program allocated $7 billion over nearly a decade in an effort to produce rapid and lasting improvements in schools identified as low performing. In this paper, we use a regression discontinuity design to estimate the effect of Ohio's SIG turnaround efforts on student achievement and school administration. The results indicate that Ohio's SIG program significantly increased reading and math achievement, with effects in both subjects of up to 0.20 standard deviations in the second year after SIG eligibility identification. Estimates for the third year are somewhat larger, in the range of one-quarter of a standard deviation. We provide evidence that these effects were primarily attributable to schools that implemented the SIG Turnaround model. We also show that SIG eligibility had a positive effect on per-pupil spending, but no average effect on administrative outcomes, including staff turnover, the number of staff members in the school, and school closure. These null overall effects mask heterogeneity across SIG models, however. Most notably, Turnaround schools experienced more turnover than they otherwise would have, whereas Transformation schools experienced less.

Keywords: School turnaround; School Improvement Grant; student achievement; teacher turnover; regression discontinuity

Introduction

Federal education policy has supported multiple initiatives designed to improve the quality of low-performing schools. Perhaps the most significant of these initiatives is the federal School Improvement Grant (SIG) program, which allocated $7 billion over nearly a decade in an effort to produce rapid and lasting improvements in the quality of low-performing schools. SIG was designed to achieve such improvements by providing schools with additional financial resources and by requiring significant changes to many aspects of schools' educational delivery, particularly their leadership and staffing, as well as their use of data to drive instructional and managerial decision-making. The extent to which the SIG program has had its intended effect, however, is an open question. A small body of work has evaluated the effectiveness of SIG, with the federally sponsored evaluation finding no significant effect on student achievement or attainment (Dragoset et al. 2017) and state-specific studies returning a mix of null and positive effects (Dee 2012; LiCalsi et al. 2015; Heissel and Ladd 2016; Henry and Guthrie 2015; Papay 2015).

We estimate the effect of Ohio's SIG turnaround efforts on student achievement and school administration. In particular, we estimate the effect of SIG eligibility on student achievement levels, per-pupil funding levels, staff turnover, and school closure. The focus on SIG eligibility—the intention-to-treat (ITT) parameter—provides insight into the effects of a SIG program. However, because nearly all SIG-eligible schools in Ohio received SIG awards, the estimated effects of SIG receipt—the treatment-on-the-treated (TOT) parameter—are very similar to the effects of SIG eligibility. We estimate these effects using a regression discontinuity (RD) design, exploiting the SIG eligibility criterion stating that schools qualified for these interventions if they ranked in the bottom 5 percent of eligible schools in terms of a weighted student proficiency rate calculated by the Ohio Department of Education (ODE). Our analyses are based on data contained in multiple sets of records provided by ODE via the Ohio Education Research Center, including individual-level records for all students, teachers, and principals in Ohio public schools between the 2008-09 and 2014-15 school years. We supplement these records with a wide range of annual school- and district-level data spanning the same time period.

Our analyses show that Ohio's SIG program had large positive effects on student achievement. These effects first emerge in the second year after a school is identified as SIG-eligible, with reading effects in excess of 0.11 standard deviations and math effects of more than 0.14 standard deviations. In the third year, SIG eligibility is estimated to increase reading achievement by at least 0.20 standard deviations and math achievement by at least 0.18 standard deviations, and some specifications suggest achievement gains of up to 0.26 standard deviations in math and 0.27 standard deviations in reading. All of these achievement results are robust to a variety of alternative analytic decisions and model specifications.

The positive achievement effects coincide with a significant infusion of cash into SIG-eligible schools. We estimate that SIG eligibility increases annual per-pupil funding by $1,500-$3,000 across the three-year grant period, which suggests that SIG funds did not supplant district funds. Further, we provide evidence that these positive achievement effects were driven by schools that implemented SIG's Turnaround model, as opposed to the Transformation model. Interestingly, our analysis also finds SIG eligibility to have little effect on several other school administrative outcomes, including staff turnover, the number of staff members in the school, and school closure. However, for several of these outcomes we provide evidence of heterogeneous effects across schools that implemented the SIG Transformation model and those that implemented the Turnaround model.

We proceed by first providing context for our analysis, reviewing existing research relevant to school turnaround initiatives and describing Ohio's SIG program in detail. We then describe the data underlying the analysis and detail the RD approach we use to estimate the effects of the two turnaround interventions. Next, we present the results, showing the effects of the turnaround intervention on student achievement and school administration. Finally, we conclude the paper by discussing the implications of our results for school turnaround policy, as well as for research on these initiatives.

Background on School Turnaround Initiatives

There is a long history of efforts to improve the quality of low-performing schools in the United States. Slavin (1989) characterized many of the early efforts as piecemeal—targeting particular aspects of educational delivery or specific student populations—and generally unsuccessful, due at least in part to the short-lived and fragmented nature of many of these initiatives (Hess 1998; Gross, Booker, and Goldhaber 2009). Partially in response to the perceived failure of this scattershot approach, comprehensive school reform (CSR) programs steadily gained popularity beginning in the late-1980s and extending through the 2000s. These initiatives were based on the notion that coordinated efforts to improve schools as a whole were more likely to have a meaningful impact than the targeted and disjointed efforts that preceded them.[1] These federal grant programs—administered primarily under Title I of the Elementary and Secondary Education Act (ESEA)—were quite specific about which interventions qualified as CSR, stipulating that they must involve evidence-based strategies for improving everything from school management and instruction to fostering parental and community involvement. Research examining the impact of various CSR programs on student achievement found mixed results.

[1] See Borman et al. 2003 for an extended description of CSR efforts.

Though many studies were of limited quality (Herman et al., 2008), there is good evidence that some CSR models had a positive impact on achievement (e.g., see Bifulco, Duncombe, and Yinger, 2005; Borman et al., 2003; Gross, Booker, and Goldhaber, 2009).

The encouraging evidence emerging from CSR evaluations contributed to the federal government's decision in the late-2000s to incentivize implementation of more aggressive school turnaround models.[2] The U.S. Department Education encouraged the adoption of such policies through the promise of additional financial resources or waivers from certain accountability provisions of No Child Left Behind (NCLB). This set of more aggressive turnaround strategies includes closing a school completely, restarting a school as a charter school or one managed by an education management organization, or, to various degrees, reconstituting a school's leadership and instructional staff through mandatory and data-driven hiring and firing processes.

A body of evidence suggests that these turnaround models could lead to improvements in student achievement. Research on school closure and charter conversions, for example, indicates that closure can benefit students in failing schools provided that affected students switch to schools of sufficiently higher quality to compensate for disruptions induced by closure or conversion (Brummet 2014; Carlson and Lavertu 2015; Carlson and Lavertu 2016a; Bross, Harris, and Liu 2016; Abdulkadiroglu et al. 2016). The turnaround models requiring reconstitution of leadership and staff are supported by work demonstrating the importance of principals (Branch, Hanushek, & Rivkin 2012; Grissom, Kalogrides, and Loeb 2015) and teachers (Hanushek 2011; Chetty, Friedman, and Rockoff 2014; Rivkin, Hanushek, and Kain 2005). Perhaps most encouraging for supporters of these models, though, is work using data

[2] "Turnaround" is the general school improvement strategy pursued under the SIG initiatives. Confusingly, it also is the label assigned to a specific SIG model. We capitalize all references to the specific SIG model and use the lower-case "turnaround" to refer to the more generic reform strategy.

from the District of Columbia’s teacher evaluation system, which shows that the replacement of low-performing teachers with higher-quality ones increased student achievement by a statistically significant and substantively meaningful magnitude (Adnot et al. forthcoming). Finally, the general focus on data-driven decision-making in these turnaround models also has support in the literature (Strunk and McEachin, 2014; Strunk, McEachin, and Westover, 2014; Carlson, Borman, and Robinson 2011).

Although research suggests that turnaround strategies could lead to improved student outcomes, it is not a foregone conclusion. Several studies have found school closure to have no meaningful effect on student achievement, apparently because students do not end up in schools of sufficiently greater quality (Brummet 2014; Engberg et al. 2012; de la Torre and Gwynne 2009). Similarly, the effectiveness of turnaround models requiring principal or teacher replacement depends on the supply of human capital. If the supply of educational personnel—teachers, principals, or school staff—is low or recruitment is difficult, as tends to be the case in low-achieving, high-poverty urban and rural districts (e.g., see Boyd, Lankford, & Wyckoff, 2007; Cowen et al., 2012; Jackson, 2009; Clotfelter, Ladd, & Vigdor, 2007, 2010), new personnel could very well be of comparable or lower quality than those they replace. Even if existing staff is replaced with personnel of comparable quality, such “churn” has itself been shown to have a negative impact on student achievement (Atteberry, Loeb, and Wyckoff 2016).

Overall, research does not provide an unambiguous prediction concerning the effects of school turnaround efforts on student outcomes. Instead, the evidence suggests that the utility of these interventions is likely to be context specific.

School Improvement Grants in Ohio

Authorized under Title I of the ESEA, the SIG program allocates formula-based grants to state education agencies (SEAs). The SEAs then make competitive awards to local school districts that demonstrate a compelling plan for using the funds to improve student achievement in SIG-eligible schools. Districts' plans must adhere to federal SIG requirements, which were significantly revamped in 2009. This redesign coincided with a one-time influx of $3 billion to the program through the American Recovery and Reinvestment Act (ARRA) in addition to the regular annual appropriation of approximately $500 million. Of this $3.5 billion, Ohio received over $130 million to award to districts with eligible schools. The state distributed this money to three cohorts of districts and their eligible schools, with each cohort receiving grant dollars over a three-year period. According to data from ODE, the median annual SIG award for each school was about $800,000, or about $2,200 per pupil, and over half of that was dedicated to personnel salaries and fringe benefits.

[Insert Table 1 about here]

Prior to receiving funds from the U.S. Department of Education (USED), ODE was required to provide USED with a list of SIG-eligible schools and outline the criteria they would use to award grants to districts. The SIG eligibility criteria were complex, consisting of the three eligibility tiers summarized in Table 1. In general, a school's eligibility was a function of its Title I receipt status under NCLB's school improvement process, grades served (i.e. elementary or middle school vs. high school), achievement levels, and graduation rates. Although schools in all three tiers were eligible to receive SIG funds, program requirements specified that districts prioritize Tier 1 and 2 schools over those in Tier 3.[3]

The achievement measures determining Tier 1 and 2 SIG eligibility are instrumental to our analyses below and thus warrant further explanation. Consistent with federal guidelines, ODE determined the lowest achieving five percent of schools using an average of two

[3] See Dragoset et al. 2017 for a comprehensive description of SIG eligibility criteria.

proficiency calculations: 1) A weighted proficiency rate in math and reading for each building in the most recent school year, and 2) An average of this weighted proficiency rate over the five most recent school years. ODE then rank-ordered schools based on a "combined proficiency rate" that weighted these two proficiency calculations equally.[4] These ranking rules were applied separately for Tier 1 and Tier 2 schools and for each cohort.

After identifying all schools eligible to receive SIG funds, ODE invited districts to submit applications that demonstrated commitment and capacity to improve student achievement in SIG-eligible schools. These applications included detailed narratives and budgets for each school that would receive funds. SIG regulations required Tier 1 and Tier 2 schools awarded SIG funds to implement one of four federally-approved turnaround models: Closure, Restart, Transformation, or Turnaround. The Closure model required a school to shut down permanently while the Restart model mandated that a school convert to a charter school or one operated by an education management organization. The Transformation and Turnaround models both required replacing the principal and implementing a number of instructional and operational reforms. The Turnaround model further entailed replacing at least 50 percent of the school staff while the Transformation model required adopting a new governance model and an educator evaluation system that included student achievement growth as a significant component.[5] Table 2 summarizes the three cohorts of SIG grants in Ohio. In particular, it lays out the timing of SIG eligibility identification and grant administration, depicts the number of districts and schools receiving SIG awards, and presents the number of Tier 1 and 2 schools that implemented each of the four SIG turnaround models. The table shows that the first cohort of schools were identified

[4] Schools could also be determined to be SIG-eligible based on graduation rates. We do not conduct an analysis of high school turnaround efforts and thus do not use this criterion in our analysis.
[5] Federal rules stipulated that a district that had nine or more Tier 1 and Tier 2 schools was not permitted to implement the Transformation model in more than 50 percent of those schools.

as eligible in the 2009-10 school year—we consider this the first treatment year for the cohort—and grant funds were first administered the following year. The second cohort of SIG awards were made in 2010-11, but there was then a three-year gap between the second and third SIG cohorts. Table 2 further shows that ODE made awards to 39 schools in 13 districts in Cohort 1, 44 schools in 29 districts in Cohort 2, and 24 schools in 11 districts in Cohort 3—we cannot include Cohort 3 in our analysis for reasons we describe in greater detail below. It also demonstrates that Tier 1 and 2 schools overwhelmingly elected to implement the Transformation and, to a lesser extent, Turnaround models.

[Insert Table 2 about here]

Data

The analyses are based on data contained in multiple sets of records. First, for each of the first two SIG cohorts, ODE provided the combined proficiency rate calculations used to rank schools to determine eligibility. These calculations are critical to the research design we employ. Unfortunately, the three-year gap between the second and third SIG cohorts, coupled with the terms of the evaluation with ODE that specified a focus on the first two cohorts, prevents us from including the final cohort in our analysis. Second, ODE provided a wide variety of annual school-level information from the 2006-07 through 2014-15 school years, including enrollment, demographic and socioeconomic composition, expenditures, attendance and mobility rates, and characteristics of the teaching staff. We also obtained several annual school-level achievement-related measures, including average standardized achievement levels in reading and math, scores on Ohio's Performance Index, and estimated value-added in reading and math. The Performance Index has a 0-120 scale and captures the achievement level of a school's students across multiple subjects (math, reading, writing, science, and social studies) on state assessments. School value-added estimates are calculated annually for use in Ohio's accountability system and are available from ODE in Normal Curve Equivalent (NCE) units, which we converted to standard deviation units for use in our analyses.[6]

Third, ODE provided student-level data via the Ohio Education Research Center (OERC). These records include identifiers for the schools and districts that students attended, as well as information on student demographic characteristics, such as sex, race/ethnicity, economic disadvantage, disability status, and English language learner (ELL) status. They also include students' scale scores on the Ohio Achievement Assessment (OAA)—administered in grades 3-8—which we standardized by subject, grade, and year.

Finally, we obtained annual staff-level records from ODE via the OERC for the 2007-08 through 2013-14 school years. These records contain information on staff members' position categories—education professional, administrator, support staff, etc.—and their specific positions within those categories, such as teacher, principal, and custodian, respectively. The records also include the individual's date of hire and, if applicable, separation, as well as the reason for the separation (e.g., resigned, retired, or terminated). We use these records to create the staff turnover measures that we employ in our analyses below.

Empirical Strategy: Regression Discontinuity Design

We leverage the eligibility criteria determining SIG eligibility—particularly the proficiency rate thresholds demarcating the bottom 5 percent of schools—to estimate the effects of Ohio's school turnaround efforts on student outcomes and school administration. As described above, schools were generally eligible for the turnaround intervention if they were in the bottom five percent of

[6] We performed this conversion by dividing the value-added score by the NCE standard deviation of 21.06. It is important to note that Ohio's value-added estimates account for up to five prior years of student test scores and, thus, should effectively control for differences in the students schools educate. It is also worth noting that ODE's publicly available 2013 and 2014 school value-added estimates are three-year averages (e.g., the estimate for 2013 is an enrollment-weighted average of estimates from 2011, 2012, and 2013). We backed out estimates for 2013 and, then, 2014 using the one-year estimates from 2011 and 2012 and official school enrollments. It is also important to note that the value-added "gain scores" we use for this analysis should not be confused with some of the value-added indices that ODE makes publicly available.

schools in terms of the combined proficiency rate.[7] Thus, an arbitrary threshold separates schools eligible to receive turnaround interventions from those that were not. For example, in the first SIG cohort a school with a combined proficiency rate of 29.15 was a Tier 1-eligible school, but one with a combined proficiency rate of 29.2 was not. We exploit this policy rule to estimate the effects of turnaround efforts using an RD design, which bases identification of the effect of turnaround on the assumption that schools just below the eligibility threshold are essentially indistinguishable from those just above the threshold.

Sample and Construction of Running Variable

We estimate the effect of SIG eligibility—the ITT parameter—for Tier 1 schools in the first two cohorts of the program.[8] For each of the first two cohorts we use the combined proficiency rate data that ODE provided to identify all potential Tier 1 schools, as well as the cutoff corresponding to the lowest five percent of those schools.

ODE selected the first cohort of SIG-eligible schools from 724 Title I-served schools in NCLB's school improvement process (i.e. potential Tier 1 schools). The bottom five percent thus consisted of 36 schools and a corresponding cutoff of 29.175 on the combined proficiency rate measure. Schools below this cutoff were Tier 1-eligible on the basis of their achievement levels whereas schools above the cutoff were not.[9] In the second SIG cohort there were 765 potential Tier 1 schools. The lowest five percent thus consisted of 38 schools and 33.66 was the corresponding cutoff on the combined proficiency rate measure. Ohio administered their SIG program such that, in theory, a school's award status in the first round of SIG awards did not

[7] As noted earlier, schools were also eligible for SIG on the basis of graduation rates. However, this applied to only a small number of schools and we thus do not use this criterion in our analyses.
[8] As noted earlier, we exclude the third SIG cohort from our analysis because ODE did not provide us with the combined proficiency rate calculations used to determine a school's distance from the SIG eligibility threshold.
[9] Specifically, schools above the 29.175 cutoff were not Tier 1-eigible on the basis of their achievement rates. A small number of schools—fewer than 10—were Tier 1-eligible because they had graduation rates lower than 60 percent.

affect its eligibility for the second round of SIG awards. That is, if a school was a SIG winner in the first round of awards, but fell in the bottom 5 percent of schools in the second round, then it was formally eligible to receive a second-round SIG award as well.[10] In practice, though, no school that received a SIG award in the first round also received an award in round two.

For each cohort—and each school within each cohort—we subtracted the combined proficiency rate cutoff identifying the lowest five percent of schools from each school's score on that measure. The resulting value captures a school's distance from the cutoff determining Tier 1 SIG eligibility—at least on the basis of achievement levels—and serves as the running variable in our analyses below. We then extracted annual student, school, and district records for all potential Tier 1 schools, starting with the year of eligibility identification and extending through the subsequent four years. Finally, to increase our ability to detect effects of SIG eligibility, we pooled the two cohorts to create a single dataset. RD designs have less statistical power than other designs, particularly when the treatment is assigned at a group—rather than individual—level, as is the case with SIG. In such cases, compared to an experimental design, RD designs require a sample four times larger in order to generate estimates with the same degree of precision (Schochet 2008). Consequently, evaluations of school-level interventions using an RD design commonly pool across time or space (or both) to increase statistical power. For example, the federal SIG evaluation pools across states (Dragoset et al. 2017), even though the program was administered separately by each state. Similarly, LiCalsi et al. (2015) and Zimmer, Henry,

[10] On the flip side, schools that were SIG-eligible in the first cohort—including those that received a SIG award—but fell outside the lowest five percent of schools in the second SIG cohort effectively serve as control schools for the second cohort of SIG-eligible schools. Our data show that eight schools fall into this category. To the extent that first-round SIG eligibility (and SIG awards) increased student achievement, then the inclusion of these schools in the control group may lead to an underestimate of SIG impacts for the second cohort. However, the inclusion of these schools is necessary to ensure the validity of the RD design, as they informed the ODE eligibility calculations. That said, we address the fact that these schools received SIG awards by treating them as "non-compliers" in our analysis estimating TOT effects (see Table A3 in the appendix).

and Kho (2017) pool across cohorts in their respective evaluations of Massachusetts' and Tennessee's turnaround efforts (although only the former study employs an RD design).

Our final, pooled dataset contains observations over a five-year period for all students attending a school with the potential for Tier 1 SIG eligibility on the basis of its achievement level. Specifically, the dataset contains more than 2.8 million observations from approximately 741,000 unique students attending potential Tier 1 schools, 62 of which were actually eligible in either the first or second cohort (or both).

[Insert Table 3 about here]

Table 3 presents summary statistics for all students in the dataset, as well as separately for those in schools below and above the eligibility threshold. The table demonstrates that students attending SIG-eligible schools are much more likely to be black, Hispanic, economically disadvantaged, English language learners, and disabled than their peers in schools above the Tier 1 eligibility cutoff. They are also less likely to be classified as a native English speaker or gifted, and they exhibit much lower achievement scores than students in SIG-ineligible schools. Finally, Table 3 illustrates that the average student below the Tier 1 eligibility threshold attends a school with a lower average enrollment, attendance rate, Performance Index score, and estimated value-added, compared to the school of the average student above the SIG eligibility cutoff.

Validity of the Design

The first step in establishing the validity of our RD design involves demonstrating a difference in the likelihood of SIG receipt for schools on either side of the eligibility threshold. Figure 1 provides such a demonstration. The markers in the figure represent the mean proportion of students attending schools that received a SIG award in defined bins below and above the eligibility threshold. The line is a smoothed, kernel-weighted mean fit separately on each side of the cutoff.[11] The figure shows that more than 75 percent of students just below the eligibility threshold attended a school that received a SIG award while only about five percent of students above the threshold were enrolled in a school that received SIG funds.[12]

[Insert Figure 1 about here]

Figure 1 illustrates a substantial discontinuity at the eligibility threshold in the probability of attending a school that received a SIG award, but the validity of the RD design could be threatened if schools were able to manipulate their performance on the criterion determining SIG eligibility—the combined proficiency rate. We believe manipulation of this metric to be unlikely for two main reasons. First, the combined proficiency rate is based on data that extends five years into the past, meaning that much of the data informing SIG eligibility were finalized long before schools were aware they would be used for that purpose. Second, the threshold determining SIG eligibility is a relative one, rather than an absolute one. Consequently, in the unlikely event that a school wanted to manipulate its performance, it would not have a clear threshold to target. Rather, the school's eligibility would depend on the simultaneous performance of other schools, which is unobservable to the potential manipulating school. Taken together, these conditions make manipulation by schools highly unlikely.

[Insert Figure 2 about here]

Nonetheless, we conducted tests designed to detect potential manipulation. First, we simply plotted the density of cases around the eligibility threshold to assess whether there is a disproportionate stacking of students on either side of the cutoff, which would be evidence of

[11] The bins in the figure are two units in width. The mean smoothing was performed using an epanechnikov kernel and a bandwidth of two. A parametric approach to estimating the discontinuity yields similar results.
[12] Parametric analysis estimates the change in probability of SIG receipt at the threshold to be between 0.7 and 0.8, depending upon the specification. All estimates are statistically significant at $p<0.01$.

potential manipulation. Figure 2 presents these plots. The left-hand panel of the figure provides a density plot on the basis of student-level data while the right-hand panel is based on school-level data. Neither panel provides evidence of disproportionate stacking around the cutoff. This visual evidence is corroborated by the results of the statistical test proposed by McCrary (2008), which fails to reject the null hypothesis of no change in the density of cases at the threshold.[13]

Although these tests provide strong evidence that there is no disproportionate stacking of cases at the threshold, it is still possible that the characteristics of students—and their corresponding schools—just below the SIG eligibility threshold are systematically different from those just above the eligibility cutoff. Any such differences would threaten the validity of the RD design. To assess the likelihood of such differences we systematically compare the observable, pre-treatment characteristics of students on either side of the eligibility threshold. We perform this comparison by estimating:

$$O_{is} = f(P_s) + \tau G_s + \varepsilon_{is} \tag{1}$$

where *O* represents an observable characteristic of student *i* in school-cohort combination *s* during the year of SIG eligibility identification. The observable characteristics are modeled as a flexible function of a school's distance from the SIG eligibility threshold $f(P_s)$, an indicator for SIG eligibility G, and an error term ε. We estimate this model over 23 separate measures that encompass students' observable demographic, achievement, and school characteristics.[14] Below we provide results from two specifications of $f(P_s)$: 1) A linear term interacted with the SIG

[13] When conducted across the full range of the running variable the coefficient returned by the test is 0.210 with a standard error of 0.241. When the test is conducted over cases within 20 points of the threshold—the bandwidth we use in our analyses below—the coefficient is 0.057 with a standard error of 0.335.

[14] Specifically, we estimate the model over the following characteristics: race/ethnicity, sex, economic disadvantage, English language learner, native English speaker, gifted, disabled, grade, reading achievement score, math achievement score, the school's Performance Index score, a charter school indicator, school enrollment, school attendance rate, the school's value-added score (separately in reading, math, and overall), the mean years of experience for teachers in the school, the percent of teachers with master's degrees in the school, and the mean salary of teachers in the school.

eligibility indicator and 2) linear and quadratic terms, each of which is interacted with the SIG eligibility indicator. We estimate the model using observations from all students attending schools within 20 points of the SIG eligibility threshold, although we obtain substantively similar results from models estimated using alternative bandwidths.[15]

[Insert Table 4 about here]

Table 4 presents the coefficient estimates and standard errors for the indicator of SIG eligibility for each observable characteristic we analyze. The results reveal no significant differences at the eligibility cutoff for any of the 23 observable characteristics when the running variable is specified as a linear term interacted with the SIG eligibility indicator. The second specification returns one estimate that is significant at $p<0.10$, but none that are significant at $p<0.05$. Considered together, the results in this section provide confidence in the validity of our RD design and thus the causal nature of the estimates in our analyses below.

Student Achievement Analysis

Leveraging our RD design, we first estimate the effect of SIG eligibility on student achievement by estimating a series of OLS models, one for each of the four years following SIG identification. The general model specification for each year is:

$$Y_{is} = f(P_s) + \tau G_s + \beta X_s + \varepsilon_{is} \quad (2)$$

where Y is achievement in reading or math for student i in school-cohort combination s. In this model $f(P_s)$ is a flexible function of the distance from the SIG eligibility threshold (i.e. the running variable described above), G is an indicator for being identified as SIG-eligible, X is a school's average achievement score in the year of SIG eligibility identification, and ε is an error term.

[15] Bandwidth selection was informed by results of the Imbens and Kalyanaraman (2012) procedure.

The parameter of interest in this model is τ, which represents the causal effect of SIG eligibility on student achievement. The estimates of τ will be unbiased as long as $f(P_s)$ is properly specified. Below we present results from a model where we specify $f(P_s)$ as a linear term interacted with the SIG eligibility indicator. In the appendix, we present similar results using a quadratic polynomial. We estimate separate models for reading and math over samples that include all students attending a school within 20 points of the SIG eligibility threshold. In robustness checks we demonstrate that substantively similar results are obtained from estimating this model over samples of alternative bandwidths. In all models we cluster standard errors by school.

We supplement the estimated effects of SIG eligibility from the series of cross-sectional models above with an analysis based on a dynamic RD technique proposed by Cellini, Ferreira, and Rothstein (2010). This technique provides the potential for greater efficiency through simultaneous estimation of the dynamic treatment effects of SIG eligibility. We implement this approach through the following regression model:

$$Y_{iskt} = f(P_s, \gamma_k) + G_s\tau_k + \alpha_k + \pi_s + \lambda_t + \varepsilon_{iskt} \tag{3}$$

where *Y* is achievement in reading or math for student *i* in school-cohort combination *s* in year-relative-to-SIG-identification *k* during calendar year *t*. Students' achievement outcomes are modeled as a flexible function of their school's distance from the SIG identification threshold—with the coefficients allowed to vary by *k*—as well as an indicator for attending a school scoring below the eligibility threshold at *k*=0. The coefficients on this indicator, represented by τ in equation (4), are the estimated effects of SIG eligibility and are also allowed to vary by *k*. The terms α_k, π_s, and λ_t are fixed effects for year-relative-to-SIG-identification, school-cohort, and calendar year, respectively. ε_{iskt} represents the error term. We again estimate separate models for reading and math over samples that include student-level observations from k=0 through k=4 for students attending a school that was within 20 points of the SIG eligibility threshold at k=0. We specify $f(P_s, \gamma_k)$ using the same functional form employed in our cross-sectional analysis. Once again, we cluster standard errors by school in all models.

[Insert Table 5 and Figures 3-4 about here]

Table 5 presents results from estimating equations (2) and (3). The top panel of the table presents results from the series of cross-sectional models while the bottom panel presents results from the dynamic models. The left-hand side of the table presents estimates of the effect of SIG eligibility on reading achievement while the right-hand side presents estimated effects on math scores. The results from the top panel of Table 5 are presented visually in Figures 3 (reading) and 4 (math). Separately for each year, these figures plot mean achievement by the distance from the eligibility threshold. Each plot also contains a separate line of best fit on each side of the eligibility threshold.

The results in the top panel of Table 5 show that SIG eligibility has no effect on achievement in the first year after eligibility identification. The point estimates are approximately 0.05 standard deviations in magnitude and insignificant across both subjects. Estimates for the second year after identification, however, are much larger in magnitude. In reading, SIG eligibility is estimated to have an effect of 0.15-0.20 standard deviations, depending upon the modeling approach. The corresponding estimates for math are 0.17-0.21 standard deviations. The estimates are statistically significant at p<0.10 or lower across both subjects and modeling approaches.

The third-year estimates are even larger than those from the second year. In reading, the estimated effects of SIG eligibility across the two modeling approaches are 0.22 and 0.27, respectively, and each is statistically significant at $p<0.10$. The analogous math effects are 0.22 and 0.26, with the former estimate significant at $p<0.10$ and the latter at $p<0.05$. The fourth-year estimates are somewhat smaller than those from the third year, particularly in the cross-sectional models. In these models, SIG eligibility is estimated to increase fourth-year reading and math scores by 0.11 and 0.09 standard deviations, respectively, which represent meaningful declines from the third-year estimates of 0.22 standard deviations. The fourth-year estimates from the dynamic model are larger in magnitude than their cross-sectional analogs. Specifically, the estimated fourth-year effects of SIG eligibility on reading and math achievement are 0.22 and 0.17 standard deviations, respectively. These estimates are only slightly smaller than the third-year estimates. Below we show that a non-parametric approach returns fourth-year estimates consistent with those from the dynamic model, which indicates that the cross-sectional estimates are the outliers. There are differences between the cross-sectional and dynamic modeling strategies that likely account for the differences between the two sets of estimates. Most notably, the dynamic modeling strategy estimates all yearly treatment effects simultaneously using a single regression, whereas the cross-sectional approach estimates these yearly effects with a series of regressions using subsets of the data. Thus, although both approaches entail estimating gains using the same pre-treatment achievement baseline, the gain estimates in the dynamic model are adjusted for differences in average achievement gains across years.

Considered together, the results in Table 5 provide strong evidence that SIG eligibility had a statistically significant—and substantively large—effect on student achievement in the second and third years after SIG eligibility identification, and likely the fourth as well. Analyses in the Appendix show that our results are insensitive to alternative specifications of the running variable (see Table A1 and Figures A1-A2) and to different bandwidths of the running variable (see Table A2). We also show that estimates of the treatment-on-the-treated (TOT) parameter are only slightly larger than the intention-to-treat (ITT) estimates presented above (see Table A3), which is unsurprising given the large proportion of eligible schools receiving SIG awards. Finally, in the appendix we show that the estimated effects of SIG eligibility on annual student achievement gains are consistent with the estimated effects on student achievement levels shown in Table 5 (see Table A4).

In addition to estimating the effects of SIG eligibility using an RD design and the parametric models presented in equations (2) and (3), we also estimate effects using the nonparametric estimator—along with the associated optimal bandwidth selection and bias-corrected inference procedures—proposed by Calonico, Cattaneo, and Titiunik (2014a; 2014b) and Calonico et al. (2017). Specifically, we estimate equation (2) with kernel-based local linear regressions, estimated separately with a triangular kernel on each side of the SIG eligibility threshold. We address the fact that standard bandwidth selectors may lead to downward bias in confidence intervals for the estimated treatment effects—and thus over-rejection of the null hypothesis of no effect—by using the bias-corrected confidence intervals proposed in Calonico, Cattaneo, and Titiunik (2014a). We account for the clustering in our running variable during the bandwidth selection process (see Bartalotti and Brummet 2017) by using the technique detailed in Calonico et al. (2017).

Table 6 presents the results of this analysis, and they are quite similar to the parametric estimates presented in Table 5. In particular, they show no significant effects of SIG eligibility in the first year after eligibility identification. However, meaningful positive estimates begin to emerge in the second year after eligibility identification, with reading estimates of 0.11-0.12 standard deviations and math estimates of 0.14-0.15 standard deviations, although the estimates are relatively imprecise and do not reach a conventional level of statistical significance. The third-year estimates are even larger than those from the second year, with estimates of about 0.20 standard deviations in both reading and math. The fourth-year estimates are comparable to, if slightly smaller than, the third-year estimates—SIG eligibility increases scores by 0.17-0.18 standard deviations in each subject. As we note above, the fourth-year estimates from this approach are closer in magnitude to those from the dynamic models in Table 5. Taken together, the results provide evidence that the positive effects of SIG eligibility persist throughout all four years following SIG eligibility identification. In the appendix, we show that school-level estimates of the effects of SIG eligibility are remarkably similar in magnitude to the student-level estimates presented in Table 6, and oftentimes more precise (see Table A5).[16]

[Insert Table 6 about here]

Turnaround vs. Transformation Model

Earlier we noted that schools receiving a SIG award were required to implement one of four federally-approved turnaround models: Closure, Restart, Transformation, or Turnaround. Table 2 demonstrates that nearly every school implemented either the Turnaround or Transformation model. To gain insight into whether the positive achievement effects seen in Tables 5 and 6 are primarily attributable to one particular model, or whether they both produce positive effects, we estimate variants of equations (2) and (3) where we interact indicators for implementing the SIG Turnaround and Transformation models with the indicator for SIG eligibility. We stress that the results of this analysis cannot be interpreted as estimates of the causal effect of implementing the Turnaround or Transformation model. Ohio schools and districts worked collaboratively to

[16] In the appendix we further show that the effect of SIG eligibility on school value-added—a measure of average annual achievement gains—are of a magnitude consistent with the effects of SIG eligibility on average achievement levels presented in Table 6 (see Tables A6-A7 for estimates from the dynamic and cross-sectional modeling approaches, respectively).

determine which SIG model a school would implement. This collaboration was designed to identify the optimal SIG model for a given school, but it also renders adoption of these models endogenous. Consequently, the results should be interpreted as mean achievement differences between students attending SIG-ineligible schools and those attending schools that respectively implemented the Turnaround or Transformation model, conditional on schools' distance from the SIG eligibility threshold.

[Insert Table 7 about here]

Table 7 presents the results of this analysis. The results make clear that—compared to SIG-ineligible schools—schools implementing the Turnaround model exhibited substantial achievement gains beginning in the second year after SIG eligibility identification. In both reading and math and across both the cross-sectional and dynamic models, the gains associated with implementation of the Turnaround model are about 0.16-0.22 standard deviations in the second year and range from 0.21-0.30 standard deviations in the third and fourth years after eligibility identification. Schools that implemented the Transformation model also exhibited achievement gains, but they are smaller in magnitude and uniformly insignificant. In the appendix we demonstrate that a difference-in-differences approach produces similar conclusions to those presented in Table 7, specifically that implementation of the Turnaround model is associated with larger achievement gains than implementation of the Transformation model (see Table A8). Together, the results provide evidence that the SIG Turnaround generated larger achievement increases than the Transformation model, although we cannot rule out the possibility that these differences are attributable to selection.

Changes in Student Composition

The SIG program was intended to improve student outcomes by increasing school effectiveness. It is possible, though, that the positive effects of SIG presented in Tables 5-6 were driven by changes in student composition at SIG-eligible schools. Relatively high-achieving students may have moved into SIG-eligible schools and produced the positive achievement effects presented above. To test for this possibility we first constructed an annual student-level measure indicating whether the student attended the same school as she did the previous year. Then, using the RD design described above, we specify this measure as the outcome in models identical in structure to the one presented in equation (2). Negative and significant coefficients on the measure of SIG eligibility would indicate that SIG-eligible schools enroll a greater proportion of new students than schools just above the eligibility threshold. Such a result would be consistent with an influx of higher achieving students into SIG-eligible schools that could drive the achievement effects shown in Tables 5-6.

Table A9 in the appendix shows, however, that the coefficient on the measure of SIG eligibility is actually positive and significant in the year of eligibility identification. In other words, SIG eligibility increases the likelihood that a student remains enrolled in the same school during the first year of SIG by about 7 percentage points. The estimates for the second through fourth years following SIG eligibility identification are also positive and range from 3-6 percentage points, depending upon the particular year, but not all are significant. These results hold when we estimate a variant of the model that excludes students attending the lowest grade of a school, as these students are highly unlikely to have attended the same school the prior year.

The significant positive effect of SIG eligibility on the probability of remaining enrolled in the same school leaves open the possibility that these schools disproportionately retained their high-achieving students. To provide evidence on this possibility we estimated a variant of the model described above where we interact the SIG eligibility indicator with a student's reading achievement score. A positive and significant coefficient on this interaction would indicate disproportionate retention of high-achieving students in SIG-eligible schools. The results in Table A10 in the appendix, however, show the coefficient on this interaction to be negative and significant in the first two years following SIG eligibility identification and insignificant thereafter. To address the possibility that the retention of higher performing students is an artifact of the positive achievement effects of SIG, the third column of the table presents results from a specification where achievement is fixed at a student's achievement level in the final pretreatment year. The coefficients on the interaction remain negative and significant in this specification, and even somewhat larger in magnitude. This provides strong evidence that low-achieving students were disproportionately retained in SIG-eligible schools in the years following eligibility identification. These results, together with the estimated effects on student test score gains and school value-added (see Appendix Tables A6-A7), suggest that the positive achievement effects of SIG eligibility were driven by true school improvement, as opposed to a changing student population.

School Administration and Staffing Analysis

School Finance

Schools identified as SIG recipients were awarded grants that ranged from $75,000 to $2 million to be spent over a three-year period. These grants were intended to increase expenditure levels at the schools receiving SIG awards. However, because grant funds generally flowed through district offices, it is possible that districts managed finances in a manner such that the grants supplanted, rather than supplemented, existing school spending. Dragoset et al. (2017) present evidence that such a scenario may have been occurring in the early years of SIG.

[Insert Table 8 about here]

We take advantage of the fact that Ohio maintains school-level expenditure data to perform analyses intended to provide insight into whether SIG grants to Ohio schools supplemented or supplanted existing expenditures. Specifically, we employ the RD design described above, coupled with the data on school expenditures, to estimate the effect of SIG eligibility on per-pupil expenditures.[17] We complement the RD estimates with results from a difference-in-differences model. This model contains fixed effects for school-by-cohort, calendar year, and year relative to SIG identification. Interactions between a SIG eligibility indicator and the indicators for year relative to SIG identification represent the difference-in-differences estimates of the effect of SIG eligibility—see the appendix for more detail on the difference-in-differences approach. Table 8 presents the results of these analyses. They illustrate that SIG eligibility has no significant effect on per-pupil expenditure levels in the year following SIG eligibility identification. That is to be expected, as the three-year grants did not begin until the following year. The results also indicate that SIG eligibility increased per-pupil expenditures by about $1,500-2,000 in the second year after eligibility identification, by $2,600-3,300 in the third year, and by $1,200-1,600 in the fourth year, which corresponds to the final year of the grant period.

These estimates are generally consistent with the size of the typical SIG award in Ohio. Indeed, award data provided by ODE indicate that the average award over the three-year grant period was about $2.25 million, which averages out to about $750,000 annually. When spread across the enrollment in SIG schools, the average award comes out to about $2,300 per student annually. However, the ODE data show SIG schools' award budgets propose spending a greater

[17] The models based on the RD design control for distance from the eligibility threshold using a linear term that is interacted with the indicator for scoring below the SIG eligibility threshold. The cross-sectional RD models also contain a covariate measuring per-pupil expenditure in the year of SIG eligibility identification.

portion of the award in the first and second years, compared to the third year of the grant. Specifically, the data indicate that the average school proposed spending approximately $2,600 per pupil in the first year of the award, $2,200 in the second, and $2,000 in the final year of the award. These annual averages do not perfectly correspond to the estimates in Table 8, but there are multiple reasons we might expect imperfect alignment. First, the averages described here are based on schools' budgeted award expenditures rather than actual award expenditures, and these two figures do not always align. Unfortunately, we do not have data on actual school-level SIG expenditures. Second, our RD approach estimates the effect of SIG eligibility at the threshold whereas the proposed expenditures represent averages across all awarded schools, not just those at the threshold. Third, some districts account for money disbursed near the end of a fiscal year on the budget for the following year.

Although we cannot completely rule out the possibility that some funds were supplanted, the results in Table 8 provide evidence that the SIG grants awarded to Ohio schools supplemented existing expenditures. In addition, the results illustrate that the magnitude of these supplements was substantial. Our data indicate that the average SIG-eligible school spent approximately $14,100 per pupil in the year of SIG eligibility identification. Compared against this base, the estimated effects correspond to spending increases of 11-14%, 18-23%, and 9-11% in the second, third, and fourth years after SIG eligibility identification, respectively.[18]

[Insert Table 9 about here]

[18] Ideally, we would be able to examine how schools spent these additional funds, exploring, for example, whether the dollars were directed toward instructional activities, student services, facilities, or other expenditure categories. Unfortunately, Ohio does not maintain school-level expenditure records with that degree of detail. The fact that Ohio maintained school-level expenditures at all during this time period is relatively unique among states—it was not until the 2015 passage of the Every Student Succeeds Act that states and districts were required to publicly report school-level expenditures.

Table 9 presents results from specifications of the RD and difference-in-differences models where we interact indicators for implementing the SIG Turnaround and Transformation models with the indicator for SIG eligibility. Across all three modeling approaches, results for schools that implemented the SIG Transformation model indicate positive and statistically significant per-pupil spending increases in the second, third, and fourth years following SIG eligibility. In contrast, only one of the twelve estimates for schools that implemented the Turnaround model are statistically significant, although several of the non-significant estimates are of a meaningful magnitude—in the range of $1,500-$3,000. The lack of significance of the Turnaround estimates is at least partially attributable to the fact that only 15 schools implemented this SIG model, which results in large standard errors. Together, the results in Table 9 provide suggestive evidence that both the Turnaround and Transformation models are associated with per-pupil spending increases, although only the estimates for the Transformation model are significant.

A natural question that emerges from the results in Table 9 is why the Transformation model entails significant spending increases but the Turnaround model does not. The requirements of each model may provide at least a partial answer to this question. Both models required a change in school leadership and a number of instructional and operational reforms. However, the Turnaround model further entailed replacing at least 50 percent of the school staff—a requirement absent from the Transformation model—while the Transformation model required schools to adopt a new governance model and an educator evaluation system that included student achievement growth as a significant component. These two components of the Transformation model were not mandated by the Turnaround model. It seems likely that the requirements of the Transformation model, particularly development of a new governance model and evaluation system, were costlier endeavors than the Turnaround mandate of staff turnover. This is particularly likely to be true if schools implementing the Turnaround model replaced terminated staff with less experienced individuals earning lower salaries.

School Staffing

In the prior section, we described how the SIG turnaround models were designed to change leadership and staffing in schools receiving SIG awards. To empirically examine how Ohio's SIG initiative affected school staffing we perform a series of analyses based on the same RD design described above. The first analysis in this vein estimates the effect of SIG eligibility on staff turnover. Specifically, using staff-level data, we estimate the effect of SIG eligibility on the probability of staff members remaining employed at the same school three and four years, respectively, after SIG eligibility identification. We estimate this effect using a variant of equation (2) where we specify an indicator of staff turnover as the outcome. As with the RD analyses in prior sections, we present results from a model where the running variable is specified as a linear term that is interacted with the SIG eligibility indicator. In the appendix we show that other specifications return substantively similar results (see Table A11 and Figure A3). We estimate the model separately for all school staff and only for teachers. We cluster standard errors by school in all models.

[Insert Table 10 about here and Figure 5 about here]

The results of this analysis are presented in the top two panels of Table 10, with corresponding visual representation in Figure 5. Table 10 shows that SIG eligibility had no significant effect on the stability of school staffing, with the "all staff" estimates very close to zero. For teachers only, the results indicate that SIG eligibility is estimated to increase the probability of being employed at the same school three and four years after potential SIG identification by 0.05, but these estimates are quite imprecise and do not approach statistical significance. Inspection of Figure 5 suggests that the null effect of SIG eligibility on staffing stability is due to high levels of staff turnover on both sides of the SIG eligibility cutoff, rather than noncompliance with the requirements of the SIG models. On average, only about 30 percent of staff at a school just above the SIG eligibility cutoff are still employed at that school four years later.

The bottom panel of Table 10 presents results from an analysis of the effect of SIG eligibility on annual staff turnover for each of the first four years following potential SIG eligibility identification. The results show that SIG eligibility had no significant effect on annual turnover of either teachers or school staff more generally. The results suggest that SIG eligibility may actually decrease teacher turnover in the years following SIG eligibility identification—three of the four point estimates are negative—but none of the estimates are sufficiently precise to reach statistical significance. Taken together, the results in Table 10 indicate that SIG had no meaningful effect of staff turnover rates, with additional analysis suggesting that the null effect is attributable to similarly high turnover rates at schools on either side of the SIG eligibility threshold.

[Insert Table 11 about here]

It is possible that the null effects presented in Table 10 mask heterogeneity across the two main SIG models, Turnaround and Transformation. Table 11 presents estimates from a specification where we interact indicators for implementing the SIG Turnaround or Transformation models with the indicator for SIG eligibility. The results provide evidence of significant heterogeneity, particularly in the case of teachers. Teachers in schools that implemented the Transformation model were 14 percentage points more likely to remain employed at the same school three years later, compared to teachers at SIG-ineligible schools. In contrast, teachers in schools that implemented the Turnaround model were about 15 percentage points less likely to remain employed at the same school three years down the line. Although these estimates cannot be interpreted causally—the decision to implement a particular SIG model is likely endogenous—these results provide important evidence that the effects of SIG on school staffing played out quite differently across the two primary SIG models.

SIG eligibility had no overall effect on staff turnover, but it is possible that it could have affected staffing levels. For example, schools and districts could use SIG funds to hire additional educational staff, which would not be reflected in turnover rates. To examine this possibility, we used the staff-level data described above to generate annual school-level counts of four staffing categories—all staff, teachers, education professionals, and administrators. We then use the RD design to estimate the effect of SIG eligibility on staffing levels across these four categories. We estimate these effects using both the cross-sectional and dynamic modeling approaches we have employed throughout the analysis. We specify the models as we did for the analysis of staff turnover, although the results are robust to alternative analytic decisions (see Table A12 in the appendix).[19]

[Insert Table 12 about here]

The results presented in Table 12 show that SIG eligibility generates a significant reduction in the number of school administrators in the first year of SIG eligibility. The number of administrators is estimated to decline by 1-1.5 FTE in that year, depending upon the modeling approach. This is perhaps unsurprising given that both the SIG Turnaround and Transformation

[19] In particular, we specify the running variable as a linear term interacted with the SIG eligibility indicator, and estimate the model over all schools within 20 points of the SIG eligibility threshold in the year of SIG eligibility identification. We cluster standard errors by school in all models.

models require a change in school leadership, and it may take some time to fully reconstitute the leadership team. There is some evidence of SIG eligibility reducing the total number of staff in a school by about 4-6 FTEs in the year following SIG identification, but only the estimate from the dynamic RD model approaches significance ($p<0.10$). Table 12 shows that the point estimates for teachers and educational professionals are positive and of a potentially meaningful magnitude, particularly in the second and third years where SIG eligibility is estimated to increase the number of FTEs in each category by 1.5-3. However, these estimated effects are quite imprecise and do not approach statistical significance. In addition, they tail off in the fourth year after SIG eligibility identification.

Table A13 in the appendix provides evidence that any increases in the number of teachers or education professionals come from schools that implemented the Transformation model. The point estimates for schools implementing the Transformation model are positive and significant for the second and third years following SIG eligibility identification whereas the point estimates for the Turnaround model in those years are negative, albeit insignificant. These results are consistent with the Transformation vs. Turnaround turnover results presented in Table 11 above.

School Closure

Our final analysis estimates the effect of SIG eligibility on the probability of school closure. There are at least two ways in which SIG eligibility could result in school closure. First, and most obviously, SIG recipients could implement the Closure or Restart turnaround models, which are two of the four federally-approved strategies that SIG awardees could implement. However, Table 2 demonstrated that almost no school elected to implement either of these models. Second, schools that are SIG-eligible but not awarded SIG funds may be closed due to lack of assistance or support for improving the quality of a school. We note that a similar phenomenon could occur with schools just above the SIG eligibility threshold—districts may decide to close these low-performing schools as they are ineligible for SIG awards that could facilitate improvement. In addition to being an interesting outcome in its own right, this closure analysis has potential implications for the interpretation of the estimated effects of SIG on student achievement. In particular, if SIG eligibility is found to have an effect—either positive or negative—on the probability of closure, then it is possible that the large positive effects of SIG eligibility on student achievement are at least partially attributable to differential sample attrition.

[Insert Table 13 about here]

We estimate the effect of SIG eligibility on school closure separately for each of the four years following SIG eligibility identification. As with the analysis of school staffing levels, we use school-level data and employ both the cross-sectional and dynamic modeling approaches, specifying the sample and models in the same way we did above. Table 13 presents the estimated effects of SIG eligibility on closure. The results illustrate that SIG eligibility has no significant effects on the probability of closure—across all years and both modeling approaches the point estimates are insignificant. In addition to demonstrating that SIG eligibility has no effect on the probability of closure, the results in Table 13 provide evidence that the large positive effects of SIG eligibility on student achievement are not attributable to differential attrition between schools that are SIG eligible and those that are not. Further analysis demonstrated that these null effects hold across both the Turnaround and Transformation models.[20]

Considered together, the results in this section indicate that SIG eligibility led to a significant increase in per-pupil expenditures, with these increased expenditures largest in schools that implemented the Transformation model. Interestingly, there was no overall effect of

[20] These results are available upon request.

SIG eligibility on school staffing—either staff turnover or staffing levels—but we provide evidence that this null mean effect masked significant heterogeneity in staffing outcomes across the two primary SIG models, Turnaround and Transformation.

Discussion and Conclusion

The federal SIG program is the latest in a long line of initiatives intended to improve the quality of low-performing schools. SIG was designed to achieve such improvements by providing schools with additional financial resources and by requiring them to make significant changes to several aspects of their educational delivery, most notably their leadership and staffing. In this paper, we begin to evaluate whether Ohio's school turnaround efforts under SIG produced the intended quality improvements. In particular, we use data from two SIG cohorts to estimate the effect of SIG eligibility on student achievement levels, as well as on school spending, staff turnover, and school staffing levels. We estimate these effects in an RD framework, leveraging the criteria stating that schools were eligible to receive a SIG award if they ranked in the bottom 5 percent of eligible schools in terms of a weighted student proficiency rate calculated by ODE.

The results of our analysis indicate that SIG eligibility had a large positive effect on the achievement of students attending those schools. These effects first emerge in the second year after a school is identified as SIG-eligible, with reading and math effects up to 0.20 standard deviations. In the third year, SIG eligibility is estimated to increase reading achievement by at least 0.20 standard deviations and math achievement by at least 0.18 standard deviations, and some specifications suggest achievement gains of up to 0.26 standard deviations in math and 0.27 standard deviations in reading. There is some evidence from the cross-sectional models that the achievement gains diminish in the fourth year following SIG eligibility identification, but our other modeling strategies—the dynamic model, the non-parametric approach, and even the cross-sectional model with different bandwidths—suggest that the gains persist through the fourth year.

Although the preponderance of the evidence points toward the persistence of these positive achievement effects through the full award period, it is worth considering the possibility that the effects diminish or disappear in the final award year, or even in post-award years. The SIG program was designed to generate lasting improvements in the quality of awarded schools. Our results indicate that Ohio's SIG program got off to a good start on this score, but they provide little indication as to whether these initial quality improvements will be maintained over the long haul. However, if prior school improvement efforts are any guide, there are reasons to think that these achievement increases may ultimately prove transitory. The history of education reform is replete with examples of promising initiatives—some of which even generate positive short-term impacts—that ultimately fizzle out (Tyack and Cuban 1995; Hess 1998). Several aspects of the education policy arena, including the impatience of policymakers and funders, the desire of superintendents to leave their mark via new policy reforms, and the general pull of the status quo, often serve to limit the staying power of even promising reform efforts. Our results do not provide evidence on the long-term impacts of Ohio's SIG initiative, but they do suggest that such inquiry would be worthwhile.

Along with providing estimates of the causal effects of SIG eligibility, our results also provide evidence of potential heterogeneity across schools that implemented the Turnaround and Transformation models. Specifically, our results indicate that the positive achievement effects are primarily attributable to schools that implemented SIG's Turnaround model. With respect to spending, we estimate that SIG eligibility increases annual per-pupil funding by $1,500-$3,000 across the three-year grant period, but provide evidence that the spending increases were somewhat larger in schools that implemented the Transformation model—although, results suggest that both Turnaround and Transformation schools realized spending increases. Finally, the null effect of SIG eligibility on staff turnover masks the fact that teachers in Turnaround schools were significantly less likely to still be employed at that school three and four years later while teachers at Transformation schools were substantially more likely to remain employed at their school. Considered together, our results are consistent with a scenario where the staff replacement requirements of the Turnaround model, coupled with the increased financial resources, generated the observed achievement increases. However, the evidence is not definitive on this score.

Our findings are at odds with those of the federally-sponsored SIG evaluation, which found SIG implementation to have no effect on student achievement or attainment (Dragoset et al. 2017). The federal evaluation was based on an RD design and drew upon data from a non-representative sample of 60 districts spread across 22 states. Such an approach provides valuable evidence about the effects of SIG across a geographically diverse set of districts, but there are also potential benefits to focusing on specific contexts. Indeed, in a federal system like ours, policy is implemented and administered by states, districts, and schools. With SIG in particular, state education agencies were given meaningful autonomy within the federally-defined parameters to administer the program. For example, Ohio was given some latitude in creating the metric on which the bottom five percent of schools were deemed SIG-eligible. Similarly, the state had autonomy in designing both the call for districts to submit proposals for SIG funds and the process for selecting those schools and districts that would receive SIG awards. More generally, the state contexts in which SIG was implemented and administered undoubtedly vary significantly, with some of those contexts more conducive than others to school turnaround policies achieving the intended school improvement.

Given this variation in context—and the potential role that context may play in the ultimate success of turnaround policies—it is important for evaluations to take the issue seriously. There are several ways this can be done. One approach involves conducting separate evaluations across distinct contexts (i.e. states) and then attempting to identify the contextual factors associated with policy success or failure after a sufficient number of evaluations have been conducted. The charter school literature has evolved in this manner, with a number of earlier studies devoted to evaluating the charter sectors in various states or cities (e.g. Witte et al. 2007; Bifulco and Ladd 2006; Zimmer et al.. 2009; Abdulkadiroğlu et al. 2011) and later work placing more emphasis on identifying the factors or contexts that facilitate successful charter schools (e.g. Angrist, Pathak, and Walters 2013; Dobbie and Fryer 2013; Fryer 2014). There are signs that the recent school turnaround literature may take a similar path. The literature now contains studies evaluating school turnaround efforts in a number of jurisdictions, including California (Dee 2012), North Carolina (Heissel and Ladd 2016; Henry and Guthrie 2015), Massachusetts (Papay 2015; LiCalsi et al. 2015), Ohio, Tennessee (Zimmer, Henry, and Kho 2017), and Los Angeles (Strunk, McEachin, and Westover 2014). Our positive results mirror those from Massachusetts (Papay 2015; LiCalsi et al. 2015) and California (Dee 2012) but are somewhat at odds with the mixed results from North Carolina (Heissel and Ladd 2016) and Tennessee (Zimmer, Henry, and Kho 2017). As this set of studies grows, it seems likely that future work will begin to systematically consider the contexts and conditions under which school turnaround strategies are and are not likely to prove effective.

This study provides evidence that school turnaround efforts can be successful in the context of Ohio, but does less to pin down the particular aspects of the context that facilitate success. Future work would do well to further analyze potential mechanisms that could contribute to the success of Ohio's SIG initiative, particularly given the effects of SIG on per-pupil spending levels and the differences between schools that implemented the Transformation and Turnaround models. When making any such assessments, however, it is important to recognize the localized nature of our estimated effects. In particular, our estimates apply to schools just below the SIG eligibility threshold—this analysis does not generalize to schools further away from the cutoff. We highlight two factors to consider in making any such judgment. On one hand, schools far away from the cutoff may lack the capacity for quality improvements that schools more proximate to the cutoff might possess. Alternatively, schools far from the cutoff may have greater room for improvement and thus may benefit from turnaround efforts to an even greater degree than schools just below the eligibility threshold. Regardless of whether our results generalize to SIG-eligible schools further down the distribution, our findings suggest that schools further up the distribution of ODE's "combined proficiency rate" might also benefit from turnaround interventions.

Finally, our results highlight the SIG program's blunt requirement that states identify the lowest five percent of schools in the state as SIG-eligible. The quality of a fifth percentile school has the potential to meaningfully vary across states—the quality of a school at this point of Massachusetts' distribution is likely to substantively differ from the quality of a school at the same point of Mississippi's—and perhaps contribute to heterogeneity in the effects of turnaround evaluations, particularly those that rely on an RD design.[21] Such considerations suggest the

[21] Evaluation of the effects of Ohio's Priority school designation, which were required under waivers issued from NCLB, provides evidence consistent with such heterogeneity (Carlson and Lavertu 2016b). The cutpoint for Priority

potential value of working to identify an empirically-driven approach applicable across contexts to identifying schools that could benefit from turnaround interventions. More generally, efforts to improve the quality of low-achieving schools are unlikely to abate in the future, and research on school turnarounds could help guide and improve such efforts.

designation was much higher than that for SIG and the evaluation found Priority designation to have no achievement effects, which contrasts with the large positive effects of SIG. That said, the interventions stemming from Priority designation differed from those under SIG. Perhaps most notable is that Priority designation did not result in any funding increases.

References

Abdulkadiroğlu, A., Angrist, J.D., Hull, P.D., Pathak, P.A. (2016). Charters without lotteries: Testing takeovers in New Orleans and Boston. *American Economic Review, 106*, 1878-1920.

Abdulkadiroğlu, A., Angrist, J. D., Dynarski, S. M., Kane, T. J., & Pathak, P. A. (2011). Accountability and flexibility in public schools: Evidence from Boston's charters and pilots. *The Quarterly Journal of Economics*, *126*(2), 699-748.

Adnot, M., Dee, T., Katz, V., & Wyckoff, J. (2016). Teacher Turnover, Teacher Quality, and Student Achievement in DCPS. *Educational Evaluation and Policy Analysis*, first published on August 29, 2016 as doi:10.3102/0162373716663646

Angrist, J. D., Pathak, P. A., & Walters, C. R. (2013). Explaining charter school effectiveness. *American Economic Journal: Applied Economics*, *5*(4), 1-27.

Atteberry, A., Loeb, S., Wickoff, J. (2016). Teacher churning: Reassignment rates and implications for student achievement. *Educational Evaluation and Policy Analysis*

Bartalotti, O., & Brummet, Q. (2017). “Regression Discontinuity Designs with Clustered Data.” In *Regression Discontinuity Designs: Theory and Applications* (Advances in Econometrics, Volume 38), ed. M. D. Cattaneo and J. C. Escanciano. Emerald Group Publishing, 383-420.

Bifulco, R., Duncombe, W., & Yinger, J. (2005). Does whole-school reform boost student performance? The case of New York City. *Journal of Policy Analysis and Management*, 24(1), 47-72.

Bifulco, R., & Ladd, H. F. (2006). The impacts of charter schools on student achievement: Evidence from North Carolina. *Education Finance and Policy*, 1(1), 50-90.

Bloom, H.S., Richburg-Hayes, L., & Black, A.R. (2007). Using covariates to improve precision for studies that randomize schools to evaluate educational interventions. *Educational Evaluation and Policy Analysis*, *29*, 30-59.

Borman, G.D., Hewes, G.M., Overman, L.T., & Brown, S. (2003). Comprehensive school reform and achievement: A meta-analysis. *Review of Educational Research, 73*, 125-230.

Boyd, D., Lankford, H., & Wyckoff, J. (2007). Increasing the effectiveness of teachers in low-performing schools. In H. F. Ladd & E. B. Fiske (Eds.), *Handbook of Research in Education Finance and Policy* (1st ed., pp. 612–630). Routledge.

Branch, G. F., Hanushek, E. A., & Rivkin, S. G. (2012). Estimating the effect of leaders on public sector productivity: The case of school principals. *National Bureau of Economic Research*, Working Paper No. 17803

Bross, W., Harris, D.N., & Liu, L. (2016). *The effects of performance-based school closure and charter takeover on student performance*. Technical Report. Education Research Alliance for New Orleans: New Orleans, LA

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

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014a). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295-2326.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014b). Robust data-driven inference in the regression-discontinuity design. *Stata Journal*, *14*(4), 909-946.

Calonico, S., Cattaneo, M. D., Farrell, M.H., & Titiunik, R. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, *17(2)*, 372-404.

Carlson, D., Borman, G.D., & Robinson, M. (2011). A multistate district-level cluster randomized trial of the impact of data-driven reform on reading and mathematics achievement. *Educational Evaluation and Policy Analysis, 33(3),* 378-398.

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

Carlson, D., & Lavertu, S. (2016a). Charter school closure and student achievement: Evidence from Ohio. *Journal of Urban Economics*, *95*, 31-48.

Carlson, D., & Lavertu, S. (2016b). *Ohio School Turnaround Interventions: Impacts and Achievement, Attainment, and Administration*. Ohio Education Research Center: Columbus, OH.

Cellini, S. R., Ferreira, F., & Rothstein, J. (2010). The value of school facility investments: Evidence from a dynamic regression discontinuity design. *The Quarterly Journal of Economics*, *125*(1), 215-261.

Chetty, R., Friedman, J.N., & Rockoff, J.E. (2014). Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood. *American Economic Review, 104*, 2633-2679

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2007). Teacher credentials and student achievement: Longitudinal analysis with student fixed effects. *Economics of Education Review*, *26*(6), 673–682.

Clotfelter, C. T., Ladd, H. F., & Vigdor, J. L. (2010). Teacher credentials and student achievement in high school: A cross-subject analysis with student fixed effects. *Journal of Human Resources*, *45*(3), 655–681.

Cowen, J. M., Butler, J. S., Fowles, J., Streams, M. E., & Toma, E. F. (2012). Teacher retention in Appalachian schools: Evidence from Kentucky. *Economics of Education Review*, *31*(4), 431–441.

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

Dee. T. (2012). School Turnarounds: Evidence from the 2009 Stimulus. Program on Education Policy and Governance. Working Paper Series. Retrieved from: http://www.hks.harvard.edu/pepg/PDF/Papers/PEPG12-04_Dee.pdf

Dobbie, W., & Fryer Jr, R. G. (2013). Getting beneath the veil of effective schools: Evidence from New York City. *American Economic Journal: Applied Economics*, *5*(4), 28-60.

Dragoset, L., Thomas, J., Herrmann, M., Deke, J., James-Burdumy, S., Graczewski, C., Boyle, A., Upton, R., Tanenbaum, C., & Giffin, J. (2017). *School Improvement Grants: Implementation and Effectiveness* (NCEE 2017-4013). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education.

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

Fryer, R. G. (2014). Injecting charter school best practices into traditional public schools: Evidence from field experiments. *The Quarterly Journal of Economics*, *129*(3), 1355-1407.

Grissom, J.A., Kalogrides, D., & Loeb, S. (2015). Using student test scores to measure principal performance. *Educational Evaluation and Policy Analysis,* 37(1), 3-38.

Gross, B., Booker, T.K., & Goldhaber, D. (2009). Boosting student achievement: The effect of comprehensive school reform on student achievement. *Educational Evaluation and Policy Analysis, 31*, 111-126.

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

Hedges, L., & Hedberg, E.C. (2007). Intraclass correlation values for planning group-randomized trials in education. *Educational Evaluation and Policy Analysis*, *29*, 60-87.

Heissel, J. A. and Ladd, H.F. (2016). School Turnaround in North Carolina: A Regression Discontinuity Analysis. Working Paper 156. Retrieved from http://www.caldercenter.org/publications/school-turnaround-north-carolina-regression-discontinuity-analysis on September 3, 2016.

Henry, G. T., Guthrie, J. E. (2015). Outcomes and Impacts of North Carolina's Initiative to Turn Around the Lowe-Achieving Schools. Consortium for Educational Research and Evaluation—North Carolina. Accessed October 27, 2016 from http://cerenc.org/wp-content/uploads/2015/09/ES-FINAL-Final-DST-Report-9-3-15.pdf

Herman, R., Dawson, P., Dee, T., Greene, J., Maynard, R., Redding, S., & Darwin, M. (2008). *Turning around chronically low-performing schools: IES practice guide*. Washington, DC: National Center for Education Evaluation and Regional Assistance.

Hess, F. M. (1998). *Spinning wheels: The politics of urban school reform*. Washington, DC: Brookings Institution Press.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *The Review of Economic Studies*, *79*(3), 933-959.

Jackson, C. K. (2009). Student demographics, teacher sorting, and teacher quality: Evidence from the end of school desegregation. *Journal of Labor Economics*, *27*(2), 213–256.

Jacob, R.T., Goddard, R.D., & Kim, E.S. (2014). Assessing the use of aggregate data in the evaluation of school-based interventions: Implications for evaluation research and state policy regarding public-use data. *Educational Evaluation and Policy Analysis, 36,* 44-66.

Konstantopolous, S. (2012). The impact of covariates on statistical power in cluster randomized designs: Which level matters more? *Multivariate Behavioral Research, 47,* 392-420.

Lafortune, J., Rothstein, J., & Schanzenbach, D.W. (2016). School finance reform and the distribution of student achievement. NBER Working Paper 22011

LiCalsi, C., Citkowicz, M., Friedman, L.B., & Brown, M. (2015). *Evaluation of Massachusetts Office of District and School Turnaround Assistance to Commissioner's Districts and Schools: Impact of School Redesign Grants*. Washington: DC: American Institutes for Research

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of econometrics*, *142*(2), 698-714.

Papay, J. (2015). The Effects of School Turnaround Strategies in Massachusetts. Presented at the 2015 conference of the Association of Public Policy and Management in Miami, FL.

Rivkin, S. G., Hanushek, E. A., & Kain, J. F. (2005). Teachers, schools, and academic achievement. *Econometrica*, *73*(2), 417-458.

Sass, T. R. (2006). Charter schools and student achievement in Florida. *Education Finance and Policy*, 1(1), 91-122.

Slavin, R. E. (1989). The PET and the pendulum. *Phi Delta Kappan, 70,* 752-58.

Strunk, K.O., & McEachin, A. (2014). More Than Sanctions: Closing Achievement Gaps Through California's Use of Intensive Technical Assistance. *Educational Evaluation and Policy Analysis*, 36(6), 281-306.

Strunk, K.O., McEachin, A., & Westover, T. N. (2014). The Use and Efficacy of Capacity-Building Assistance of Low-Performing Districts: The Case of California's District Assistance and Intervention Teams. *Journal of Policy Analysis and Management*, 33(3), 719-751.

Tyack, D. B., & Cuban, L. (1995). *Tinkering toward utopia*. Cambridge: Harvard University Press.

Witte, J., Weimer, D., Shober, A., & Schlomer, P. (2007). The performance of charter schools in Wisconsin. *Journal of Policy Analysis and Management*, *26*(3), 557-573.

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

Zimmer, R., Henry, G.T., & Kho, A. (2017). The Effects of School Turnaround in Tennesee's Achievement School District and Innovation Zones. *Educational Evaluation and Policy Analysis*, *39(4)*, 670-696.

Tables and Figures

Table 1. SIG Eligibility Criteria, by Tier

Tier	Ohio Eligibility Criteria
Tier 1	1. Receive Title I funds under NCLB "School improvement" process AND 2. Lowest achieving 5 percent of schools over 5 year period OR Secondary schools with 5-year graduation rate less than 60%
Tier 2	1. Secondary schools that were Title I-eligible, but not Title I recipients, under NCLB "School improvement" process AND 2. Lowest achieving 5 percent of schools over 5 year period OR Secondary schools with 5-year graduation rate less than 60%
Tier 3	1. Receive Title I funds under NCLB "School improvement" process AND 2. Not Tier 1-eligible (i.e. did not meet the achievement or graduation rate criteria)

Table 2. Ohio SIG Award Process, by Cohort

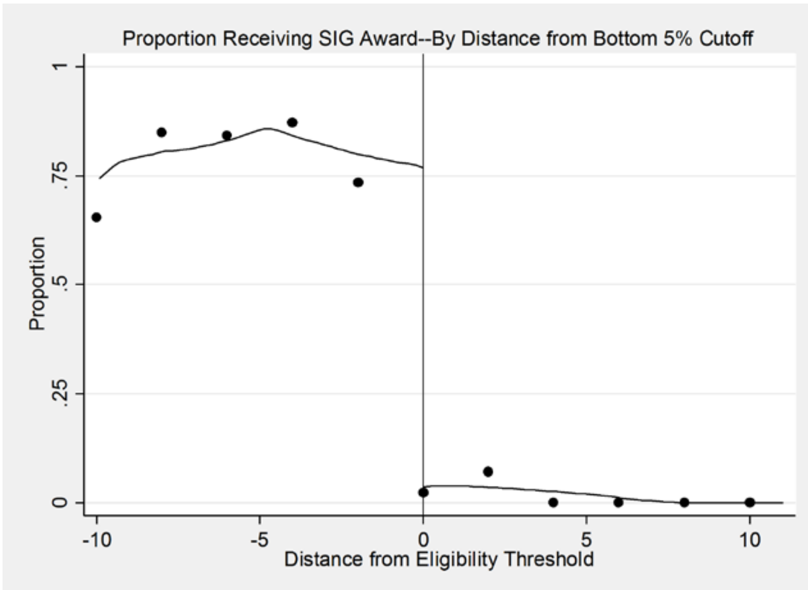
Characteristic	**Cohort 1**	**Cohort 2**	**Cohort 3**
Timing of eligibility identification	2009-10 school year	2010-11 school year	2013-14 school year
Years included in analysis	2009-10 to 2012-13 sch. yrs.	2010-11 to 2013-14 sch. yrs.	Cohort 3 not included
Timing of grant administration	2010-11 to 2012-13 sch. yrs.	2011-12 to 2013-14 sch. yrs.	2014-15 to 2016-17 sch. yrs.
Number of districts w/1+ SIG school	13	29	11
Number of SIG schools	39	44	24
Turnaround Model for Tier 1 & 2 Schools			
Closure	0	0	0
Restart	0	1	0
Turnaround	7	8	9
Transformation	25	33	14
Other	7	2	1

NOTE: The "Other" category includes Tier III schools that were not required to implement a turnaround model, as well as a small number of schools that closed shortly after identification and one school that implemented both Turnaround and Transformation models.

Table 3. Sample means, by position relative to SIG eligibility threshold

Characteristic	All observations	Observations below eligibility threshold	Observations above eligibility threshold
White	0.517	0.103	0.530
Black	0.356	0.757	0.343
Hispanic	0.055	0.089	0.054
Asian	0.009	0.004	0.009
Female	0.487	0.453	0.489
Economic disadvantage	0.707	0.925	0.700
ELL	0.036	0.066	0.035
Native English speaker	0.952	0.911	0.954
Gifted	0.083	0.024	0.085
Disabled	0.171	0.250	0.169
Reading Score (z-score)	-0.386	-1.224	-0.362
Math Score (z-score)	-0.416	-1.265	-0.392
School Performance Index	82.250	57.290	83.026
Charter school	0.142	0.296	0.137
School enrollment	861.874	379.889	876.868
School attendance rate	94.195	92.664	94.243
School reading value-added	0.313	-0.193	0.330
School math value-added	1.461	-0.433	1.525
School overall value-added	0.887	-0.313	0.927
School teacher yrs. experience	13.460	10.556	13.550
School teacher % MA	55.654	40.374	56.126
School teacher avg. salary	54,076.770	52,329.710	54,130.830
N	596,871	18,005	578,866

Figure 1. Mean Proportion of students in schools receiving SIG awards, by distance from SIG eligibility threshold



Note: Markers in the figure represent the mean proportion of schools receiving SIG awards in each bin of ODE "combined proficiency rate" measure used to determine SIG eligibility. Each bin is two units in width. The line is a smoothed, kernel-weighted mean calculated separately on each side of the cutoff.

Figure 2. Distribution of students (left panel) and schools (right panel)

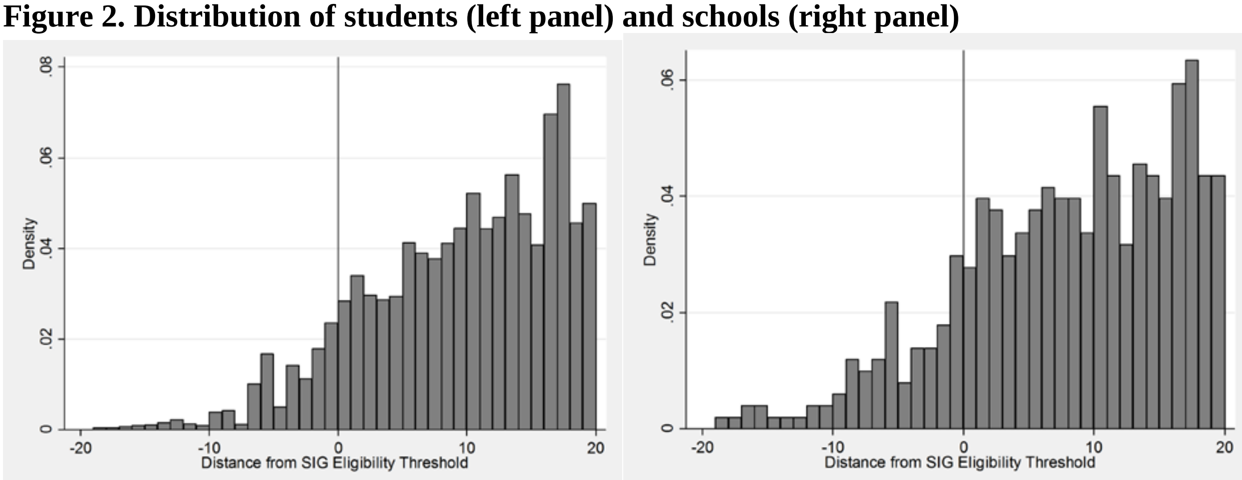


Table 4. Coefficients and standard errors on SIG eligibility indicator from OLS model predicting observable characteristics

Outcome Variable	(1) Coef. (Std. Error)	(2) Coef. (Std. Error)
White	0.024	0.057
	(0.030)	(0.042)
Black	-0.104	-0.167
	(0.100)	(0.124)
Hispanic	0.081	0.104
	(0.084)	(0.103)
Asian	0.002	-0.002
	(0.002)	(0.003)
Female	0.003	-0.005
	(0.011)	(0.015)
Economic disadvantage	-0.026	-0.037
	(0.025)	(0.035)
English language learner	0.037	0.015
	(0.058)	(0.076)
Native English	-0.074	-0.061
	(0.076)	(0.098)
Gifted	0.002	-0.009
	(0.006)	(0.008)
Disabled	0.02	0.007
	(0.032)	(0.042)
Grade	-0.147	-0.241
	(0.331)	(0.484)
Standardized reading score	-0.033	-0.076*
	(0.034)	(0.043)
Standardized math score	-0.037	-0.055
	(0.029)	(0.040)
Sch. Performance Index	-0.763	-0.674
	(0.607)	(0.851)
Charter school	0.011	0.226
	(0.113)	(0.149)
Enrollment	89.689	104.324
	(55.877)	(71.017)
Attendance rate	-0.106	-0.657
	(0.769)	(0.876)
Reading Value-added	0.002	0.145
	(0.750)	(1.034)

Table 4. Coefficients and standard errors on SIG eligibility indicator from OLS model predicting observable characteristics

Outcome Variable	**(1) Coef. (Std. Error)**	**(2) Coef. (Std. Error)**
Math Value-added	-0.723	-0.030
	(0.775)	(1.034)
Total Value-added	-0.361	0.058
	(0.630)	(0.795)
Mean Yrs. Teacher Exp.-Sch.	0.055	-2.413
	(1.565)	(2.065)
Pct. Teachers w/MA-Sch.	1.253	-7.971
	(5.051)	(6.427)
Mean Teacher Salary-Sch.	2447.76	-5867.316
	(3676.793)	(4871.696)
N	153,625	153,625
Linear term	X	X
Linear term interacted with SIG indicator	X	X
Quadratic term		X
Quadratic term interacted with SIG indicator		X

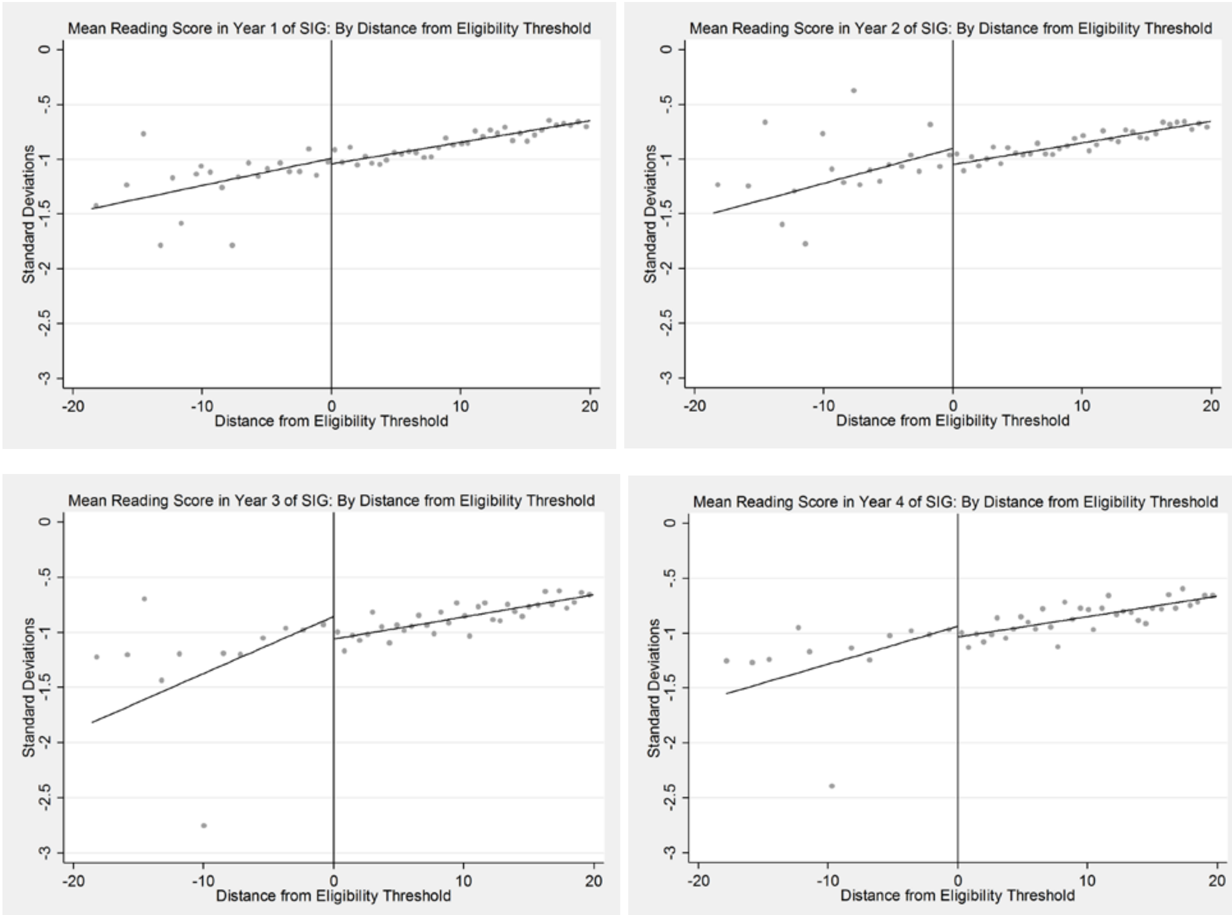
NOTE: Standard errors clustered by school in parentheses below coefficient estimates. All estimates based on analytic sample containing all students in schools within 20 points of the SIG eligibility threshold. Model controls for distance from the reclassification threshold using the specification denoted in the table. *p<0.10; **p<0.05; ***p<0.01

Table 5. Coefficients and standard errors for indicators of SIG eligibility from models predicting student achievement levels, by subject, year, and modeling approach

	Reading				Math			
	Year 1 of SIG	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**	**Year 1 of SIG**	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**
	Cross-sectional Models							
SIG Eligibility	0.053	0.156*	0.215*	0.111	0.046	0.171*	0.215*	0.088
	(0.052)	(0.090)	(0.119)	(0.086)	(0.051)	(0.095)	(0.112)	(0.089)
N	71,231	65,273	60,281	57,433	71,217	65,233	60,312	57,437
N Schools	298	261	250	239	298	261	250	239
	Dynamic Model							
SIG Eligibility	0.078	0.204*	0.269*	0.221*	0.084	0.208**	0.257**	0.171
	(0.064)	(0.114)	(0.140)	(0.131)	(0.053)	(0.103)	(0.120)	(0.107)
N	330,102				330,082			
N Schools	308				308			

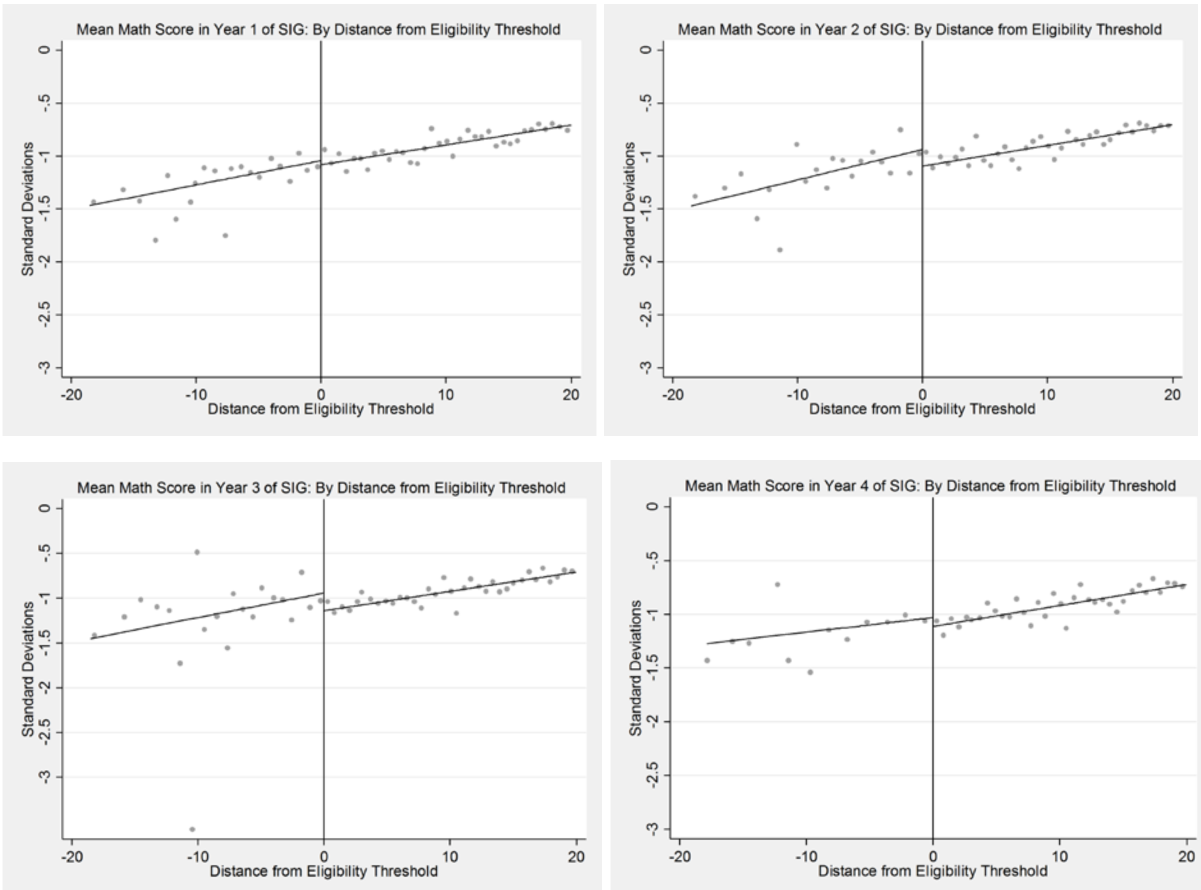
NOTE: Standard errors clustered by school in parentheses below coefficient estimates. Estimates based on analytic samples containing all students in schools within 20 points of the SIG eligibility threshold in the year of SIG eligibility identification. Models control for distance from the eligibility threshold using a linear term interacted with the SIG eligibility indicator. Cross-sectional models contain a school-level measure of baseline achievement. Dynamic models based on the approach proposed by Cellini, Ferreira, and Rothstein (2010). *p<0.10; **p<0.05; ***p<0.01

Figure 3. Mean Reading Achievement, by Distance from the SIG Eligibility Threshold



Note: Markers in each panel of the figure represent mean reading achievement of students attending schools in each bin of ODE "combined proficiency rate" measure used to determine SIG eligibility. Each panel also contains a line of best fit that is fitted separately on each side of the SIG eligibility threshold.

Figure 4. Mean Math Achievement, by Distance from the SIG Eligibility Threshold



Note: Markers in each panel of the figure represent mean math achievement of students attending schools in each bin of ODE "combined proficiency rate" measure used to determine SIG eligibility. Each panel also contains a line of best fit that is fitted separately on each side of the SIG eligibility threshold.

Table 6. Coefficients and standard errors for indicators of SIG eligibility from non-parametric estimator predicting student achievement levels, by subject and year

	Reading				Math			
	Year 1 of SIG	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**	**Year 1 of SIG**	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**
SIG Eligibility- Conventional	-0.012	0.124	0.203*	0.173*	-0.084	0.138	0.179	0.164
Estimate	(0.065)	(0.086)	(0.123)	(0.097)	(0.074)	(0.122)	(0.147)	(0.110)
SIG Eligibility- Bias-corrected	-0.024	0.107	0.199	0.176*	-0.104	0.147	0.186	0.170
estimate	(0.065)	(0.086)	(0.123)	(0.097)	(0.074)	(0.122)	(0.147)	(0.110)
N	281,802	273,189	262,964	257,263	281,720	273,105	262,666	256,656
Effective *N*	13,619	11,272	10,236	10,039	13,121	14,314	13,320	9,598
N Schools	145	116	90	105	141	141	127	91
Bandwidth for estimate	4.219	3.942	3.922	4.034	4.067	5.227	5.153	3.719
Bandwidth for bias correction	7.144	6.614	5.658	6.649	6.806	8.656	7.929	5.897

NOTE: Standard errors clustered by school in parentheses below coefficient estimates. Estimates produced by the approach proposed by Calonico, Cattaneo, and Titiunik (2014a). *p<0.10; **p<0.05; ***p<0.01

Table 7. Coefficients and standard errors for SIG Turnaround and Transformation models from models predicting student achievement levels, by subject, year, and modeling approach

	Reading				Math			
SIG Model	**Year 1 of SIG**	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**	**Year 1 of SIG**	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**
	Cross-sectional Model							
Treat X Turnaround	0.016	0.191***	0.269***	0.220**	-0.018	0.159*	0.240**	0.211**
	(0.046)	(0.061)	(0.095)	(0.098)	(0.068)	(0.084)	(0.103)	(0.097)
Treat X Trans	0.095	0.159	0.137	0.068	0.073	0.165	0.178	0.063
	(0.067)	(0.107)	(0.108)	(0.088)	(0.055)	(0.111)	(0.122)	(0.098)
N	71,287	65,327	60,328	57,489	71,274	65,285	60,359	57,495
N Schools	298	262	250	240	298	262	250	240
	Dynamic Model							
Treat X Turnaround	0.047	0.207**	0.296***	0.261**	0.040	0.218*	0.292**	0.285**
	(0.055)	(0.098)	(0.112)	(0.116)	(0.081)	(0.128)	(0.125)	(0.111)
Treat X Trans	0.099	0.178	0.174	0.134	0.090	0.182	0.205	0.113
	(0.072)	(0.132)	(0.133)	(0.123)	(0.058)	(0.118)	(0.133)	(0.117)
N	330,102				330,082			
N Schools	308				308			

NOTE: Standard errors clustered by school in parentheses below coefficient estimates. Estimates represent mean achievement differences between SIG-ineligible schools and schools that implemented the Turnaround and Transformation models, respectively. Estimates based on analytic samples containing all students in schools within 20 points of the SIG eligibility threshold in the year of SIG eligibility identification. Models control for distance from the eligibility threshold using a linear term interacted with the SIG eligibility indicator. Cross-sectional models contain a school-level measure of baseline achievement. Dynamic models based on the approach proposed by Cellini, Ferreira, and Rothstein (2010). *p<0.10; **p<0.05; ***p<0.01

Table 8. Coefficients and standard errors from models predicting per-pupil expenditures, by year relative to SIG eligibility identification

	Year 1 of SIG	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**
	Cross-sectional RD Models			
SIG Eligibility	-43.04	1,692.144*	2,910.849**	1,558.97
	(665.90)	(1,012.39)	(1,356.82)	(1,136.55)
N	496	453	261	279
N Schools	299	266	212	239
	Dynamic RD Model			
SIG Eligibility	351.54	1,980.886**	2,572.86**	3289.86*
	(517.46)	(925.22)	(1,011.89)	(1,965.60)
N	2,389			
N Schools	306			
	Difference-in-Differences Model			
SIG Eligibility	-346.49	1,559.65	3289.86*	1527.77*
	(341.34)	(1,000.09)	(1,965.60)	(883.74)
N	6,936			
N Schools	850			

NOTE: Standard errors clustered by school in parentheses below coefficient estimates. All estimates based on analytic sample containing all schools within 20 points of the SIG eligibility threshold in the year of SIG eligibility identification. Models control for distance from the eligibility threshold using a linear term interacted with the SIG eligibility indicator. Cross-sectional models contain a school-level measure of baseline achievement. Dynamic models based on the approach proposed by Cellini, Ferreira, and Rothstein (2010). *p<0.10; **p<0.05; ***p<0.01

Table 9. Coefficients and standard errors for SIG Turnaround and Transformation models from models predicting per-pupil expenditures, by year and modeling approach

	Year 1 of SIG	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**
	Cross-sectional RD Models			
Treat X Turnaround	-1291.58	209.65	654.36	-1,885.63
	(1320.430)	(1844.570)	(1893.088)	(2149.808)
Treat X Transformation	425.23	2835.00***	4284.34***	2865.64**
	(767.180)	(1027.724)	(1429.631)	(1166.466)
N	496	453	261	279
N Schools	299	266	212	239
	Dynamic RD Model			
Treat X Turnaround	902.90	2707.78	2865.95	1843.83
	(822.777)	(1676.443)	(1874.721)	(1303.887)
Treat X Transformation	860.18	3320.84**	4094.97***	2775.79*
	(720.817)	(1317.166)	(1508.280)	(1558.934)
N	2,389			
N Schools	306			
	Difference-in-Differences Model			
Treat X Turnaround	412.649	1454.093	2008.691	1401.24***
	(593.926)	(1283.390)	(1502.108)	(504.702)
Treat X Transformation	66.068	1621.92***	2692.29***	2040.48***
	(306.889)	(442.460)	(538.951)	(605.401)
N	6,936			
N Schools	850			

NOTE: Standard errors clustered by school in parentheses below coefficient estimates. All estimates based on analytic sample containing all schools within 20 points of the SIG eligibility threshold in the year of SIG eligibility identification. Models control for distance from the eligibility threshold using a linear term interacted with the SIG eligibility indicator. Cross-sectional models contain a school-level measure of baseline achievement. Dynamic models based on the approach proposed by Cellini, Ferreira, and Rothstein (2010). *p<0.10; **p<0.05; ***p<0.01

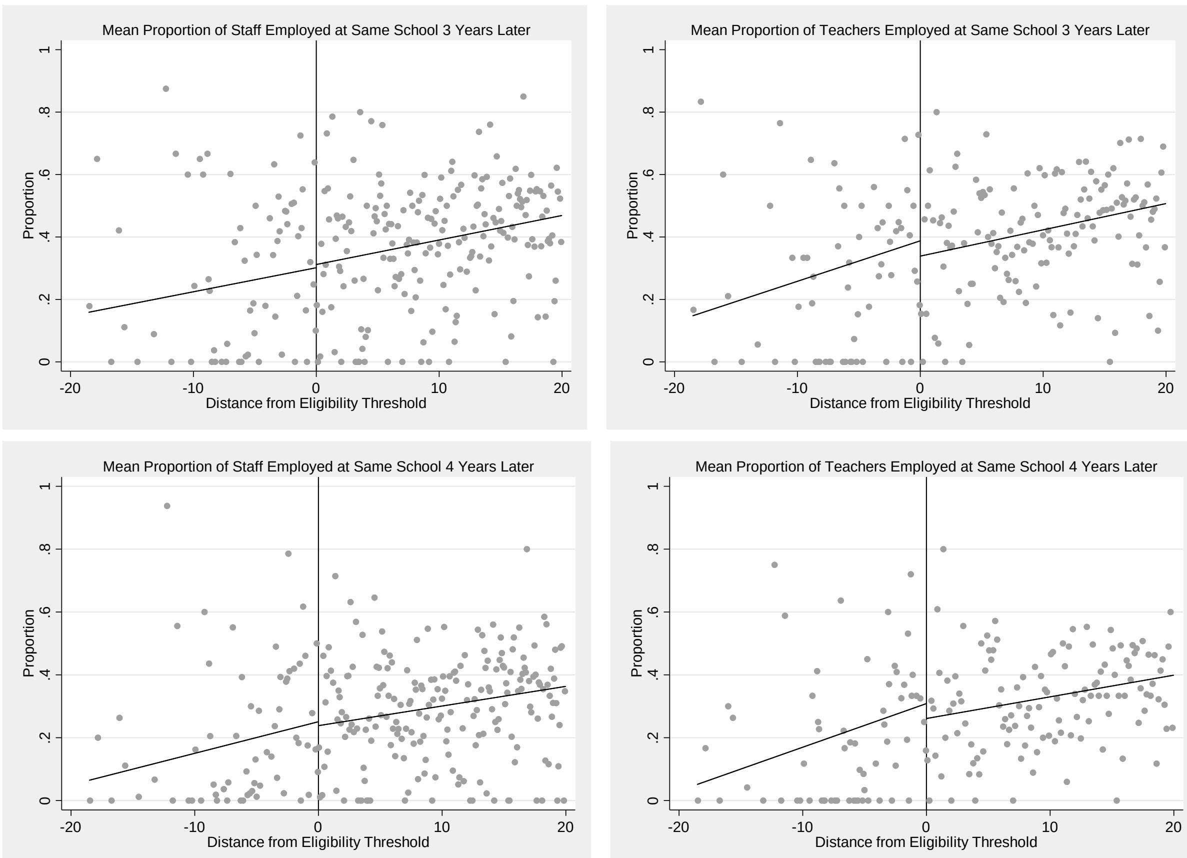
Table 10. Coefficients and standard errors for indicators of SIG eligibility from models predicting staff turnover

	[1] All staff	[2] Teachers
	Employed Three Years After SIG Eligibility	
Effect of SIG Eligibility	-0.010	0.049
	(0.060)	(0.058)
N	21,858	10,354
N Schools	307	305
	Employed Four Years After SIG Eligibility	
Effect of SIG Eligibility	0.012	0.047
	(0.049)	(0.051)
N	21,858	10,354
N Schools	307	307
	Year-by-Year Turnover	
Turnover 1 Year After SIG Eligibility	0.025	-0.037
	(0.031)	(0.030)
N	24,799	11,632
N Schools	303	303
Turnover 2 Years After SIG Eligibility	-0.019	0.005
	(0.023)	(0.030)
N	22,715	10,639
N Schools	268	268
Turnover 3 Years After SIG Eligibility	-0.028	-0.014
	(0.021)	(0.033)
N	21,551	10,023
N Schools	254	254
Turnover 4 Years After SIG Eligibility	-0.039	-0.052
	(0.026)	(0.035)
N	20,430	10,023
N Schools	245	254

NOTE: Standard errors clustered by school in parentheses below coefficient estimates. All estimates based on analytic sample containing staff in schools within 20 points of the SIG eligibility threshold in the year of SIG eligibility identification. Model controls for distance from the reclassification threshold using linear term interacted with the SIG eligibility indicator . *p<0.10; **p<0.05; ***p<0.01.

Figure 5. Mean Proportion of Staff Employed at School Three and Four Years After SIG Eligibility Identification, by Distance from the SIG Eligibility Threshold

All Staff *Teachers*



Note: Markers in each panel of the figure represent mean proportion of staff still employed at the same school three or four years later in each bin of ODE "combined proficiency rate" measure used to determine SIG eligibility. Each panel also contains a line of best fit that is fitted separately on each side of the SIG eligibility threshold.

Table 11. Coefficients and standard errors for SIG Turnaround and Transformation models from models predicting staff turnover, by year

	All Staff		Teachers	
	Turnaround	Transformation	Turnaround	Transformation
	Employed Three Years After SIG Eligibility			
Coef. For SIG model	-0.034	0.045	-0.145***	0.143***
	(0.075)	(0.067)	(0.055)	(0.049)
N	21,858		10,354	
N Schools	307		305	
	Employed Four Years After SIG Eligibility			
Coef. For SIG model	0.017	0.059	-0.072	0.119**
	(0.070)	(0.055)	(0.052)	(0.048)
N	21,858		10,354	
N Schools	307		305	

NOTE: Standard errors clustered by school in parentheses below coefficient estimates. Estimates represent mean differences in the probability of remaining employed between SIG-ineligible schools and schools that implemented the Turnaround and Transformation models, respectively. Estimates based on analytic samples containing all students in schools within 20 points of the SIG eligibility threshold in the year of SIG eligibility identification. Models control for distance from the eligibility threshold using a linear term interacted with the SIG eligibility indicator. *p<0.10; **p<0.05; ***p<0.01

Table 12. Coefficients and standard errors for indicators of SIG eligibility from models predicting staffing levels, by staff position, year, and modeling approach

	Cross-sectional Models				Dynamic Models			
	Year 1 of SIG	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**	**Year 1 of SIG**	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**
	All Staff							
SIG Eligibility	-3.797 (2.738)	-0.175 (2.940)	1.039 (3.084)	0.127 (3.107)	-5.668* (3.216)	-2.398 (3.342)	-1.441 (3.415)	-1.948 (3.397)
N	499	459	436	419	2,216			
N Schools	303	272	257	247	307			
	Teachers							
SIG Eligibility	0.954 (0.681)	1.613 (1.149)	1.742 (1.467)	0.987 (1.267)	0.665 (0.839)	1.291 (1.465)	1.523 (1.886)	0.919 (1.924)
N	499	459	436	419	2,216			
N Schools	303	272	257	247	307			
	Education Professionals							
SIG Eligibility	0.548 (0.851)	2.092 (1.450)	2.815 (1.885)	1.923 (1.883)	-0.070 (1.038)	1.624 (1.711)	2.413 (2.305)	1.790 (2.511)
N	499	459	436	419	2,216			
N Schools	303	272	257	247	307			
	Administrators							
SIG Eligibility	-0.855** (0.402)	-0.265 (0.434)	-0.293 (0.510)	-0.357 (0.551)	-1.576** (0.797)	-1.123 (0.838)	-1.099 (0.747)	-1.018 (0.773)
N	499	459	436	419	2,216			
N Schools	303	272	257	247	307			

NOTE: Standard errors clustered by school in parentheses below coefficient estimates. All estimates based on analytic sample containing all students in schools within 20 points of the SIG eligibility threshold in the year of SIG eligibility identification. Models control for distance from the SIG eligibility threshold with a linear term interacted with the SIG eligibility indicator. Cross-sectional models contain a baseline measure of staffing levels. Dynamic models based on the approach proposed by Cellini, Ferreira, and Rothstein (2010). *p<0.10; **p<0.05; ***p<0.01.

Table 13. Coefficients and standard errors for indicators of SIG eligibility from models predicting school closure, by subject, year, and specification

	Year 1 of SIG	**Year 2 of SIG**	**Year 3 of SIG**	**Year 4 of SIG**
		Cross-sectional RD Models		
SIG Eligibility	-0.010	-0.014	-0.023	0.004
	(0.047)	(0.072)	(0.081)	(0.093)
N	504	504	504	504
N Schools	308	308	308	308
		Dynamic RD Model		
SIG Eligibility	-0.003	-0.010	-0.021	0.006
	(0.051)	(0.075)	(0.084)	(0.097)
N		3,027		
N Schools		308		

NOTE: Standard errors clustered by school in parentheses below coefficient estimates. All estimates based on analytic sample containing all schools within 20 points of the SIG eligibility threshold in the year of SIG eligibility identification. Models control for distance from the SIG eligibility threshold with a linear term interacted with the SIG eligibility indicator. Dynamic models based on the approach proposed by Cellini, Ferreira, and Rothstein (2010) *p<0.10; **p<0.05; ***p<0.01.