U.S. Department of Education
January 2017

School Improvement Grants: Implementation and Effectiveness

Lisa Dragoset
Jaime Thomas
Mariesa Herrmann
John Deke
Susanne James-Burdumy
Mathematica Policy Research

Cheryl Graczewski
Andrea Boyle
Rachel Upton
Courtney Tanenbaum
Jessica Giffin
American Institutes for Research

Thomas E. Wei
Project Officer
Institute of Education Sciences

ies INSTITUTE OF EDUCATION SCIENCES

NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE

School Improvement Grants: Implementation and Effectiveness

January 2017

Lisa Dragoset
Jaime Thomas
Mariesa Herrmann
John Deke
Susanne James-Burdumy
Mathematica Policy Research

Cheryl Graczewski
Andrea Boyle
Rachel Upton
Courtney Tanenbaum
Jessica Giffin
American Institutes for Research

Thomas E. Wei
Project Officer
Institute of Education Sciences

NCEE 2017-4013
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

U.S. Department of Education
John King
Secretary

Institute of Education Sciences
Ruth Curran Neild
Deputy Director for Policy and Research
Delegated Duties of the Director

National Center for Education Evaluation and Regional Assistance
Audrey Pendleton
Acting Commissioner

January 2017

This report was prepared for the Institute of Education Sciences under Contract ED-IES-10-C-0077. The project officer is Thomas E. Wei in the National Center for Education Evaluation and Regional Assistance.

IES evaluation reports present objective information on the conditions of implementation and impacts of the programs being evaluated. IES evaluation reports do not include conclusions or recommendations or views with regard to actions policymakers or practitioners should take in light of the findings in the report.

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.

This report is available on the IES website at http://ies.ed.gov/ncee.

Alternate Formats: Upon request, this report is available in alternate formats, such as Braille, large print, audiotape, or computer diskette. For more information, please contact the Department's Alternate Format Center at 202-260-9895 or 202-205-8113.

ACKNOWLEDGMENTS

Many people contributed in significant ways to the Impact Evaluation of School Improvement Grants. First, we would like to thank members of the evaluation's Technical Work Group—Thomas Cook, Thomas Fisher, Guido Imbens, Brian Jacob, Thomas Kane, Sean Reardon, Eric Smith, Jeffrey Smith, James Spillane, Elizabeth Stuart, and Jonathan Supovitz—who imparted valuable input at critical junctures.

At Mathematica Policy Research, important contributions were made by Phil Gleason, who provided a thoughtful, critical review of the report; Eric Zeidman, Kristin Hallgren, and Nancy Duda, who led our large team of dedicated staff who recruited districts into the study; Irma Perez-Johnson and Kristin Hallgren, who led the many committed individuals who interviewed districts; David DesRoches, who led the school survey effort with assistance from Lauren Maul, who tracked the school sample during data collection; Hermine Marcovici and her team at Mathematica's Survey Operation Center, who expertly followed up with schools; Elias Walsh, who provided a careful review of the findings from the regression discontinuity analyses; Mark Brinkley, who led the development and management of the web school survey with assistance from Jason DiBartolo, who specified requirements for the survey, Bea Jones, who maintained a tracking database, and Roland Scurato, who developed programs to send email reminders to schools; Cassie Pickens Jewell, who developed and managed project databases and worked to resolve questions about Memorandums of Understanding with districts; Jacob Hartog and Marykate Zukiewicz, who identified interview and survey questions that aligned with the SIG application; Alexandra Killewald, who assisted with the selection of the study districts and schools; Amanda Beatty and Mai Miksic, who compiled prior research on the types of reforms promoted by SIG; John Chen, Mason DeCamillis, Emily Evans, Matthew Jacobus, Malik Mubeen, Przemyslaw Nowaczyk, Luis Rodriguez, and Lisa Shang, who provided excellent programming assistance in cleaning and analyzing data; and Leah Hackleman-Good, Jennifer Littel, and Colleen Fitts, who expertly edited and produced the report.

At American Institutes for Research, we thank Liz Grant, who provided expert leadership; Mike Garet, who provided valuable quality assurance review; and Jizhi Zhang and Cong Ye, who provided excellent programming assistance in cleaning and analyzing data.

At Social Policy Research Associates, we thank Sukey Leshnick and Castle Sinicrope, who supervised dedicated teams of recruiters and interviewers from Social Policy Research Associates.

Finally, we would like to extend a special thanks to the many school districts and schools who participated in the study. Without their strong support and participation, this study would not have been possible.

CONTENTS

TABLES

FIGURES

EXECUTIVE SUMMARY

In response to the recession that began in 2007, the U.S. Congress passed, and President Barack Obama signed into law, the American Recovery and Reinvestment Act of 2009 (Pub. Law 111-5). At an estimated cost of $831 billion, this economic stimulus package sought to save and create jobs, provide temporary relief to those adversely affected by the recession, and invest in education, health, infrastructure, and renewable energy. States and school districts received $100 billion to secure teachers' jobs and promote innovation in schools. This funding included $3 billion for School Improvement Grants (SIG), one of the Obama administration's signature programs and one of the largest federal government investments in an education grant program. The SIG program awarded grants to states that agreed to implement one of four school intervention models—transformation, turnaround, restart, or closure—in their lowest-performing schools. Each of the models prescribed specific practices designed to improve student outcomes, including outcomes for high-need students such as English language learners (ELLs) (U.S. Department of Education 2010a).

Given the importance of the SIG program and sizable investment in it, the Institute of Educations Sciences (IES) commissioned this evaluation to focus on four primary questions:

- Did schools implementing a SIG-funded model use the improvement practices promoted by SIG, and how did that compare to use of those practices by schools not implementing a SIG-funded model?
- Did use of SIG-promoted practices include a focus on ELLs, and did that focus on ELLs differ between schools implementing a SIG-funded model and schools not implementing one?
- Did receipt of SIG funding to implement a school intervention model have an impact on outcomes for low-performing schools?
- Was the type of school intervention model implemented related to improvement in outcomes for low-performing schools?

The Every Student Succeeds Act of 2015 (ESSA) made changes to the SIG program that gives states and districts much more flexibility in determining how to turn around their lowest-achieving schools. For example, the U.S. Department of Education (ED) can no longer require the use of particular school intervention models, and funds previously set aside for SIG now flow through the regular Title I formula (Klein 2015). Despite these changes, findings on the first two questions remain useful to policymakers considering the future direction of funds for low-performing schools because they identify practice areas that these schools have and have not yet addressed. Further, findings on the first question provide a useful policy context for interpreting findings on the third question of whether the $3 billion federal investment had a positive impact on student achievement. For example, if use of the practices promoted by SIG was similar between schools that received grants and schools that did not, then it seems less likely that SIG would have a subsequent impact on student achievement. Findings on the fourth question, which shed light on whether certain models were associated with larger student achievement gains than other models, remain relevant for educators and administrators considering future evidence-based approaches for turning around low-performing schools.

This is the final report for this evaluation of SIG. Three earlier briefs focused on: (1) implementation of three interrelated levers for school improvement—granting low-performing schools operational authority, supporting them, and monitoring their progress (Herman et al. 2014); (2) low-performing schools' adoption of individual practices and combinations of practices promoted by SIG (Herrmann et al. 2015); and (3) states' capacity to support school turnaround (Tanenbaum et al. 2015). An earlier report covered all major topic areas that SIG promoted, examining the extent to which schools implementing a SIG-funded model and schools not implementing one reported using these practices in spring 2012, and whether use differed across these two groups of schools (Dragoset et al. 2015). This final report builds on the earlier briefs and report by including an additional year of data (spring 2013) and by examining whether receipt of SIG funding had an impact on student outcomes.

Key findings

SIG allowed grantees to implement one of four school intervention models (transformation, turnaround, restart, or closure). These models promoted the use of many improvement practices in four main areas: (1) adopting comprehensive instructional reform strategies, (2) developing and increasing teacher and principal effectiveness, (3) increasing learning time and creating community-oriented schools, and (4) having operational flexibility and receiving support. It is worth knowing to what extent schools implementing these models with SIG funds (referred to as SIG-funded models throughout this report) actually used these practices, and how that compares to other schools. We examined the use of these SIG-promoted practices in two ways: (1) we conducted a descriptive analysis that compared use of these practices for 290 schools that implemented a SIG-funded model in 2012–2013 and 190 schools that did not, and (2) we used a regression discontinuity design (RDD) with data from 460 schools to examine whether implementation of a SIG-funded model in 2010–2011 had an impact on use of these practices.

We also examined whether the SIG program had an impact on student outcomes. We used an RDD to calculate the overall impact of implementing any of the four SIG-funded models on test scores, high school graduation, and college enrollment. A sample of 190 schools eligible for SIG and 270 schools that were not eligible for SIG was used in the analysis.

Finally, we examined whether certain intervention models were associated with larger student achievement gains than other models. We conducted a correlational analysis that examined the relationship between the type of model implemented and changes in student achievement over time. A sample of 270 schools that implemented a SIG-funded model in 2010–2011 was used in the analysis.

Key findings included:

- **Although schools implementing SIG-funded models reported using more SIG-promoted practices than other schools, we found no evidence that SIG *caused* those schools to implement more practices.** Our descriptive analysis found that schools implementing a SIG-funded model used significantly more SIG-promoted practices than other schools (22.8 of the 35 practices examined [65 percent] versus 20.3 practices [58 percent], a difference of 2.5 practices). Our more rigorous RDD analysis found a similar difference of 3.3 practices, but it was not statistically significant. Therefore, we are unable to conclude that SIG *caused* the observed difference in use of practices.

- **Across all study schools, use of SIG-promoted practices was highest in comprehensive instructional reform strategies and lowest in operational flexibility and support.** In the comprehensive instructional reform strategies area, study schools reported using, on average, 7.1 of the 8 SIG-promoted practices examined (89 percent). In the operational flexibility and support area, study schools reported using, on average, 0.87 of the 2 SIG-promoted practices examined (43 percent).
- **There were no significant differences in use of English Language Learner (ELL)-focused practices between schools implementing a SIG-funded model and other schools.**
- **Overall, across all grades, we found that implementing any SIG-funded model had no significant impacts on math or reading test scores, high school graduation, or college enrollment.**
- **When we compared student achievement gains from different models in elementary grades (2nd through 5th), we found no evidence that one model was associated with larger gains than another. For higher grades (6th through 12th), the turnaround model was associated with larger student achievement gains in math than the transformation model.** However, factors other than the SIG model implemented, such as baseline differences between schools implementing different models, may explain these differences in achievement gains.

Background

The SIG program aimed to support the implementation of school intervention models in low-performing schools. Although SIG was first authorized in 2001, this evaluation focused on SIG awards granted in 2010, when roughly $3.5 billion in SIG awards were made to 50 states and the District of Columbia, $3 billion of which came from the American Recovery and Reinvestment Act of 2009. States identified the low-performing schools eligible for SIG based on criteria specified by ED and then held competitions for local education agencies seeking funding to help turn around eligible schools.

For the 2010 SIG competition, ED required states to categorize schools into three eligibility tiers based on the school's level (elementary or secondary), Title I status,[1] and achievement or graduation rate. These tiers helped prioritize the distribution of SIG funds at the local level and determined the practices to be used for school turnaround. In general, SIG eligibility Tiers I and II included schools with the lowest achievement and most persistent achievement problems in each state.

[1] Title I, Part A (Title I) of the Elementary and Secondary Education Act provides financial assistance to local educational agencies and schools with many children from low-income families. A school receiving Title I funds that fails to meet adequate yearly progress targets can be assigned a Title I status of "in need of improvement, corrective action, or restructuring."

ED required that each SIG-awarded school under Tier I or Tier II implement one of four school intervention models. These models required specific practices:

- **Transformation.** This model required schools to replace the principal, adopt a teacher and principal evaluation system that accounted for student achievement growth as a significant factor, adopt a new governance structure, institute comprehensive instructional reforms, increase learning time, create community-oriented schools, and have operational flexibility.
- **Turnaround.** This model required schools to replace the principal, replace at least 50 percent of the school staff, institute comprehensive instructional reforms, increase learning time, create community-oriented schools, and have operational flexibility.
- **Restart.** This model required schools to convert to a charter school or close and reopen under the management of a charter management organization or education management organization.
- **School closure.** This model required districts to close schools and enroll their students in higher-achieving schools within the district.

These required practices can be grouped into the four main topic areas promoted by SIG. Table ES.1 lists these four broad areas and the objectives promoted by SIG within each. For example, replacing the principal falls under the “identifying and rewarding effective teachers and principals and removing ineffective ones” objective within the “developing and increasing teacher and principal effectiveness” topic area. The objectives listed in the table cover all practices promoted by SIG, which different models designated as either required or permissible. For example, adopting a teacher and principal evaluation system that accounted for student achievement growth was required under the transformation model but permissible under the turnaround and restart models. For detailed information about the practices that each model required, see Appendix F, Table F.1.

We did not limit our examination to the specific practices required by each model. We instead focused on all required or permissible practices under the transformation or turnaround models because (a) both models prescribed a large set of overlapping practices, (b) restart model schools could choose to use any of those practices, and (c) an earlier report from this study (Herrmann et al. 2014) already presented findings on the implementation of required practices by schools using different models.

Table ES.1. SIG objectives, by topic area

Implementing comprehensive instructional reform strategies
Using Data to Identify and Implement an Instructional Program
Promoting the Continuous Use of Student Data
Conducting Periodic Reviews to Ensure that the Curriculum is Being Implemented with Fidelity
Implementing a New School Model (Such As an Academy with a Theme Focused on Science, Technology, Engineering and Math)
Providing Supports and Professional Development to Staff to Assist ELL Students and Students with Disabilities
Using and Integrating Technology-Based Supports
Tailoring Strategies for Secondary Schools

Developing and increasing teacher and principal effectiveness
Using Rigorous, Transparent, and Equitable Evaluation Systems
Identifying and Rewarding Effective Teachers and Principals and Removing Ineffective Ones
Providing High-Quality, Job-Embedded Professional Development or Supports
Implementing Strategies to Recruit, Place, and Retain Staff
Increasing learning time and creating community-oriented schools
Using Schedules and Strategies That Provide Increased Learning Time or Increasing the Number of Hours per Year That School Was in Session
Engaging Families and Communities and Providing a Safe School Environment That Meets Students' Social, Emotional, and Health Needs
Having operational flexibility and receiving support
Having Primary Responsibility for Budget, Hiring, Discipline, or School Year Length Decisions
Receiving Technical Assistance and Support

Source: SIG application.

ELL = English language learner.

Research questions and study design

Including the four primary research questions listed earlier (and italicized below), this report was guided by a total of seven research questions in three broad areas:

Use of SIG-promoted practices

1. *Did schools implementing a SIG-funded model use the improvement practices promoted by SIG, and how did that compare to use of those practices by schools not implementing a SIG-funded model?*
2. Did receipt of SIG funding to implement a school intervention model have an impact on the number of SIG-promoted practices used by low-performing schools?
3. Did schools' use of SIG-promoted practices change over time?
4. *Did use of SIG-promoted practices include a focus on ELLs, and did that focus on ELLs differ between schools implementing a SIG-funded model and schools not implementing one?*
5. Did use of these ELL-focused improvement practices differ based on the prevalence of ELL students in the school or the achievement gap between ELL and other students?

Whether SIG-funded intervention models improved student outcomes

6. *Did receipt of SIG funding to implement a school intervention model have an impact on outcomes for low-performing schools?*

Whether the type of model was related to improvements in student outcomes

7. *Was the type of school intervention model implemented related to improvement in outcomes for low-performing schools?*

Here we describe the study sample, the data collected, and the methods we used to analyze the data. The sample for the SIG evaluation included 22 states and approximately 60 districts.

Data on the use of SIG-promoted practices came from surveys of approximately 480 school administrators conducted in spring 2012 and spring 2013. Data on student outcomes came from student-level administrative data obtained from states and districts. We analyzed the implementation of SIG through both a descriptive analysis (that compared survey responses from 290 schools that implemented SIG-funded models in 2012–2013 to 190 schools that did not) and a more rigorous RDD analysis (that compared survey responses from 190 schools eligible for SIG funds in 2010–2011 to 270 schools that were not eligible for SIG). We examined the overall impact of implementing any of the four SIG-funded models using an RDD analysis that compared test score data from the 190 schools that met the SIG eligibility criteria to the 270 schools that missed the cutoff for eligibility. We used a descriptive, correlational analysis of 270 schools that implemented a SIG-funded model in 2010–2011 to examine whether certain intervention models appeared more effective than others at improving student achievement.

Prior to receiving a grant, SIG-funded model schools had baseline characteristics similar to those of other study schools

Interpreting the differences between schools that implemented a SIG-funded model and schools that did not requires understanding the characteristics of these two groups of schools at baseline (during the 2009–2010 school year, which was prior to SIG funding receipt).

- **Schools that implemented a SIG-funded model and schools that did not had similar observable characteristics prior to receipt of the 2010 SIG awards.**[2] These included several student and school demographic characteristics, such as race/ethnicity, percentage of students eligible for free or reduced-price lunch, Title I eligibility, location, and school level (elementary, middle, high school).
- **Study schools implementing a SIG-funded model were generally not representative of all U.S. schools implementing such models**. The schools in our study that were implementing SIG-funded models were more disadvantaged and more likely to be in an urban area than U.S. schools nationally that were implementing such models. In particular, study schools implementing a SIG-funded model had higher percentages of students eligible for free or reduced-price lunch than U.S. schools nationally implementing such models (83 percent versus 77 percent), and were more likely to be located in an urban area (88 percent versus 58 percent). Because the SIG sample is not representative of schools nationwide, the findings here may not apply to all schools nationally.

Detailed Findings

Schools implementing SIG-funded models reported using more SIG-promoted practices than other schools, but we found no evidence that SIG *caused* those schools to implement more practices

We examined whether implementing a SIG-funded model was associated with using more SIG-promoted practices. This is an important first step in understanding the extent to which the SIG program might improve student achievement. If schools implementing a SIG-funded model

[2] The baseline characteristics examined came from the Common Core of Data. To limit respondent burden, the school administrator surveys focused primarily on practices schools were using in spring 2012 and spring 2013.

used the same practices as similar schools that did not implement a SIG-funded model, it is less likely that any changes in outcomes for SIG schools—positive or negative—could be attributed to the program.

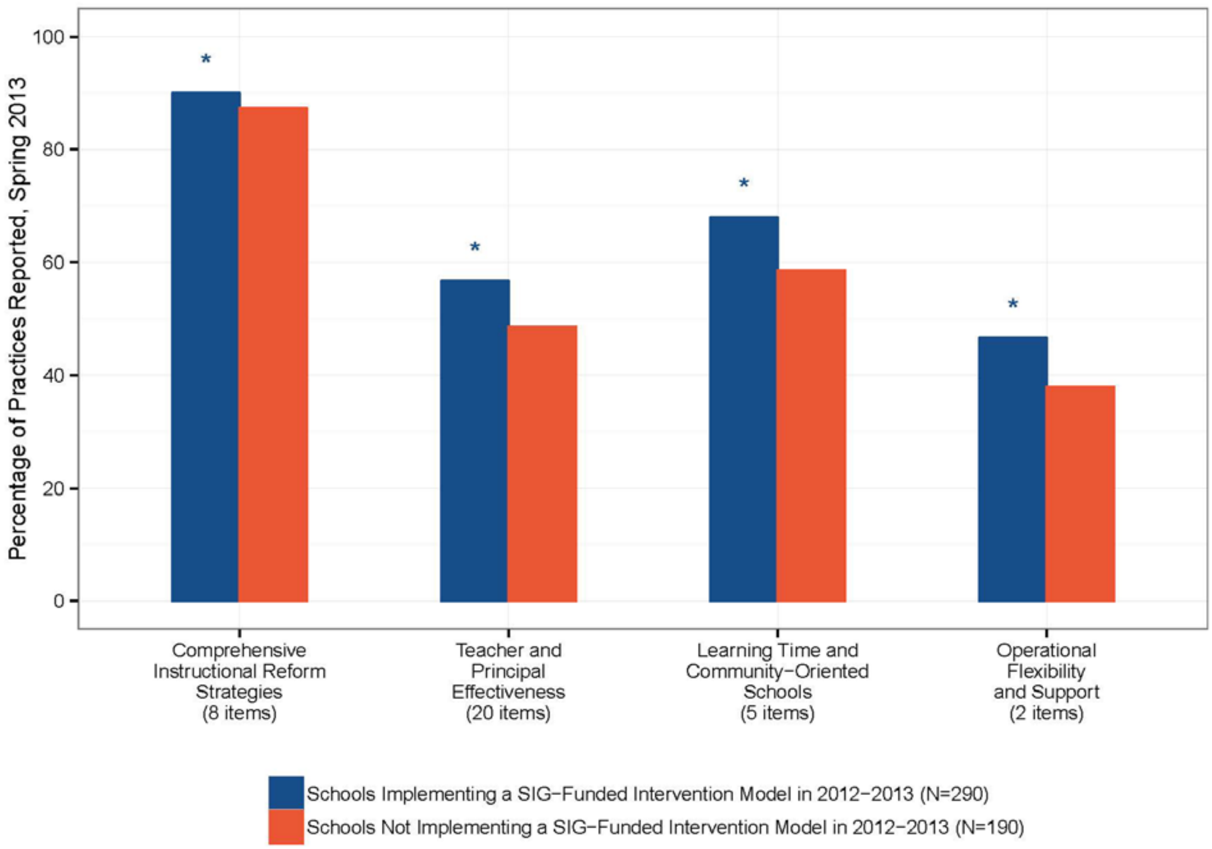
Using a descriptive analysis, we found that in spring 2013, schools that implemented a SIG-funded model reported using more practices, on average, than schools that did not in the following areas (Figure ES.1):

- Comprehensive instructional reform strategies (90 percent of practices reported by SIG-funded model schools compared to 87 percent for schools not implementing such models)
- Teacher and principal effectiveness (57 percent of practices reported by SIG-funded model schools compared to 49 percent for schools not implementing such models)
- Learning time and community-oriented schools (68 percent of practices reported by SIG-funded model schools compared to 59 percent for schools not implementing such models)
- Operational flexibility and support (47 percent of practices reported by SIG-funded model schools compared to 38 percent for schools not implementing such models)

Adding up the differences across the four areas, schools implementing a SIG-funded model reported using more SIG-promoted practices overall (65 percent of the 35 practices examined, or 22.8 practices) than schools not implementing one (58 percent of the 35 practices examined, or 20.3 practices), a difference of 7 percentage points (2.5 practices). It is not clear whether a difference of this size would be meaningful in its overall influence on improvement practices and school outcomes.

The spring 2013 findings presented in this report were generally the same as the spring 2012 findings presented in an earlier report from this evaluation (Dragoset et al. 2015).

Figure ES.1. Use of practices promoted by SIG, by topic area



Source: Surveys of school administrators in spring 2013.

Note: The total number of practices (shown in parentheses below each bar) differed by topic area. This figure reads as follows (using the first bar on the left as an example): schools implementing a SIG-funded model reported using 90 percent of the practices in the comprehensive instructional reform strategies area.

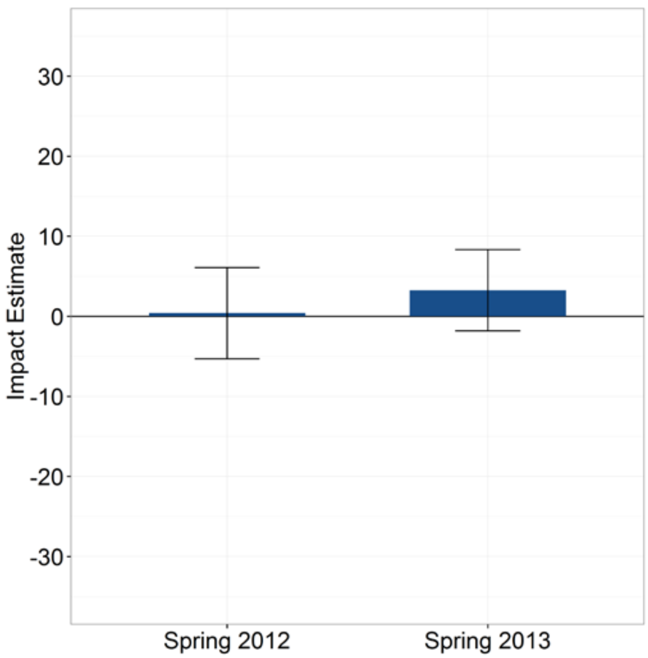
*Significantly different from schools not implementing a SIG-funded model in 2012–2013 at the 0.05 level, two-tailed test.

In addition to the descriptive analysis just described, we also used a more rigorous RDD analysis to examine whether implementing a SIG-funded model had an impact on schools' use of practices. For schools near the SIG eligibility cutoff, we found that the implementation of a SIG-funded model had no significant impact on the total number of SIG-promoted practices used by schools in either spring 2012 or spring 2013 (Figure ES.2). The differences between schools that just met the SIG eligibility criteria and those that just missed the criteria were 0.4 practices (1 percentage point) in spring 2012 and 3.3 practices (9 percentage points) in spring 2013. Although these differences were similar in size to the differences we observed in the descriptive analysis (particularly for 2013), they were not statistically significant. One likely reason why these differences were statistically significant in the descriptive analysis but not in the RDD analysis is that the RDD analysis was less able than the descriptive analysis to detect differences in the number of practices used.[3] Therefore, although our analyses show that schools implementing

[3] The minimum detectable differences for the RDD analysis were 5.8 practices in spring 2012 and 5.2 practices in spring 2013. In contrast, the minimum detectable difference was 0.8 practices in spring 2013 for the descriptive analysis presented in this report.

SIG-funded models used more SIG-promoted practices than other schools, we are unable to conclude that SIG *caused* those observed differences.

Figure ES.2. Impacts of SIG-funded models on number of SIG-promoted practices used



Source: State and district administrative records; surveys of school administrators in spring 2012 and 2013.

Note: Units are the number of practices used, out of 35 practices examined. Black lines show 95 percent confidence intervals. This figure reads as follows (using the first bar on the left as an example): in spring 2012, schools that implemented a SIG-funded model used 0.4 more practices than schools that did not implement such a model, but this difference was not statistically significant. The results shown in this figure were calculated using the RDD methods described in Chapter II and Appendix A.

Across all study schools, use of SIG-promoted practices was highest in the comprehensive instructional reform strategies area and lowest in the operational flexibility and support area

Use of SIG-promoted practices was highest in the comprehensive instructional reform strategies area, in which schools reported using, on average, 7.1 of the 8 SIG-promoted practices examined (89 percent). Use of SIG-promoted policies and practices was lowest in the operational flexibility and support area. In that area, schools reported using, on average, 0.87 of the 2 SIG-promoted practices examined (43 percent).

Across all topic areas, the use of individual practices varied widely. Nearly all study schools reported using benchmark or interim assessments at least once per year (a practice in the comprehensive instructional reform strategies area). In contrast, very few study schools reported (1) using teacher evaluation results to inform decisions about compensation, (2) using principal evaluation results to inform decisions about compensation, or (3) using financial incentives to recruit and retain effective principals (practices in the teacher and principal effectiveness topic area).

In three of four areas, changes over time in use of SIG-promoted practices did not significantly differ between schools implementing a SIG-funded model and schools not implementing one

In three areas—comprehensive instructional reform strategies, teacher and principal effectiveness, and operational flexibility and support—there were no differences between the two groups of schools with respect to changes over time in practices used. In the fourth area—learning time and community-oriented schools—the schools implementing a SIG-funded model reported a decrease of 14 percent of practices between 2011–2012 and 2012–2013, compared to a decrease of 4 percent for schools not implementing such a model.

There were no significant differences in use of ELL-focused practices promoted by SIG between schools implementing a SIG-funded model and schools not implementing one

Both groups of schools (those implementing a SIG-funded model and those not implementing one) reported using 52 percent of the ELL-focused practices examined.

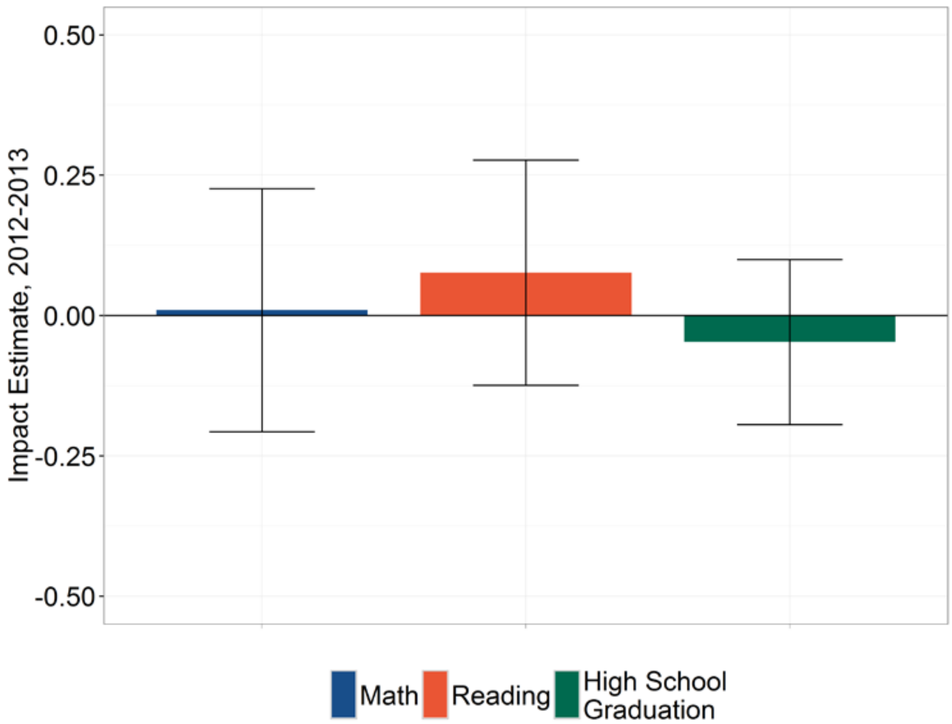
Use of ELL-focused practices did not differ based on the prevalence of ELL students in the school, but SIG-funded model schools with higher ELL achievement gaps used these practices more than schools with lower gaps

The differences in use of ELL-focused practices between schools with higher and lower ELL populations were not significant (0.4 practices among schools implementing a SIG-funded model and 0.3 practices among schools not implementing one). However, among schools implementing a SIG-funded model, schools with higher ELL achievement gaps reported using significantly more ELL-focused practices than schools with lower ELL achievement gaps (0.3 more practices).

SIG-funded models had no significant impact on test scores, high school graduation, or college enrollment

We found no effect of SIG-funded models on student outcomes for schools near the SIG eligibility cutoff. When we examined the impacts of SIG-funded models on math and reading test scores, high school graduation, and college enrollment for 2010–2011, 2011–2012, and 2012–2013, we found no significant impacts (Figure ES.3 shows results for 2012–2013; Appendix A presents results for earlier years [2010–2011 and 2011–2012]). For 2012–2013, the impact on math test scores was 0.01 standard deviations, the impact on reading test scores was 0.08 standard deviations, and the impact on high school graduation was -5 percentage points. We were unable to calculate an impact on college enrollment for 2012–2013 due to insufficient sample sizes, but we found no significant impacts on college enrollment for the other two school years (the impacts for 2010-2011 and 2011-2012 were -11 and 2 percentage points). For all of these student outcomes, we found no significant impacts within student and school subgroups.

Figure ES.3. Impacts of SIG-funded models on student outcomes



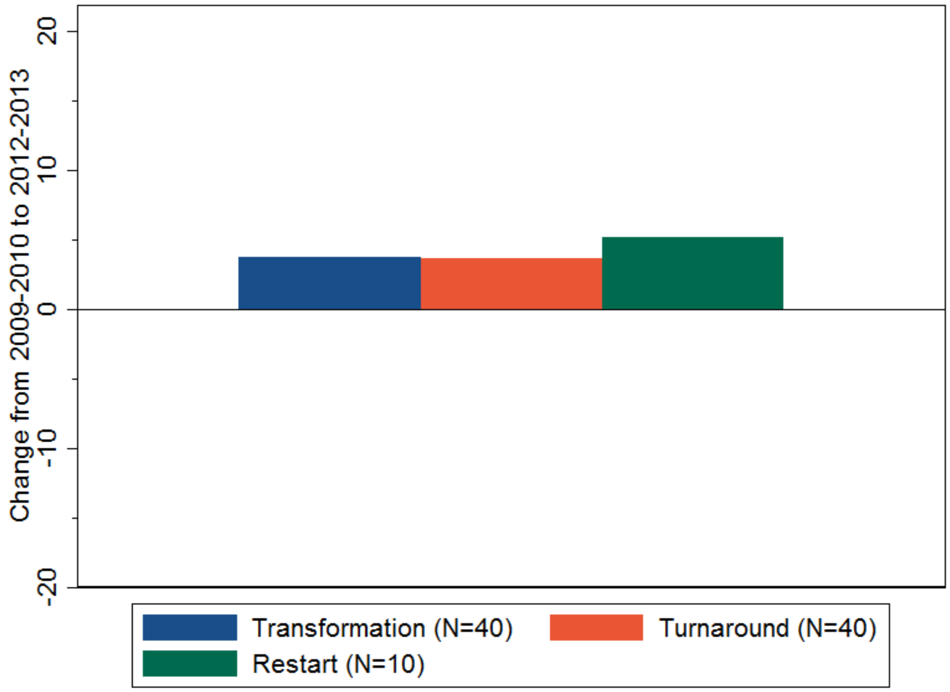
Source: State and district administrative records.

Note: Units for test scores are effect sizes (test scores were standardized to have a standard deviation of 1). Units for high school graduation are percentage points/100. For example, an impact of 0.1 indicates an increase of 10 percentage points. Black bars show 95 percent confidence intervals. The results shown in this figure were calculated using the RDD methods described in Chapter II and Appendix A.

In elementary grades, there was no evidence that one model was associated with larger student achievement gains than another

For elementary grades (2nd through 5th), we found no evidence that one intervention model was associated with larger student achievement gains than another. Between 2009–2010 (the year prior to SIG implementation) and 2012–2013, there were no significant differences in math or reading gains between schools implementing different models (Figure ES.4 presents math results; see Appendix B for reading results). This finding was also true for the two other outcome years we examined (2010–2011 and 2011–2012) and across all sensitivity analyses (see Appendix B).

Figure ES.4. Changes in math test scores in elementary grades, by model



Source: State administrative data.

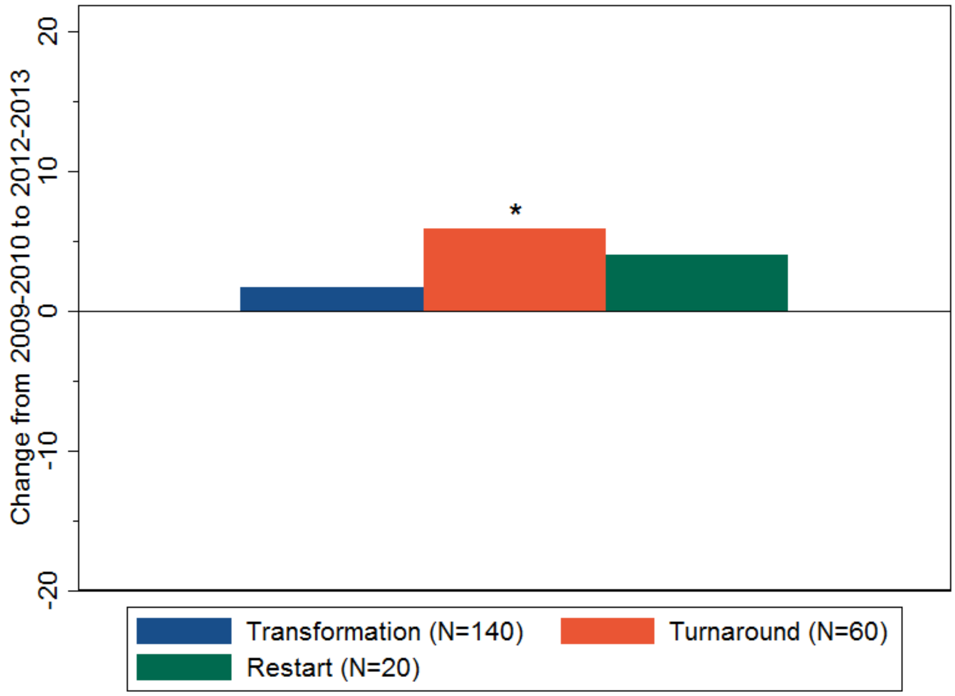
Notes: This figure depicts regression-adjusted changes in math test scores between the baseline year (2009–2010) and 2012–2013 in grades 2 through 5. Changes in math test scores were regression-adjusted for state and grade using a linear model. The key finding (that no model was associated with larger student achievement gains than another) remained the same when we calculated changes in math test scores in a way that accounted for student mobility. Units are normal curve equivalents (NCEs). The NCEs reported in this figure correspond to the following effect sizes (ESs): transformation ES = 0.18, turnaround ES = 0.17, restart ES = 0.25. There were no statistically significant differences between schools implementing different models.

In higher grades, the turnaround model was associated with larger student achievement gains in math than the transformation model

For higher grades (6th through 12th), the implementation of the turnaround model was associated with larger student achievement gains than the transformation model. In particular, between 2009–2010 and 2012–2013, turnaround schools experienced larger gains in math than transformation schools (Figure ES.5 shows math results; see Appendix B for reading results).

However, factors other than the SIG model implemented, such as baseline differences between schools implementing different models, may explain these differences in achievement gains. In particular, turnaround schools served more economically disadvantaged and lower-achieving students at baseline than transformation schools. This finding suggests that turnaround schools may have been fundamentally different from transformation schools prior to SIG, meaning that any number of explanations (other than the model implemented) could account for the different achievement gains.

Figure ES.5. Changes in math test scores in higher grades, by model



Source: State administrative data.

Notes: This figure depicts regression-adjusted changes in math test scores between the baseline year (2009–2010) and 2012–2013 in grades 6 through 12, using changes calculated in a way that accounted for student mobility. Changes in math test scores were regression-adjusted for state and grade using a linear model. When we calculated changes in math test scores in a way that did not account for student mobility, we found that both the turnaround and restart models were associated with larger student achievement gains than the transformation model. Units are normal curve equivalents (NCEs). The NCEs reported in this figure correspond to the following effect sizes (ESs): transformation ES = 0.08, turnaround ES = 0.28, restart ES = 0.19.

*Significantly different from transformation model.

Conclusions

The findings in this report suggest that the SIG program did not have an impact on the use of practices promoted by the program or on student outcomes (including math or reading test scores, high school graduation, or college enrollment), at least for schools near the SIG eligibility cutoff. In higher grades (6th through 12th), the turnaround model was associated with larger student achievement gains in math than the transformation model. However, factors other than the SIG model implemented, such as unobserved differences between schools implementing different models, may explain these differences in achievement gains.

These findings have broader relevance beyond the SIG program. In particular, the school improvement practices promoted by SIG were also promoted in the Race to the Top program. In addition, some of the SIG-promoted practices focused on teacher evaluation and compensation policies that were also a focus of Teacher Incentive Fund grants. All three of these programs involved large investments to support the use of practices with the goal of improving student outcomes. The findings presented in this report do not lend much support for the SIG program having achieved this goal, as the program did not appear to have had an impact on the practices used by schools or on student outcomes, at least for schools near the SIG eligibility cutoff.

However, it is important to keep in mind that our impact estimates only apply to schools near the SIG eligibility cutoff. They correspond to what might be expected if a policy change slightly shifted the cutoff for SIG eligibility, slightly increasing or decreasing the number of schools eligible for SIG funds. We cannot say whether SIG had an impact on use of practices or student outcomes for schools far away from the cutoff.

I. INTRODUCTION

The American Recovery and Reinvestment Act of 2009 (ARRA) provided an unprecedented amount of federal funds for education in an effort to lessen the effects of the nation's economic downturn and to make a lasting investment in education. Through $97.4 billion in ARRA funds, the federal government sought to save education jobs, fund a new wave of innovation in education, and support comprehensive efforts to turn around low-performing schools. The School Improvement Grants (SIG) program received an additional $3 billion through ARRA. Through formula-based grants to states, SIG focused on turning around low-performing schools (formally referred to as "persistently lowest-achieving schools" in SIG guidance) using one of four school intervention models. The SIG application criteria laid out school improvement practices in four main areas: (1) adopting comprehensive instructional reform strategies, (2) developing and increasing teacher and principal effectiveness, (3) increasing learning time and creating community-oriented schools, and (4) having operational flexibility and receiving support.

To learn about the effectiveness of SIG, the U.S. Department of Education's (ED) Institute of Education Sciences (IES) commissioned an evaluation of the program. The SIG evaluation is based on a descriptive analysis of school-level education practices and a regression discontinuity design to assess the effect of SIG on student outcomes and use of SIG-promoted practices.

At the request of the Office of English Language Acquisition (OELA) at ED, part of the evaluation also focuses on how districts and schools have addressed the needs of English language learners (ELLs) as they used the practices promoted by SIG. ELLs are of particular interest to this evaluation because (1) they are historically lower achieving than non-ELLs[4] and (2) the SIG program placed particular emphasis on prioritizing the academic achievement of high-needs students, including ELLs (U.S. Department of Education 2010a).

This is the final report for this evaluation of SIG. Three earlier briefs focused on: (1) implementation of three interrelated levers for school improvement—granting low-performing schools operational authority, supporting them, and monitoring their progress (Herman et al. 2014); (2) low-performing schools' adoption of individual practices and combinations of practices promoted by SIG (Herrmann et al. 2015); and (3) states' capacity to support school turnaround (Tanenbaum et al. 2015). An earlier report covered all major topic areas that SIG promoted, examining the extent to which schools implementing a SIG-funded model and schools not implementing one reported using these practices in spring 2012, and whether use differed across these two groups of schools (Dragoset et al. 2015). This final report builds on the earlier briefs and report by including an additional year of data (spring 2013) and by examining whether receipt of SIG funding had an impact on student outcomes.

In this chapter, we provide background information about the SIG program, present prior research on SIG, and provide an overview of our evaluation and the contents of this report.

[4] Since 2002, ELLs' reading test scores have been below those of non-ELLs on the National Assessment of Educational Progress (National Center for Education Statistics, *The Condition of Education*. Accessed February 17, 2014, at https://nces.ed.gov/programs/coe/indicator_cgf.asp).

A. Scope, purpose, timing, and size of SIG funding

The SIG program aimed to support the implementation of school intervention models in low-performing schools. After the influx of funds from ARRA, SIG awarded grants to six cohorts of schools before it was consolidated into Title I. This evaluation focused on SIG awards granted in 2010 (cohort 1), when roughly $3.5 billion in SIG awards were made, with $3.0 billion of that funding from ARRA. To receive SIG in 2010, state education agencies (SEAs) had to submit applications to ED identifying SIG-eligible schools (based on criteria specified by ED) and specifying the criteria the SEA would use to make SIG subgrants to eligible districts. SIG funds were awarded in grants to states (apportioned by a formula based on Title I allocations). States were then required to distribute 95 percent of those funds through competitive subgrants to local education agencies (LEAs, which are typically school districts) for implementation of school intervention models in eligible schools over the course of three school years starting with 2010–2011.

For the 2010 SIG competition, ED required states to categorize eligible schools into three eligibility tiers based on each school's level (elementary or secondary), eligibility for and receipt of Title I program funds,[5] and achievement or graduation rate. These tiers helped prioritize the distribution of SIG funds at the local level and determined the strategies to be used for school turnaround. Tier I and II schools had to be prioritized over Tier III schools for awards. The SIG eligibility tiers were highly complex (and described in detail in Appendix A). In short, schools eligible for SIG under Tier I and Tier II were generally those who were persistently low performing (in the lowest-achieving 5 percent of schools in the state or high schools with a graduation rate under 60 percent). Tier III schools were also persistently low-performing schools but did not meet the Tier I or II requirements.

ED required SIG-awarded schools under Tier I or Tier II to implement one of four school intervention models, each of which required specific practices (summarized below).

Model requirements	Transformation	Turnaround	Restart	Closure
Replace principal	X	X		
Institute comprehensive instructional reforms	X	X		
Increase learning time	X	X		
Create a community-oriented school	X	X		
Have operational flexibility	X	X		
Adopt a teacher and principal evaluation system that accounts for student growth as a significant factor	X			
Adopt a new governance structure	X			
Replace at least 50 percent of school staff		X		
Convert to a charter school or close and reopen under a charter or education management organization			X	

[5] Title I, Part A (Title I) of the Elementary and Secondary Education Act provides financial assistance to LEAs and schools with many children from low-income families. Title I funds are allocated using formulas based primarily on census poverty estimates and the cost of education in each state.

Model requirements	Transformation	Turnaround	Restart	Closure
Close the school and enroll students in higher-achieving schools in the district				X

The distribution of the 2010 SIG grantees from Tiers I, II, and III across model types and the distribution of award amounts are shown in Table I.1. The maximum grant amount was $2 million per year for three years ($6 million total). The most commonly selected school intervention model was the transformation model (implemented by 50 percent of schools) with a median award per school of $2.1 million in total over three years. The turnaround model was the second most popular school intervention model (14 percent) with a median award of $2.7 million. The restart and closure models were selected for just 3 percent and 1 percent of schools. About a third of schools received awards to implement Tier III strategies, and the median award among those schools was $300,000 (Hurlburt et al. 2011). Federal rules did not require Tier III schools to implement one of the four ED-specified school intervention models. Instead, ED provided districts with the flexibility to decide which strategies to implement in Tier III schools. ED's only requirement was that the strategies should be research based and designed to address the particular needs of each Tier III school.

Table I.1. SIG funding awarded in 2010 and number of schools implementing each intervention model

	School intervention model					
	Transformation	Turnaround	Restart	Closure	Tier III strategies[a]	Total
Number of schools implementing each intervention model						
Tier I	354	138	24	8	0	524
Tier II	255	40	9	8	0	312
Tier III	14	0	0	0	403	417
Total	623	178	33	16	403	1,253
Distribution of award amounts (over three years)						
10th percentile	$942,892	$1,236,632	$1,187,500	$31,935	$60,190	n.a.
50th percentile	$2,100,000	$2,684,490	$2,167,965	$50,000	$300,000	n.a.
90th percentile	$5,114,190	$5,190,000	$5,490,491	$254,323	$900,405	n.a.

Source: IES database of SIG grantees; Hurlburt et al. (2011).

Note: The SIG awards summarized in this table are from the round of state applications due to the U.S. Department of Education on February 8, 2010. The award amount percentiles are based on the total award amount per school.

[a] Tier III strategies refer to all school improvement strategies used by SIG-awarded Tier III schools. Federal rules did not require Tier III schools to implement one of the four ED-specified school intervention models.

n.a. = not applicable.

B. Prior research on SIG documents implementation progress and challenges, and some evidence of effectiveness

Prior research on SIG has focused on two broad categories: (1) documenting the implementation of SIG by states and schools, and (2) assessing the effectiveness of the SIG program. Four themes emerged from the research on SIG implementation:

1. Schools receiving SIG funding generally adopted more SIG-promoted practices than other schools. Several reports—including two from this study—found that use of SIG-promoted practices was higher in SIG schools than other schools, or reported that SIG schools used those practices more intensively than other schools (Dragoset et al. 2015, Herrmann et al. 2014, Center on Education Policy 2012a).
2. Use of some SIG-promoted practices was increasing over time more broadly. Though not specifically about SIG, one study used a nationally representative set of schools to examine reforms that were SIG priorities. The study found an increase in the percentage of schools that reported implementing reforms related to standards and assessments and data systems between 2009–2010 and 2011–2012, but no change in implementation of reforms related to educator evaluation and compensation (Troppe et al. 2015).
3. States reported providing schools and districts with various types of assistance with SIG implementation, but many states also planned to provide some assistance to low-performing, non-SIG schools (Center on Education Policy 2011a, 2012b).[6]
4. Common challenges encountered during SIG implementation included difficulties attracting and retaining high-quality teachers and principals, particularly in rural areas; difficulties using data to inform and differentiate instruction; and limited state capacity to provide assistance (Center on Education Policy 2012a; GAO 2011; U.S. Department of Education 2011a–j).

Two main themes emerged from the research on the effectiveness of SIG:

1. Studies that used more rigorous methods generally found a positive relationship between SIG and student achievement (Dee 2012, Gold et al. 2012, LiCalsi et al. 2015). In one study, this finding was particularly true for schools that implemented the turnaround model (Dee 2012). However, these studies focused on individual states (California and Massachusetts) or cities (Philadelphia), as opposed to examining a broader set of schools from many states.
2. Less rigorous descriptive studies reported improvement among SIG schools nationwide, in math and reading achievement (U.S. Department of Education 2012) and in narrowing achievement gaps between SIG schools and other schools (Council of the Great City Schools 2015).

C. Mixed evidence on relationship between practices promoted by SIG and student outcomes

Though research on SIG is limited, a large body of literature examines the effectiveness of the school improvement practices promoted by SIG and school turnaround more broadly. Overall, this literature provides mixed evidence on whether these practices improve student outcomes. In all four SIG areas (listed above), both experimental and non-experimental studies found mixed results. Some studies found that the practices promoted by SIG in those areas were associated with improved student outcomes, but other studies found no relationship between

[6] Other studies investigated states' methods for selecting SIG schools and their plans for monitoring and supporting them (Center on Education Policy 2011a, 2011b, 2012a, 2012b; Government Accountability Office 2011; Hurlburt et al. 2011, 2012; U.S. Department of Education, Office of the Inspector General 2012).

these practices and student outcomes.[7] Studies on the effectiveness of school turnaround more broadly (excluding SIG-funded intervention models but including school intervention models similar to those promoted by SIG) also found mixed results.[8]

D. Evaluation focus

This evaluation seeks to address gaps in the existing literature by documenting the implementation of SIG-promoted practices by both SIG and non-SIG schools and rigorously assessing the effectiveness of SIG using a large sample of schools from many states. As noted above, few studies on the implementation of SIG-promoted practices examined whether the practices used by SIG schools differed from those used by other schools, and there is no rigorous, large-scale evidence on whether SIG improves student outcomes.

This evaluation examines seven research questions in three broad areas:

Use of SIG-promoted practices

1. Did schools implementing a SIG-funded model use the improvement practices promoted by SIG, and how did that compare to use of those practices by schools not implementing a SIG-funded model?
2. Did receipt of SIG funding to implement a school intervention model have an impact on the number of SIG-promoted practices used by low-performing schools?
3. Did schools' use of SIG-promoted practices change over time?
4. Did use of SIG-promoted practices include a focus on ELLs and did that focus on ELLs differ between schools implementing a SIG-funded intervention model and schools not implementing one?
5. Did use of these ELL-focused improvement practices differ based on the percentage of ELL students in the school or the achievement gap between ELL and other students?

Whether SIG-funded intervention models improved student outcomes

6. Did receipt of SIG funding to implement a school intervention model have an impact on outcomes for low-performing schools?

[7] Examples include Abdulkadiroglu et al. 2011; Allen et al. 2011; Angrist et al. 2011; Betts et al. 2005; Black et al. 2009; Carlson et al. 2011; Clark 2009; Clark et al. 2013; Constantine et al. 2009; Cortes et al. 2012; Decker et al. 2004; Dobbie and Fryer 2011; Furgeson et al. 2012; Garet et al. 2010; Glazerman et al. 2006; Gleason et al. 2010; Henderson et al. 2007; Henderson et al. 2008; James-Burdumy et al. 2005; May and Robinson 2007; Quint et al. 2008; Slavin et al. 2011; and Steinberg 2014.

[8] Examples include Bifulco et al. 2003; Borman et al. 2003; Booker et al. 2009; de la Torre and Gwynne 2009; Dobbie and Fryer 2011; Fryer 2014; Heissel and Ladd 2016; Hoxby et al. 2009; Hoxby and Rockoff 2005; Gleason et al. 2010; Kemple 2015; Player and Katz 2013; Strunk et al. 2012; Tuttle et al. 2013; Zimmer and Buddin 2006; and Zimmer et al. 2012.

Whether the type of model was related to improvements in student outcomes

7. Was the type of school intervention model implemented related to improvement in outcomes for low-performing schools?

To answer the first, third, fourth, and fifth research questions, we conducted descriptive analyses of data from surveys of school administrators. To answer the second and sixth research questions, we used a regression discontinuity design that compared outcomes in schools that just met the SIG eligibility criteria to outcomes in schools that just missed the eligibility cutoff. To answer the seventh research question, we used a correlational analysis that examined differences in outcome gains of schools implementing different intervention models.

The overall sample for the evaluation was purposively selected to support the estimation of impacts of SIG-funded models on student outcomes; it includes 22 states, approximately 60 districts, and approximately 480 schools.[9] Though the results from this evaluation of SIG are not necessarily generalizable to SIG schools nationwide, they are nonetheless important because they add to the limited knowledge base about the implementation and impacts of SIG-funded school turnaround efforts.

E. Report structure

In Chapter II, we describe the study sample, design, and data collected to address these research questions. In Chapter III, we provide baseline information on the SIG sample. In Chapter IV, we present findings on schools' use of SIG-promoted practices. In Chapter V, we examine the change over time in schools' use of SIG-promoted practices. In Chapter VI, we examine whether SIG-funded intervention models improved student outcomes. In Chapter VII, we present findings on the use of SIG-promoted practices related to ELLs. In Chapter VIII, we discuss the findings from this report. In Appendices A through G, we provide additional results, including details on responses to individual survey questions.

[9] Following reporting requirements established by the U.S. Department of Education's National Center for Education Statistics, we rounded all district and school sample sizes to the nearest 10.

II. SAMPLE, DATA, AND METHODS

In this chapter, we describe the study sample, the data collected, and the methods we used to analyze the data for the evaluation.

A. Study sample

We selected the sample for the evaluation with two goals in mind: (1) we wanted a geographically diverse sample and (2) we wanted a sample that would support estimating impacts using a regression discontinuity design (RDD). To efficiently support the RDD analysis, we prioritized states and districts that had the largest number of schools eligible for SIG, and that had a high proportion of SIG-eligible schools actually receiving SIG funds to implement an intervention model. Based on those criteria, we selected a sample of 22 states and approximately 60 districts for the evaluation. Because the sample was not randomly selected, caution should be taken when interpreting the results. In particular, readers should not assume that the findings presented in this report necessarily generalize to SIG schools nationwide.

After selecting the overall sample, we focused on three different subsets of schools within these states and districts (Table II.1 shows how these subsets of schools align with the study's research questions):

- **To describe the use of SIG-promoted practices by schools that implemented a SIG-funded intervention model and schools that did not, we analyzed data from 290 schools that implemented a SIG-funded model in 2012–2013 compared to 190 low-performing schools that did not**.[10] The sample for this analysis included more than a third of all schools nationwide implementing a SIG-funded model in 2012–2013. This sample provides the most comprehensive information on the practices being used in spring 2013.
- **To examine whether SIG-funded models had an impact on the use of SIG-promoted practices and student outcomes, we analyzed data from intervention and comparison groups that were based on SIG eligibility criteria.** The intervention group consisted of 190 schools that were eligible for SIG because they were *below* the SIG eligibility cutoff. These intervention schools represent approximately one-fourth of SIG grant recipients from the 2010 cohort. The comparison group consisted of 270 schools *above* the SIG eligibility cutoff. This sample was used for this analysis so that we could obtain rigorous estimates of the impact of SIG on practices and outcomes. This analysis (as well as the correlational analysis described in the next bullet) focused on SIG-funded models implemented in 2010–2011 because that was the year in which the large ARRA-funded SIG grants were awarded. In contrast, the descriptive analysis in the first bullet focused on SIG-funded models implemented in 2012–2013 because it was designed to provide an overall sense of whether, during the most recent year for which we had data (2012–2013), schools implementing a SIG-funded model were using different practices than other schools.

[10] Low-performing schools (formally referred to as "persistently lowest-achieving schools" in SIG guidance) are generally schools that fell in the lowest 5 percent in academic achievement in the state (or, for high schools, that had a graduation rate lower than 60 percent) for at least two years.

- **To examine the relationship between changes in student outcomes and the type of school intervention model implemented, we analyzed data from 270 schools that implemented a SIG-funded model in 2010–2011**. This analysis excluded schools that did not implement a SIG-funded model because it focused on comparing the outcome changes associated with various models to each other.

Table II.1. Samples of schools used to address research questions

Intervention group sample	Comparison group sample	Research questions addressed with sample	Sample justification
Descriptive analysis			
290 schools implementing a SIG-funded model in 2012–2013	190 schools not implementing a SIG-funded model in 2012–2013	• Did schools implementing a SIG-funded model use the improvement practices promoted by SIG, and how did that compare to use of those practices by schools not implementing a SIG-funded model? • Did schools' use of SIG-promoted practices change over time? • Did use of SIG-promoted practices include a focus on ELLs, and did that focus on ELLs differ between schools implementing a SIG-funded model and schools not implementing one? • Did use of these ELL-focused improvement practices differ based on the prevalence of ELL students in the school or the achievement gap between ELL and other students?	Comparing schools that did and did not implement a SIG-funded model in 2012–2013 (the most recent year we have data) provides the broadest and most recent description of use of practices by the two groups of schools
Regression discontinuity analysis			
190 schools eligible for large ARRA-funded SIG awards in 2010–2011 (in Tiers I and II and below SIG eligibility cutoff)	270 schools not eligible for large ARRA-funded SIG awards in 2010–2011 (in Tier III or not eligible in 2010–2011 and above SIG eligibility cutoff)	• Did receipt of SIG funding to implement a school intervention model have an impact on the number of SIG-promoted practices used by low-performing schools? • Did receipt of SIG funding to implement a school intervention model have an impact on outcomes for low-performing schools?	Focusing on schools that are just on either side of the eligibility cutoff allows us to obtain rigorous estimates of the impact of SIG funding on practices and outcomes
Correlational analysis			
270 schools that implemented a SIG-funded model in 2010–2011	Not applicable	• Was the type of school intervention model implemented related to improvement in outcomes for low-performing schools?	Comparing schools that implemented different SIG-funded models in 2010–2011 allows us to compare changes over time in outcomes for those models

Source: Mathematica Policy Research.

We present information on the baseline characteristics of each sample in Chapter III.

B. Data collection

We used the following data sources for the evaluation:

- **School surveys.** To describe the use of SIG-promoted practices and estimate impacts on their use, we obtained data from web-based surveys of school administrators conducted in spring 2012 and spring 2013. The surveys—sent to the principal of each school—collected information about the SIG models and specific practices reported by schools, as well as supports they reported receiving from states and districts. The SIG objectives in each area and the practices within each area for which we had survey data are detailed in Table II.2.[11,12] The school survey included questions addressing six ELL-focused practices aligned with SIG objectives (Table II.3). Eighty-seven percent of schools in the sample responded to the spring 2012 survey and 93 percent responded to the spring 2013 survey.
- **Common Core of Data (CCD).** To provide baseline and other contextual information, we gathered publicly available data from the CCD, which includes annual data about each public school, local education agency, and state in the country. We obtained 2009–2010 CCD data on school-level characteristics such as total enrollment and the percentages of students in each race/ethnicity category, of students eligible for free or reduced-price lunch, and of schools eligible for Title I.

Table II.2. SIG objectives and practices addressed by school administrator survey questions, by topic area

Topic area	SIG objectives	Practices addressed by school administrator survey questions
Implementing comprehensive instructional reform strategies	Using data to identify and implement an instructional program	• Use data to evaluate instructional programs
	Promoting the continuous use of student data	• Use data to inform instruction • Use benchmark or interim assessments at least annually
	Conducting periodic reviews to ensure that the curriculum is being implemented with fidelity	No items in school survey aligned with this objective[a]
	Implementing a new school model (such as a themed academy)	No items in school survey aligned with this objective[a]
	Providing supports and professional development to staff to assist ELLs and students with disabilities	• Implement strategies for ELLs to master content

[11] The school administrator survey protocols are available at https://www.mathematica-mpr.com/-/media/publications/pdfs/spring_2012_school_administrator_survey.pdf and https://www.mathematica-mpr.com/-/media/publications/pdfs/education/spring_2013_school_administrator_survey.pdf.

[12] The spring 2013 survey questions specified that the 2012–2013 school year was the time period of interest. In particular, the results from spring 2013 represent *only* the practices used during the specific time period between spring 2012 and spring 2013, not the total set of practices used at any point after receipt of SIG funds. Some practices (such as reviewing the strengths and competencies of existing instructional staff and hiring a significant number of new staff) might be one-time events. In these cases, if the school used the practice in the 2011–2012 school year and did not use it again in the following school year, it might have responded “yes” to the question on the spring 2012 survey and “no” on the spring 2013 survey. It is important to keep this in mind when interpreting the findings, because any observed decreases between 2011–2012 and 2012–2013 in the number of practices used may be at least partially because of one-time events in 2011–2012 not needed again the following year. Finally, it is important to keep in mind that because the use of practices is a dynamic process, it is possible that some practices used by schools at the time of the survey were no longer in use after the survey.

Topic area	SIG objectives	Practices addressed by school administrator survey questions
	Using and integrating technology-based supports	• Increase technology access or use computer-assisted instruction
	Tailoring strategies for secondary schools	• Track postsecondary preparation or use project-based learning • Create small learning communities or academies • Track progress to high school graduation
Developing and increasing teacher and principal effectiveness	Using rigorous, transparent, and equitable evaluation systems	• Use student achievement growth • Use multiple evaluation measures
	Identifying and rewarding effective teachers and principals and removing ineffective ones	• Use evaluations to inform compensation • Review competencies of staff or replace instructional staff
	Providing high-quality, job-embedded professional development or supports	• Provide multiple-session professional development events[b] • Provide professional development on Common Core State Standards (CCSS), state standards, or turnaround • Professional development involves working collaboratively or is facilitated by school leaders • Provide professional development on student learning needs • Design professional development with school staff[b] • Use data to evaluate professional development[b]
	Implementing strategies to recruit, place, and retain staff	• Provide financial incentives or flexible work conditions • Use evaluation results to inform reductions in force or have policies that allow principal authority to hire staff[b]
Increasing learning time and creating community-oriented schools	Increasing learning time	• Use schedules or strategies to increase learning time
	Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs	• Change parent or community involvement strategies • Provide professional development on working with parents or cultural sensitivity or increase volunteers or safety measures • Change discipline policies • Use data to guide nonacademic supports
Having operational flexibility and receiving support	Having operational flexibility	• Have autonomy on budgeting, hiring, discipline, or school year length
	Receiving technical assistance and support	• Receive training or technical assistance to support school improvement or use data to improve instruction

Source: SIG application; surveys of school administrators in spring 2012 and spring 2013.

[a] The number of questions included in the school administrator survey was purposefully limited to reduce the time it took to complete the survey. We initially developed the interview questions based on an examination of the SIG application criteria. To ensure that the interview was of a reasonable length, we then pared down the initial list of questions through a deliberative process with the Institute of Education Sciences and the SIG Program Office, to assess their priorities for the types of questions to include. The survey did not include any questions about this objective.

[b] The school administrator survey did not ask about this practice for principals.

ELL = English language learner.

Table II.3. SIG objectives and the ELL-focused practices aligned with those objectives that were addressed by school administrator survey questions, by topic area

Topic area	SIG objectives	ELL-focused practices addressed by school administrator survey questions
Implementing comprehensive instructional reform strategies	Using data to identify and implement an instructional program	No items in school survey aligned with this objective for ELLs[a]
	Promoting the continuous use of student data	• Use data on ELLs to inform instruction
	Conducting periodic reviews to ensure that the curriculum is being implemented with fidelity	No items in school survey aligned with this objective for ELLs[a]
	Implementing a new school model (such as a themed academy)	No items in school survey aligned with this objective for ELLs[a]
	Providing supports and professional development to staff to assist ELLs and students with disabilities	• Implement strategies for ELLs to master content
	Using and integrating technology-based supports	No items in school survey aligned with this objective for ELLs[a]
	Tailoring strategies for secondary schools	No items in school survey aligned with this objective for ELLs[a]
Developing and increasing teacher and principal effectiveness	Using rigorous, transparent, and equitable evaluation systems	No items in school survey aligned with this objective for ELLs[a]
	Identifying and rewarding effective teachers and principals and removing ineffective ones	No items in school survey aligned with this objective for ELLs[a]
	Providing high-quality, job-embedded professional development or supports	No items in school survey aligned with this objective for ELLs[a]
	Implementing strategies to recruit, place, and retain staff	• Offer financial incentives for teachers with ELL expertise • Offer financial incentives for principals with ELL expertise
Increasing learning time and creating community-oriented schools	Increasing learning time	No items in school survey aligned with this objective for ELLs[a]
	Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs	• Provide additional services for ELLs
Having operational flexibility and receiving support	Having operational flexibility	No items in school survey aligned with this objective for ELLs[a]
	Receiving technical assistance and support	• Receive supports to use data on ELLs to improve instruction

Source: SIG application; surveys of school administrators in spring 2012 and spring 2013.

[a] The number of questions included in the school administrator survey was purposefully limited to reduce the time it took to complete the survey. We initially developed the interview questions based on an examination of the SIG application criteria. To ensure that the interview was of a reasonable length, we then pared down the initial list of questions through a deliberative process with the Institute of Education Sciences and the SIG Program Office, to assess their priorities for the types of questions to include. The survey did not include any questions about this objective.

ELL = English language learner.

- **Student-level administrative data.** To estimate the impact of SIG-funded models on student outcomes, we obtained student-level administrative data from states and districts. We examined the following outcomes for the 2010–2011, 2011–2012, and 2012–2013 school years: standardized test scores on state math and reading assessments, high school graduation rates, and college enrollment rates. We used data from 2009–2010 (the year prior to the SIG awards that we focused on in this study) as covariates in our analyses and to compare baseline characteristics of the schools in our sample. We used data from all four school years to examine the relationship between changes in student outcomes and the type of school intervention model implemented.

- **District interviews.** To provide context for the analysis of *schools'* reported use of practices promoted by SIG, we conducted district interviews in spring 2012 and spring 2013 that summarize the extent to which *districts* reported using the practices promoted by SIG.[13] We also obtained school-level budget information during the spring 2012 district interview. We used those data for two purposes: (1) to compare the baseline characteristics of the intervention and comparison schools used in our RDD analyses and (2) to estimate the impact of SIG-funded models on per-pupil spending. The spring 2012 and spring 2013 district interviews had a 100 percent response rate.

C. Analysis methods

In the remainder of the chapter, we describe the methods used to examine the study's research questions (listed in Chapter I).

1. Comparing the practices reported by schools implementing a SIG-funded model and schools not implementing one

In this section, we describe the descriptive methods we used to compare the practices reported by schools implementing a SIG-funded model and schools not implementing one. The purpose of these comparisons is not to determine whether receipt of SIG to implement a school intervention model *caused* schools to use particular practices. Rather, the purpose is to determine whether schools implementing a SIG-funded model used the practices promoted by the four SIG models and how that compares with the use of those practices by schools not implementing a SIG-funded model. This analysis is an important complement to the more rigorous RDD analysis described below that examines the impact of SIG-funded models on practices. In particular, whereas the RDD analysis focuses on schools that are just on either side of the eligibility cutoff, this descriptive analysis looks at the full set of schools that completed surveys for this study. Therefore, this analysis provides the broadest possible examination of the practices being implemented by these two types of schools.

When interpreting the results, please note the following caveat: relative to the RDD findings, this descriptive analysis provides less rigorous evidence on whether implementing SIG-funded models affected the use of SIG-promoted practices. In particular, there could be reasons other than the implementation of a SIG-funded model that explain differences in use of practices that emerge from the descriptive analyses. For this reason, the findings from this analysis should be interpreted with caution.

[13] The implementation data collection focused on the school and district levels because SIG was a school-level intervention and districts played an important role in applying to the state for SIG funds on behalf of their low-performing schools.

Additional caveats to keep in mind are: all of the findings related to use of these practices are based on self-reported data, 2 of the 15 SIG objectives were not addressed by our study instruments, we did not collect information about the quality, fidelity, scope, or intensity with which the practices were implemented, and the sample of schools was not randomly selected. For these reasons, the findings from the analysis of these practices should be interpreted with caution.

This report focuses on broad differences in use of practices by schools that implemented a SIG-funded model and schools that did not, rather than examining whether schools used the practices required by their model more specifically. In particular, the practices that we examined were either required or permissible under the SIG transformation and turnaround models. We chose this focus for several reasons. First, both models prescribed a large set of overlapping practices; all practices that were required or permissible under the transformation model were also permissible under the turnaround model (Appendix F, Table F.1). Second, although schools that implemented the restart model under SIG were required to convert to or close and reopen as a charter school, these schools could still choose to use the practices that were either required or permissible by SIG under the transformation and turnaround models. In addition, only 20 of the 480 schools in the sample implemented the restart or closure model, so it is unlikely that their inclusion would have a substantial effect on the overall results. Finally, another report from this study (Herrmann et al. 2014) already provides more specific information on the implementation of required practices by schools that implemented the transformation and turnaround models.

In this section, we first describe how we formed the two groups that are the basis for the comparisons presented in this analysis. We then describe how we summarized the large number of findings from school surveys and how we analyzed the extent to which schools focused on ELLs in their use of practices promoted by SIG.

a. SIG comparisons

As noted above, in this analysis we compared two groups of schools in the 22 states that had data to support the estimation of impacts of SIG-funded models on student outcomes:

- Schools implementing a SIG-funded model: 290 schools that indicated they received SIG funding and were implementing one of the four school intervention models in spring 2013
- Schools not implementing a SIG-funded model: 190 schools that did not receive SIG funding or received SIG funding but were not implementing one of the four intervention models

The construction of these two groups of schools was driven by the fact that this evaluation is focused on SIG-funded intervention models. More specifically, because this evaluation focuses on Tier I and II SIG schools and the SIG "intervention" for those schools consisted of using SIG funds to implement one of four ED-specified school intervention models, the analysis focuses on comparing schools that implemented a SIG-funded model to schools that did not. We placed 50 schools that received SIG funding but were not implementing a SIG model into the second group (that is, the group of schools not implementing a SIG-funded model) because they would not be expected to have used the practices promoted by the four SIG models. We also placed into this second group 30 schools that reported implementing a SIG model without SIG funding because the goal of this analysis is to compare the practices used by schools implementing SIG-funded models with practices used by similar schools that were not implementing these SIG-funded models.

b. Summarizing findings from the school survey

Given the large number of questions in the survey, it was difficult to discern broad patterns or form overall conclusions by separately examining responses to individual questions. Therefore, we analyzed data from the survey using methods designed to provide information about broad patterns observed in the data. Readers interested in the responses to specific survey questions can refer to Appendix E.

Examining use of practices in spring 2013. To summarize the large amount of data collected, we identified school survey questions that aligned with the practices that SIG sought to affect. Throughout the report, we use the term *SIG-promoted practices* to mean practices aligned with the SIG application criteria. We determined how many practices each school reported using and then calculated the average number of practices for the two groups of schools (those that implemented a SIG-funded model and those that did not). We then tested whether differences in the average number of practices reported were statistically significant between the two groups. Throughout the report, when we say that one group of schools reported using more/fewer practices than the other group within a particular topic or subtopic area, we are always reporting findings that were statistically significant. In contrast, we did not conduct statistical tests to assess whether differences in individual practices between the two groups of schools were statistically significant. See Appendix C for more details on our methods for summarizing schools' use of practices.

Examining changes over time in use of practices. In addition to examining SIG-promoted practices used in 2012–2013, we examined whether use of SIG-promoted practices changed between 2011–2012 and 2012–2013. In Chapter V, we present figures showing the change between the two years in the average number of practices in place for each group of schools (those implementing a SIG-funded model in 2012–2013 and those not implementing one in that year). To conduct this analysis, we averaged the number of practices across each group of schools separately for each school year. We then subtracted the mean number of practices for 2011–2012 from the mean number for 2012–2013 and tested whether the resulting number differed between the two groups of schools. In other words, this analysis tested whether the change between the two years in number of practices used differed between schools implementing a SIG-funded model and schools not implementing one.

This analysis of changes in practices over time focused on approximately 400 of the 480 schools that were included in the analysis of practices reported in 2012–2013. Because this study primarily focused on analyzing the impacts of SIG funds awarded in 2010, the analysis of change in practices over time included schools that received grants in 2010 and continued to implement SIG-funded models for all three years of the grant (2010–2011, 2011–2012, and 2012–2013), compared to non-SIG schools that did not implement a SIG-funded model in any of those three years. As a result, 60 of the 290 schools that implemented a SIG-funded model in 2012–2013 were excluded from the analysis of change in practices over time because they did not implement a SIG-funded model in all three years (2010–2011, 2011–2012, and 2012–2013). In addition, 20 of the 190 schools that did not implement a SIG-funded model in 2012–2013 were excluded from the analysis of change in practices over time because they implemented a SIG-funded model in an earlier year (either 2010–2011 or 2011–2012).

c. ELL-focused analyses of the school surveys

We examined the extent to which schools focused on ELLs in their use of SIG-promoted practices based on the same approach described above for analyzing data from the school survey. The only difference was that the summary measures included only practices that were explicitly focused on ELLs. We also examined whether use of these ELL-focused policies and practices differed by the size of the ELL population and the ELL/non-ELL achievement gap (which we refer to in this report as the *ELL achievement gap* for simplicity). We took the following steps to conduct these analyses:

- We first identified ELL-focused school survey questions that aligned with the practices that SIG sought to affect. We then determined how many ELL-focused practices each school reported using.
- Next, we categorized each school according to whether it had an above-median or below-median ELL population and an above- or below-median ELL achievement gap, where ELL population is defined as the percentage of students who are ELLs. We classified schools as having higher (above-median) or lower (below-median) ELL populations using student-level administrative data from 2009–2010 that contained indicators for whether each student participated in a program for ELLs. We classified schools as having higher or lower ELL achievement gaps based on their gaps on the state math assessment, using student-level administrative data from 2009–2010.[14] Specifically, we calculated the ELL achievement gap as average achievement for non-ELLs minus average achievement for ELLs. We used administrative data from 2009–2010 because it was the year prior to the round of SIG awards on which we focused in this report. To calculate these variables, we first used student-level data to compute the ELL population and the ELL achievement gap for each school in our sample. We then used these school-level values to determine the median ELL population and median ELL achievement gap for the schools in our sample.
- We then examined use of ELL-focused practices for those school groups (above-median ELL population, below-median ELL population, above-median gap, below-median gap). Throughout this report, we use *schools with higher ELL populations* to refer to schools with above-median ELL populations, *schools with higher ELL achievement gaps* to refer to schools with above-median ELL achievement gaps, *schools with lower ELL populations* to refer to schools with below-median ELL populations, and *schools with lower ELL achievement gaps* to refer to schools with below-median ELL achievement gaps. For all the ELL-focused analyses of school survey data, we excluded schools that had no ELLs.
- Finally, we conducted two types of statistical tests:
  - The first was to determine whether there were differences in the number of ELL-focused practices used *between* each group of schools. For example, we compared SIG-funded

[14] The math and reading gaps were highly correlated (0.6 for schools and 0.9 for districts), so the choice of subject was unlikely to make a large difference in the composition of the higher and lower groups.

model schools with lower ELL populations and ELL achievement gaps to schools that also had lower ELL populations and gaps but did not implement a SIG-funded model.

- The second was to determine whether there were differences in the number of ELL-focused practices *within* each group of schools. For example, we compared SIG-funded model schools with lower ELL populations and gaps to SIG-funded model schools with higher ELL populations and gaps.

2. Examining whether SIG affected use of SIG-promoted practices and student outcomes

In this section, we provide a general description of the RDD method, provide an overview of the SIG eligibility tier definitions that support the use of an RDD approach, and summarize the specific RDD approach we used in this study to estimate impacts. Detailed information about the RDD methods we used to estimate impacts is provided in Appendix A.

a. General description of RDD method

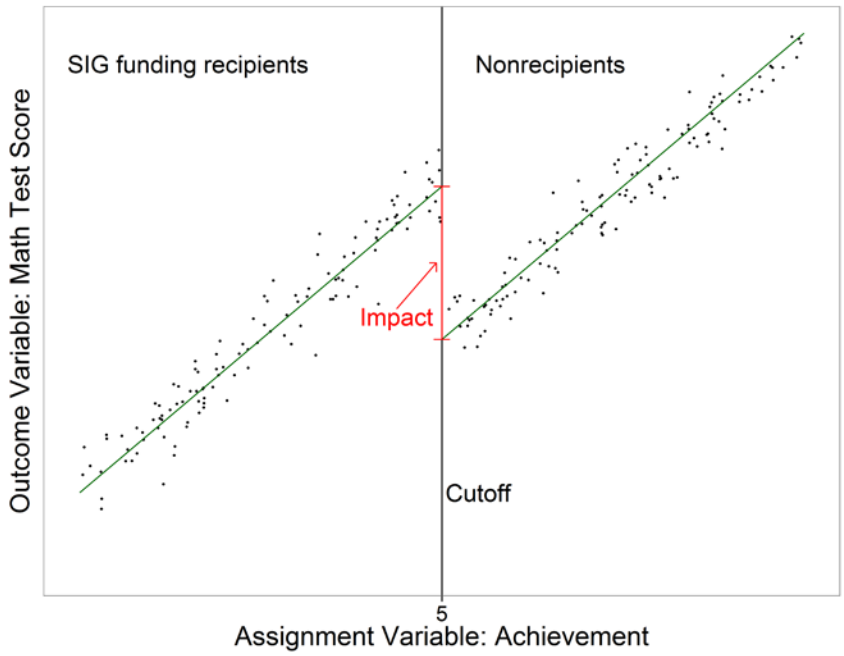
An RDD is possible when a continuous variable, called the *assignment variable*, is used to assign study units (in this case, schools) to receive an intervention. For example, schools with assignment variable values that are below a cutoff value might be assigned to the intervention group and schools with values above the cutoff value might be assigned to the comparison group (that does not receive the intervention). Broadly speaking, those two groups are then compared to estimate the impact of the intervention.

For RDD to be feasible, the cutoff value needs to truly differentiate between schools that do and do not receive the intervention. For example, if the proportion of schools receiving the intervention at a cutoff value on an assignment variable is the same above and below the cutoff, then the cutoff value is not actually differentiating which schools receive the intervention, and RDD would not be feasible. However, it is not necessary for *all* schools below the cutoff to receive the intervention, and for *no* schools above the cutoff to receive it (which would be a sharp RDD). An RDD can still be feasible even if not all schools below the cutoff received the intervention and if some schools above the cutoff did (which would be a fuzzy RDD).

Figure II.1 illustrates the RDD graphically using a hypothetical example. In this example, schools are ranked according to their school-level average achievement; schools with achievement at the fifth percentile or below receive SIG funds, and schools with achievement above the fifth percentile do not. This figure plots student outcomes (math test scores one year after SIG was implemented in intervention schools) against the assignment variable (baseline school-level achievement). Each dot represents a school. The two lines show the relationship between the outcome and the assignment variable for the intervention and comparison groups. The estimated impact on math test scores is the vertical distance between the two lines at the cutoff value of 5 (that is, the fifth percentile of school-level average achievement).

Similar to a randomized controlled trial (RCT), the RDD is a rigorous design that enables us to establish whether SIG-funded models caused schools to use SIG-promoted practices and caused student achievement to improve. Unlike an RCT, in which the estimated impact of the intervention applies to all schools in the study, the RDD impact estimate applies only to schools near the cutoff value of the assignment variable. Therefore, this estimate does not necessarily represent the impact of the intervention on schools far away from the cutoff value of the assignment variable.

Figure II.1. Hypothetical example of the regression discontinuity method



Source: Simulated data.

b. SIG eligibility tier definitions that support an RDD

The definitions of Tier I and Tier II SIG eligibility provide the opportunity to use an RDD for this evaluation. In particular, those definitions involve a cutoff on school-level achievement at the fifth percentile. Schools with achievement below that cutoff (Tier I and Tier II schools) form the intervention group, and schools on the other side of the cutoff (Tier III schools and non-SIG-eligible schools) form the comparison group. The difference in outcomes for schools just above and just below the cutoff is the RDD impact.

SIG eligibility rules also involve a cutoff on the graduation rate. We decided not to use the graduation rate variable in our benchmark approach, based on the results of diagnostic analyses showing evidence of discontinuities in the density of the graduation rate at the RDD cutoff value. However, as a sensitivity analysis, we estimated impacts that include the graduation rate cutoff and found that our results and conclusions did not change (see Appendix A for more details).

The intervention in this evaluation is defined as receiving SIG funds for implementing one of the four school intervention models specified by ED (which we determined based on information from the school administrator survey and the administrative data). As noted above, to use RDD to examine this intervention, the RDD cutoff value needs to truly differentiate receipt of the intervention between two groups of schools. In this study, we find this to be the case for two reasons. First, Tier I and II schools had to be prioritized over Tier III schools when awarding SIG funds, so a substantially higher proportion of schools in Tiers I and II received

SIG grants than those in Tier III. Second, ED required each Tier I or II school that received a SIG grant to implement one of the four ED-specified school intervention models, whereas Tier III schools were not required to do so.

The RDD we used was a fuzzy RDD, meaning that not all schools below the cutoff received SIG funds to implement an intervention model, and some schools above the cutoff did. In the intervention group, 85 percent of schools below the cutoff received SIG funds to implement an intervention model. In the comparison group, just 10 percent of schools received SIG funds to implement an intervention model; these were Tier III schools that implemented a model even though they were not required to do so.

The SIG eligibility tier definitions create eight opportunities to use an RDD. The tier definitions and RDD opportunities that they create are highly complex. For readers interested in these details, please see Appendix Table A.1 for the full definitions of the SIG eligibility tiers and Appendix Table A.2 for details on the eight RDD opportunities we considered (including the intervention and comparison groups and applicable RDD assignment variable for each opportunity).

For readers who prefer a higher-level summary of the tiers, we provide that here. There were two different sets of tier definitions: the original definitions published in the *Federal Register* on December 10, 2009, and a set of expanded tier definitions from the Appropriations Act, which was signed into law on December 16, 2009. Both sets of definitions focused on persistently low-performing schools, broadly defined as follows:

- The original Tier I and II definitions focused on schools with achievement in the lowest 5 percent in the state or with graduation rates below 60 percent; the original Tier III definition focused on low-performing Title I schools in improvement status that were above those cutoffs. Although cutoffs for SIG eligibility could involve school achievement and graduation rates, our benchmark RDD analyses used only school achievement as an assignment variable.
- The expanded tier definitions increased the set of low-performing schools that could receive SIG. The expanded Tier I and II definitions permitted awards to schools with achievement in the lowest 20 percent in the state or that had not made adequate yearly progress for two consecutive years, and the expanded Tier III definition permitted awards to low-performing schools that did not meet the expanded Tier I or II requirements.

The eight RDD opportunities can be summarized as follows: five of the eight opportunities involve the achievement assignment variable and three involve the graduation rate assignment variable, six of the eight opportunities involve secondary schools and two involve elementary schools, and half of the opportunities involve Tier I schools as the intervention group and the other half involve Tier II schools as the intervention group.

As noted above, the RDD estimate does not necessarily represent the impact of the intervention on schools far away from the cutoff value of the assignment variable. In the context of this study, this means that we are focused on the impact on outcomes for schools that were on the cusp of being included in SIG eligibility Tiers I and II. This impact corresponds to what might be expected if a policy change expanded or contracted Tiers I and II through a small increase or decrease in the cutoff on school-level achievement used to define Tiers I and II.

c. Our methods for estimating RDD impacts

The RDD component of this study can be characterized as a set of several separate grade-specific RDD analyses that we aggregated to obtain the overall estimate of the impact of SIG-funded models. For each grade and each outcome, we first standardized the outcome variable by subtracting the mean and dividing by the standard deviation of the outcome variable, using state-grade-level means and standard deviations. We then estimated a separate RDD impact for each grade and outcome. We calculated the overall impact of SIG-funded models on each outcome as a weighted average of the grade-specific impacts.

The main impact findings presented in this report are based on the study team's benchmark approach for estimating impacts. The approach follows the What Works Clearinghouse (WWC) evidence standards for regression discontinuity designs (U.S. Department of Education 2010b, 2015) and was selected before data analysis began. It was chosen because it performed better in simulations than alternative approaches. Consistent with the WWC RDD standards, we also conducted several diagnostic analyses to assess the rigor of the RDD and examine whether the underlying assumptions required in an RDD held in this study. In Appendix A, we describe the benchmark approach, the smallest impacts that the benchmark approach could detect, the alternative approaches that we used to assess the sensitivity of the findings, and the robustness of the study's findings to those methods.[15]

In addition to estimating impacts for each grade, we calculated impacts on student outcomes (math and reading test scores, high school graduation, and college enrollment) separately for each of the following policy-relevant subgroups:

- ELLs and non-ELLs
- Elementary and secondary schools
- Title I-receiving secondary schools in improvement, corrective action, or restructuring; and secondary schools that were eligible for, but did not receive, Title I funds[16]

[15] We carried out two types of analyses that handled student mobility in different ways. In the first type (the benchmark analysis presented in the body of the report), we included only students who were present at the end of the school year when standardized tests were administered. This impact estimate was place-based because it analyzed the students who were actually present in each place (school) at the time tests were administered, rather than students who were slated to attend each school. In the second analysis (a sensitivity test presented in Appendix A), we controlled for student mobility by focusing on the students who were slated to attend each school (based on the school they attended in the baseline year and typical school feeder patterns in the district) rather than the school they actually attended. If the intervention affected student mobility systematically, then the place-based impact estimates comingle impacts on student achievement with impacts on mobility, but the impact estimates from the sensitivity analysis isolate the impacts on student achievement. Details of these two approaches are provided in Appendix A.

[16] These subgroups were of interest because each one represented a separate RDD opportunity with distinct intervention and comparison groups under the original tier definitions. The first group—Title I-receiving secondary schools in improvement, corrective action, or restructuring—were divided into Tier I (the intervention group) and

- Schools in early Race to the Top (RTT), later RTT, and non-RTT states[17]

The smallest impacts that the benchmark approach could detect were within the range of impacts on academic achievement measures from past studies. In a review of meta-analyses, Hill (2008) found mean effect sizes ranging from 0.22 to 0.27 standard deviations across grade levels. The smallest impacts our benchmark approach could detect ranged from 0.19 to 0.22 standard deviations for test score outcomes, from 0.15 to 0.26 standard deviations for high school graduation, and from 0.27 to 0.39 standard deviations for college enrollment.

3. Examining whether the type of school intervention model implemented was related to changes in outcomes

In this analysis, we focused on examining the relationship between the type of school intervention model implemented and changes in student outcomes. Ideally, one would like to obtain rigorous evidence on whether certain intervention models are more effective than others at improving student achievement. However, it is not possible to answer that question using a rigorous approach. An RCT is not possible because schools are not randomly assigned to implement particular models, and an RDD is not feasible because we do not know which model schools in the comparison group would have implemented if they had received SIG funding. Therefore, we conducted a correlational analysis to examine the relationship between the type of model a school implemented and the changes in student outcomes experienced by that school.

This type of correlational analysis cannot conclusively establish which models are most effective at improving student achievement. Therefore, interpreting the results requires caution: specific school intervention models might not have caused the observed changes in outcomes. It is possible that factors other than the model implemented, such as baseline differences between schools implementing different models, may explain any differences in achievement gains that we observed.

In this section, we first describe our main analysis which analyzed changes over time in math and reading test scores for schools implementing different models. We then describe two additional analyses we conducted to examine whether factors other than the model implemented may explain changes in test scores over time: an analysis of changes in the composition of students attending schools implementing different models and an analysis of baseline characteristics of schools implementing different models. Detailed explanations of our analysis methods, including regression models and a full description of the sensitivity analyses conducted, appear in Appendix B.

Tier III (the comparison group). The second subgroup—secondary schools that were eligible for, but did not receive, Title I funds—were divided into Tier II (the intervention group) and SIG-ineligible schools (the comparison group).

[17] Early RTT states received grants in the first two rounds of the competition. Round 1 winners were announced on March 29, 2010, and round 2 winners were announced on August 24, 2010. Later RTT states received grants in the third round. Round 3 winners were announced on December 22, 2011. Non-RTT states did not receive grants. The RDD analysis sample includes 7 of the 12 early RTT states, 4 of the 7 later RTT states, and 10 of the 32 non-RTT states, so the RTT versus non-RTT analysis is not fully representative of all the RTT and non-RTT states.

a. Main analysis of changes in math and reading test scores

In our main analysis, for each school that implemented a SIG-funded model, we examined how changes in math and reading test scores over time were correlated with specific models (transformation, turnaround, and restart). We excluded schools that implemented the closure model from our main analysis because it examined math and reading test scores in years after SIG implementation began—that is, after these schools had closed.[18] Our analyses (see Chapter VI for findings) focus on the change in test scores between the baseline year—that is, the 2009–2010 school year, the year prior to SIG implementation—and the latest outcome year for which we had data (2012–2013). This analysis focused on students who attended the schools in 2009–2010, compared to students who attended the schools in 2012–2013. (Results for the other outcome years, 2010–2011 and 2011–2012, are shown in Appendix B.)

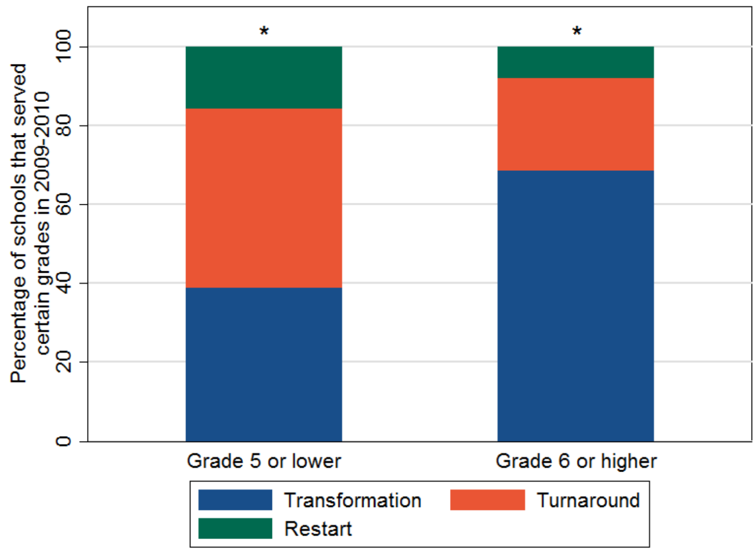
We examined the relationship between test scores and model separately for two groups of schools: schools that included grade 6 or any higher grade (up through grade 12) and schools that did not (that is, schools that only included grade 5 or lower). We conducted the analysis in this way for two reasons. First, we found that schools in these two grade ranges implemented a different mix of models. For example, turnaround model schools made up the largest percentage of schools serving grade 5 or lower, while transformation model schools made up the largest percentage of schools serving grade 6 or higher (Figure II.2). Second, we found that the relationship between test scores and model differed between these two groups, so we concluded that analyzing them separately was the most appropriate approach.

b. Analysis of changes in the composition of students attending schools implementing different models

Because school intervention models might influence which schools students attend, we examined changes in the composition of students attending schools implementing different models. As noted above, our main analysis focused on students who attended the study schools in 2009–2010, compared to students who attended the study schools in 2012–2013. Because of this, the main analysis comingles effects of intervention models on the academic outcomes of individual students with effects on the composition of the students attending these schools. If we found a positive relationship between implementation of a particular model and improved student achievement, it could be for either of two reasons. First, implementation of that model might improve student achievement because schools implementing that model are more effective at educating children. Second, implementation of that model might change which students attend the schools implementing that model because, for example, more motivated parents might send their children to these schools even if the schools are not any better at educating children.

[18] However, we did include closure schools in all tables and figures that show *baseline characteristics* (that is, characteristics from the 2009–2010 school year, prior to SIG implementation) to provide an overall sense of what schools implementing each of the four models (transformation, turnaround, restart, and closure) looked like at baseline and how they compared to each other at baseline.

Figure II.2. Percentages of schools that served certain grades, by model



Source: State administrative data.

Notes: This figure depicts regression-adjusted percentages of schools implementing different models that served grade 5 or lower and grade 6 or higher at baseline (the 2009–2010 school year). Each bar adds to 100 percent. Percentages were regression-adjusted for state using a logit model.

*Percentages of schools serving grade 5 or lower and grade 6 or higher were significantly different across models.

To investigate whether the second reason might help explain any changes in test scores we might observe in the main analysis, we analyzed changes in the composition of students attending schools implementing different models. In particular, we analyzed whether the student body composition of schools implementing different models changed over time with respect to several variables, including the percentage of students eligible to receive free or reduced-price lunch and the average achievement level of the students, as measured by math test scores from 2009–2010 before SIG funding was received. For example, for schools implementing the transformation model, we calculated the difference in 2009–2010 math test scores between students attending those schools in 2009–2010 and students attending those same schools in 2012–2013. We did the same for schools implementing the turnaround model. We then tested whether the change in average achievement levels between the two years was statistically significantly different for transformation and turnaround schools.

In these analyses, we found some compositional changes over time between models. For example, we found that disadvantaged students left restart schools between 2009–2010 and 2012–2013 in greater proportions than they left schools implementing the transformation model. To address this issue, we conducted a sensitivity analysis to determine whether and how much of the differences in outcome changes between models that we observed could be due to student mobility. The sensitivity analysis involved re-estimating our main model using outcome changes calculated in a way that accounted for student mobility. We present the results of this sensitivity analysis in Chapter VI. Appendix B describes this sensitivity analysis, as well as others we conducted, in more detail.

c. Analysis of baseline characteristics of schools implementing different models

Because different types of schools might choose to implement different models, we investigated whether there were differences in the baseline characteristics of schools depending on which model they were implementing. This analysis is important, because differences in outcomes over time might not be due to the model being implemented but rather could be due to pre-existing differences between schools implementing different models. We examined several baseline characteristics in 2009–2010, including the percentage of schools serving grade 6 or higher (shown in Figure II.2 above), the percentage of students eligible for free or reduced-price lunch, and average baseline test scores.

III. UNDERSTANDING THE CONTEXT FOR SCHOOL IMPROVEMENT GRANTS

As we described in Chapter II, the analyses in this report used different samples of schools from 22 states and approximately 60 districts to address three broad categories of research questions. Each analysis used the sample of schools that best addressed its goals. In this chapter, we present information about the samples of states, districts, and schools used in this report. To interpret the findings, it is important to understand how these samples compare to broader sets of states, districts, and schools in the United States, as well as how the intervention and comparison groups within each analysis compare. The first comparison indicates whether findings from this study might generalize to states, districts, and schools nationwide. The second comparison gives insight into whether baseline differences between the intervention and comparison schools could explain any differences in outcomes observed between the two groups. In Section A, we compare the baseline characteristics of study states and districts with those of all states and districts in the United States that had schools implementing SIG-funded models. In Section B, we compare the baseline characteristics of the intervention schools in each analysis with those of the comparison schools and those of all schools in the United States that were implementing SIG-funded models.

A. Baseline characteristics of states and districts

We compared the states and districts where our study schools were located to all states and districts in the United States that had schools implementing a SIG-funded model in 2010–2011. (All states had schools implementing a SIG-funded model.) For these comparisons, we examined characteristics for the 2009–2010 school year, the year prior to the SIG awards that are the focus of this study. These comparisons indicate whether our findings might generalize more broadly to states and districts nationwide.

1. Study states and all states in the United States had similar baseline characteristics

The characteristics of the states in our study did not differ significantly from those of all states in the United States (Table III.1). The lack of significant differences between these two groups suggests our findings might generalize to states nationwide.

Table III.1. Baseline (2009–2010) characteristics of study states and all states

	Study States	All States[a]
Average percentage of students by racial/ethnic category		
White, non-Hispanic	55.3	61.8
Black, non-Hispanic	19.5	15.8
Hispanic	18.3	13.7
Asian	3.8	4.6
Other	3.1	4.1
Percentage of students eligible for free or reduced-price lunch	48.0	45.5
Percentage of schools that are Title I eligible	68.1	67.8
Percentage of schools by location:		

	Study States	All States[a]
Urban	30.0	23.3
Suburban	25.7	22.5
Town	14.3	16.0
Rural	30.0	38.2
Number of States	**22**	**51**

Source: Common Core of Data, 2009–2010.

Note: Data from 2008–2009 were used for states with data missing in 2009–2010. Data from 2007–2008 were used for states with data missing in both 2009–2010 and 2008–2009. Data from 2009–2010 were used whenever possible because that was the school year just before the first year of implementation of the ARRA-funded SIG models. Percentages of students and schools are unweighted state-level averages. There were no statistically significant differences between study states and all states at the 0.05 level using a two-tailed test.

[a] Includes 50 states and the District of Columbia, all of which contained schools implementing a SIG-funded intervention model in 2010–2011.

2. Study districts had lower percentages of white students and were more likely to be urban than all districts in the United States that had schools implementing a SIG-funded model

Unlike study states, study districts did significantly differ from districts nationwide that had schools implementing a SIG-funded model, in terms of students' race and school location (Table III.2). For example, the districts in our study had a lower percentage of students who were non-Hispanic white (19.5 versus 33.8 percent) and had schools that were more likely to be located in an urban area (68.2 versus 39.6 percent) than the group of districts nationwide. This suggests that our findings may not necessarily generalize to districts nationwide.

Table III.2. Baseline (2009–2010) characteristics of study districts and all U.S. districts with schools implementing a SIG-funded model

	Study Districts	Districts in the United States with at Least One School Implementing a SIG-Funded Intervention Model in 2010–2011
Average percentage of students by racial/ethnic category		
White, non-Hispanic	19.5*	33.8
Black, non-Hispanic	38.7	30.3
Hispanic	32.0	24.4
Asian	3.3	2.7
Other	6.5	8.9
Percentage of students eligible for free or reduced-price lunch	72.4	68.3
Percentage of schools that are Title I eligible	81.4	81.3
Percentage of districts by location:		
Urban	68.2*	39.6
Suburban	17.3	18.7
Town	5.7	12.0
Rural	8.8*	29.7

	Study Districts	Districts in the United States with at Least One School Implementing a SIG-Funded Intervention Model in 2010–2011
Number of Districts	**60**	**420**

Source: Common Core of Data, 2009–2010; IES database of SIG grantees.

Note: Data from 2008–2009 were used for districts with data missing in 2009–2010. Data from 2007–2008 were used for districts with data missing in both 2009–2010 and 2008–2009. Data from 2009–2010 were used whenever possible because that was the school year just prior to the first year of implementation of the ARRA-funded SIG models. Percentages of students and schools are unweighted district-level averages. The percentages of districts with at least one school implementing a SIG-funded model are based on schools' planned implementation as of 2009–2010 for cohort 1 grantees, and only include Tier I and Tier II schools.

*Significantly different from districts in the United States with at least one school implementing a SIG-funded intervention model in 2010–2011 at the 0.05 level, two-tailed test.

B. Baseline characteristics of schools used in the analyses

The analyses in this report used different samples of schools from these states and districts to best address their research question (see Table II.1). For each analysis sample, we conducted two comparisons of baseline characteristics: we compared the intervention and comparison schools (where applicable), and we compared the intervention schools implementing a SIG-funded intervention model to all schools in the United States that were implementing these models.

1. Intervention and comparison schools had similar baseline characteristics

For both the descriptive and RDD analyses, we found no evidence that baseline differences could explain any differences in observed outcomes between the intervention and comparison group schools.

In the descriptive analysis, schools implementing a SIG funded model in 2012–2013 and those not implementing one had similar characteristics at baseline. The two groups of schools had similar student and school demographics, such as race/ethnicity, percentage of students eligible for free or reduced-price lunch, Title I eligibility, location, and school level (elementary, middle, high school) (Table III.3).

As expected, intervention schools differed from comparison schools on characteristics related to the intervention. Schools implementing a SIG-funded model in 2012–2013 were more likely than schools not implementing one to be planning to implement one of the four models at the time of the SIG application, to be eligible for SIG under Tier I or II (which had higher priority for SIG awards than Tier III schools), and to be in SIG cohorts 1, 2, or 3.

Table III.3. Baseline (2009–2010) characteristics of schools used in the descriptive analysis and of all U.S. schools implementing a SIG-funded intervention model

	Study Schools Implementing a SIG-Funded Intervention Model in 2012–2013	Study Schools Not Implementing a SIG-Funded Intervention Model in 2012–2013	All U.S. Schools Implementing a SIG-Funded Intervention Model in 2012–2013
Characteristics from 2009–2010 Common Core of Data			
Average percentage of students by racial/ethnic category			
White, non-Hispanic	8.8*	8.7	19.0
Black, non-Hispanic	54.8*	50.4	45.4
Hispanic	31.4	34.9	27.5
Asian	1.9	2.1	2.1
Other	3.0*	3.9	6.0
Average percentage of students eligible for free or reduced-price lunch	82.8*	80.9	77.3
Percentage of Title I eligible schools	94.4*	94.3	90.5
Percentage of schools by location			
Urban	87.5*	87.6	58.4
Suburban	6.6*	5.7	17.0
Town or rural	5.9*	6.7	24.7
Percentage of schools by level			
Elementary	29.5	32.0	27.1
Middle	19.8	17.5	20.1
High	49.0	45.9	46.2
Other	1.7*	4.6	6.6
Eligibility tier and planned intervention model at time of SIG application			
Percentage of schools by model			
Transformation	57.3†*	8.2	73.3
Turnaround	34.3†*	5.7	20.5
Restart or closure[a]	8.4†	3.1	6.1
Percentage of schools by SIG cohort			
Cohort 1	82.7†*	8.8[b]	57.4
Cohort 2	12.8†*	2.1[b]	34.7
Cohort 3	4.5†*	0.0	8.0
Percentage of schools by eligibility tier			
Tier I	66.2†	21.1	63.8
Tier II	22.9†*	11.2	35.1
Tier III	10.9†*	67.7	n.a.
Number of Schools	**270–290**	**160–190**	**1,420–1,450**

Source: Common Core of Data, 2009–2010; IES database of SIG grantees; surveys of school administrators in spring 2012 and spring 2013; interviews with district administrators in spring 2012 and spring 2013.

Note: Percentages of students are unweighted school-level averages. U.S. schools implementing a SIG-funded model were restricted to schools in Tiers I and II because ED required that each Tier I or II school receiving SIG implement one of four models (whereas Tier III schools receiving SIG were not required to do so), so schools in Tiers I and II are more similar to the group of study schools implementing a SIG-funded model than Tier III schools are. The national percentages of schools implementing each of the four models are based on schools' planned implementation as of 2009–2010 for cohort 1 grantees and as of 2010–2011 for cohort 2 grantees. Data from 2009–2010 were used whenever possible because that was the school year just before the first year of implementation of the ARRA-funded SIG models. Data from 2008–2009 were used for schools with data missing in 2009–2010, and data from 2007–2008 were used for schools with data missing in both 2009–2010 and 2008–2009. To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for town and rural school locations and for restart and closure models. Ranges are provided for the sample sizes because missing data varied across items.

[a] Schools that had already implemented the closure model as of spring 2013 were not surveyed and were not included in the analysis. Schools that were planning to implement the closure model but had not yet closed as of spring 2013 were surveyed and included in the analysis, for reasons described in Appendix C.

[b] Cohort 1 and cohort 2 schools that were not implementing a SIG-funded model in 2012–2013 are schools that replied "no" to either question TA1 on the school survey (which asked whether the school received SIG funds for school improvement efforts in the current school year), or question TA7 (which asked whether the school was using one of the four ED-specified models), or both. (Note that a school had to reply "yes" to both questions to be considered implementing a SIG-funded model in our analysis.) For the schools for which we had information from another source (specifically, the district interview), that source corroborated the information provided in the school survey.

†Significantly different from study schools not implementing a SIG-funded intervention model in 2012–2013 at the 0.05 level, two-tailed test.

*Significantly different from schools in the United States implementing a SIG-funded intervention model in 2012–2013 at the 0.05 level, two-tailed test.

IES = Institute of Education Sciences; n.a. = not applicable.

In the regression discontinuity analysis, there were no differences in baseline characteristics between the intervention and comparison groups at the cutoff values. Our RDD analysis involved comparing schools just above and below the SIG eligibility cutoff for the assignment variable. To confirm that these comparisons were valid, we examined whether the intervention and comparison schools at the cutoff had similar baseline characteristics. More specifically, we compared the average baseline characteristics for intervention and comparison schools after adjusting for the assignment variable. The results show that intervention and comparison schools at the cutoff did indeed have similar baseline characteristics (Table III.4).

As expected, intervention schools were significantly more likely than comparison schools to implement a SIG-funded intervention model in 2010–2011, which was the SIG award year of interest for the RDD analysis (Table III.4). In 2010–2011, 85 percent of intervention schools implemented a SIG-funded model compared to 10 percent of comparison schools. Since not all schools below the cutoff implemented the intervention and some schools above the cutoff did, the RDD analysis we conducted was a *fuzzy RDD*.

Table III.4. Baseline characteristics of schools in the RDD analysis

	Intervention Schools	Comparison Schools
Baseline characteristics in 2009–2010		
Math achievement	-0.78	-0.72
Reading achievement	-0.72	-0.77
Percentage of students eligible for free or reduced-price lunch	76.8	80.1
Percentage of students who are English language learners	15.0	17.5
Percentage of students who are white	15.8	12.1
Percentage of schools that implemented a SIG-funded model in 2010–2011	85.0*	9.9

	Intervention Schools	Comparison Schools
Number of schools	**190**	**270**
Number of students attending schools	**145,270**	**152,000**

Source: State and district administrative records.

Notes: Administrative data from 2009-2010 were aggregated to the school level, using the same set of students who were in the benchmark impact analysis for each school. The units for math and reading achievement are effect sizes (test scores were standardized to have a standard deviation of 1). After adjusting for the assignment variable, there were no systematic differences in baseline characteristics between the intervention and comparison groups at the cutoff values (there was a systematic difference in the percentage of schools implementing a SIG-funded model, as expected).

*Significantly different from comparison schools at the 0.05 level, two-tailed test.

2. Intervention schools were more disadvantaged and more likely to be urban than SIG schools nationwide

For all three analysis samples, we compared the characteristics of the intervention schools with those of all schools in the United States that were implementing SIG models. We found that the intervention schools were not representative of SIG schools nationwide. In particular, they had more economically disadvantaged students, were more likely to be located in urban areas, and were generally more likely to be implementing the turnaround model than all SIG schools in the United States. This pattern suggests that the findings in this report may not necessarily generalize to SIG schools nationwide.

In the descriptive analysis, study schools implementing a SIG-funded model in 2012–2013 differed from U.S. schools implementing such models in 2012–2013 on nearly all of the baseline characteristics examined. For example, study schools implementing a SIG-funded model in 2012–2013 were significantly more likely than U.S. schools implementing such models to be economically disadvantaged (with 82.8 versus 77.3 percent of students eligible for free or reduced-price lunch) and located in an urban area (87.5 versus 58.4 percent) (Table III.3). Study schools implementing a model were significantly less likely than SIG schools nationwide to be implementing a transformation model (57.3 versus 73.3 percent) and more likely to be implementing a turnaround model (34.3 versus 20.5 percent).

In the RDD analysis, intervention schools were not comparable to all schools in the United States that were implementing SIG-funded models in 2010–2011. For example, RDD intervention schools were significantly more likely than SIG schools nationwide to be economically disadvantaged (with 86.3 versus 78.3 percent of students eligible for free or reduced-price lunch) and located in an urban area (82.8 versus 58.2 percent) (Table III.5). Because the RDD focused on an achievement assignment variable (and achievement scores were more commonly available in grades 3–8), RDD intervention schools were also less likely to be high schools than SIG schools nationwide (38.5 versus 48.6 percent). In addition, RDD intervention schools were significantly less likely to be implementing a transformation model than SIG schools nationwide (45.3 versus 73.8 percent).

In the correlational analysis, schools in the analysis sample implementing a SIG-funded model in 2010-2011 differed from U.S. schools implementing SIG-funded models in that year on nearly all of the baseline characteristics examined. For example, similar to the findings from the other analyses, schools in the correlational analysis were significantly more likely than SIG schools nationwide to be economically disadvantaged (with 84.3 versus 78.3 percent of students eligible for free or reduced-price lunch) and located in an urban area (87.7 versus 58.2 percent) (Table III.5). Among both schools in the correlational analysis and SIG schools nationwide, the most popular model was the transformation model, followed by the turnaround, restart, and closure models. However, the correlational analysis included higher percentages of turnaround and restart schools than in schools implementing SIG-funded models nationwide.

Table III.5. Baseline (2009–2010) characteristics of RDD intervention schools, schools in the correlational analysis, and of all U.S. schools implementing a SIG-funded intervention model in 2010–2011

	RDD Intervention Schools	Schools in the Correlational Analysis	All U.S. Schools Implementing a SIG-Funded Intervention Model in 2010–2011
Characteristics from 2009–2010 Common Core of Data			
Average percentage of students by racial/ethnic category			
White, non-Hispanic	9.0*	9.1*	19.3
Black, non-Hispanic	54.0*	54.7*	45.0
Hispanic	30.7	30.9	26.7
Asian	2.2	1.9	2.2
Other	4.2	3.4*	6.8
Average percentage of students eligible for free or reduced-price lunch	86.3*	84.3*	78.3
Percentage of Title I eligible schools	94.3*	94.4*	89.3
Percentage of schools by location			
Urban	82.8*	87.7*	58.2
Suburban	10.4*	6.3*	16.3
Town or rural	6.8*	5.9*	25.4
Percentage of schools by level			
Elementary	37.0*	32.0*	24.3
Middle	20.8	19.7	19.8
High	38.5*	46.1	48.6
Other	3.6	2.2*	7.2
Eligibility tier and planned intervention model at time of SIG application			
Percentage of schools by model			
Transformation	45.3*	59.1*	73.8
Turnaround	26.0	27.9*	20.0
Restart	-[a]	9.7*	4.1
Closure	-[a]	3.3	2.1
Restart or closure	4.2	-[a]	6.2
Percentage of schools by eligibility tier			
Tier I	63.1	68.7*	62.1
Tier II	29.1*	27.6*	37.9

	RDD Intervention Schools	Schools in the Correlational Analysis	All U.S. Schools Implementing a SIG-Funded Intervention Model in 2010–2011
Tier III	7.8*	3.7*	n.a.
Number of Schools	**180–190**	**270**	**810–820**

Source: Common Core of Data, 2009–2010; IES database of SIG grantees; state and district administrative records.

Note: Percentages of students are unweighted school-level averages. U.S. schools implementing a SIG-funded model were restricted to schools in Tiers I and II because ED required that each Tier I or II school receiving SIG implement one of four models (whereas Tier III schools receiving SIG were not required to do so), so schools in Tiers I and II are more similar to the RDD intervention schools than Tier III schools are. The national percentages of schools implementing each of the four models are based on schools' planned implementation as of 2009–2010. Data from 2009–2010 were used whenever possible because that was the school year just before the first year of implementation of the ARRA-funded SIG models. Data from 2008–2009 were used for schools with data missing in 2009–2010, and data from 2007–2008 were used for schools with data missing in both 2009–2010 and 2008–2009. To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for town and rural school locations. Ranges are provided for the sample sizes because missing data varied across items.

[a] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for the restart and closure models for RDD intervention schools.

*Significantly different from schools in the United States implementing a SIG-funded intervention model in 2010–2011 at the 0.05 level, two-tailed test.

IES = Institute of Education Sciences; n.a. = not applicable.

IV. SCHOOLS' USE OF PRACTICES PROMOTED BY SCHOOL IMPROVEMENT GRANTS

To understand the extent to which a grant program like SIG might improve student achievement, it is important to first understand whether schools that implemented a SIG-funded intervention model used the practices promoted by the program. If these schools used the same practices as similar schools that did not implement a SIG-funded model, it is unlikely that any changes in outcomes for SIG schools—positive or negative—could be attributed to the program.

In this chapter, we assess the extent to which schools implementing a SIG-funded model and those not implementing one reported using SIG-promoted practices. We conducted two separate analyses—each using different analysis methods and samples—to examine schools' use of these practices. Both analyses address the same basic research question (whether schools implementing a SIG-funded model used more SIG-promoted practices than schools not implementing one), and each analysis has unique advantages. Because the analyses complement each other, examining the results of both enables a fuller understanding of the answer to the research question.

First, we conducted a descriptive analysis that compared the use of SIG-promoted practices in schools that implemented a SIG-funded model in 2012–2013 and in schools that did not.[19] This analysis was designed to provide a sense of whether, at a particular point in time (spring 2013), schools implementing a SIG-funded model were using school improvement practices to a greater extent than other schools. It cannot conclusively establish whether receipt of SIG to implement a school intervention model *caused* schools to use SIG-promoted practices. Factors other than implementing a SIG-funded model, such as unobserved differences between the two groups of schools, may explain observed differences in the use of SIG-promoted practices. However, the advantage of this analysis is that it sheds light on the extent to which a large group of low-performing schools used SIG-promoted practices during the 2012–2013 school year.

Second, we used a regression discontinuity design (RDD) to examine whether implementation of a SIG-funded model in 2010–2011 had an impact on schools' use of SIG-promoted practices in spring 2012 and spring 2013. The advantage of this analysis is that RDD is a rigorous design that enables us to determine whether SIG-funded models *caused* schools to use SIG-promoted practices. However, in contrast to the descriptive analysis above, the RDD impacts apply only to schools near the SIG eligibility cutoff—that is, schools near the lowest 5 percent of achievement in the state; we do not know whether the impacts of SIG that we present in this chapter also apply to schools far from these cutoff values.

[19] The analyses presented in this report do not distinguish between required and permissible practices; according to the SIG application criteria, required practices are those that schools implementing a particular SIG model *must* use, and permissible practices are those that schools implementing a particular SIG model *may* use. For a detailed examination of whether low-performing schools adopted the practices that were required and/or permissible under the transformation and turnaround models, please see Herrmann, M., L. Dragoset, and S. James-Burdumy. "Are Low-Performing Schools Adopting Practices Promoted by School Improvement Grants?" NCEE 2015-4001. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, October 2014.

In Sections A through D of this chapter, we present findings from the descriptive analysis for each of the four topic areas promoted by SIG. In Section E, we summarize those findings. In Section F, we present findings from the RDD analysis. In Section G, we compare and contrast the findings from the two analyses. Appendix A provides more details about the RDD analysis methods and findings. Appendix C provides more detailed findings from the descriptive analysis, such as the use of practices by subtopic.

To provide context for school reports about SIG-promoted practices used, Appendix D presents findings on the extent to which *districts* reported using SIG-promoted practices. In some cases, schools may use SIG-promoted practices because districts use them or require schools to use them. For example, districts might require multiple performance measures for teacher and principal evaluations or provide additional supports and programs to students with disabilities. All districts in the study sample included schools that were and were not implementing a SIG-funded intervention model, so districts' use of SIG-promoted practices might reduce differences in use of practices by the two groups of schools in the same district.

A. Schools implementing a SIG-funded model used more comprehensive instructional reform strategies than schools not implementing one

One goal of SIG is to promote the use of instructional practices that have the potential to increase academic rigor and student achievement. The SIG application criteria focused on practices to reform instruction in seven subtopics: (1) Using Data to Identify and Implement an Instructional Program; (2) Promoting the Continuous Use of Data to Identify and Address the Needs of Individual Students; (3) Conducting Periodic Reviews of the Curriculum; (4) Implementing a New School Model; (5) Providing Supports and Professional Development (PD) to Staff to Assist Both English language learners (ELLs) and Students with Disabilities; (6) Using and Integrating Technology-Based Supports; and (7) Tailoring Strategies for Secondary Schools. We collected data on five of these subtopics through school survey questions that asked about eight practices aligned with SIG objectives in these areas (Table IV.1). Because none of the questions from the school surveys aligned with the third or fourth subtopic, we excluded these subtopics from the analysis.[20]

[20] To limit the survey length, we sought input from IES and the SIG program office on which questions were of greatest interest.

Table IV.1. Practices aligned with SIG objectives on comprehensive instructional reform strategies, by subtopic

Using data to identify and implement an instructional program
Using data to evaluate instructional programs (for example, measuring program effectiveness)
Promoting the continuous use of student data
Using data to inform and differentiate instruction
The typical English/language arts or math teacher used benchmark or interim assessments at least once per year
Providing supports and professional development to staff to assist ELLs and students with disabilities
Implementing strategies (including additional supports or professional development) to ensure that students with limited English proficiency acquire language skills to master academic content
Using and integrating technology-based supports
Increased access to technology for teachers, or the typical English/language arts teacher used computer-assisted instruction
Tailoring strategies for secondary schools
Secondary school monitored students' college readiness (such as enrollment in Advanced Placement courses), including providing supports (such as project-based learning) so that low-achieving students can take advantage of these types of opportunities[a]
The school was divided, or grades within the secondary school were subdivided, into small learning communities or field/career-oriented academies[a]
Secondary school tracked student progress toward (and readiness for) high school graduation[a]

Source: SIG application; surveys of school administrators in spring 2012 and 2013.

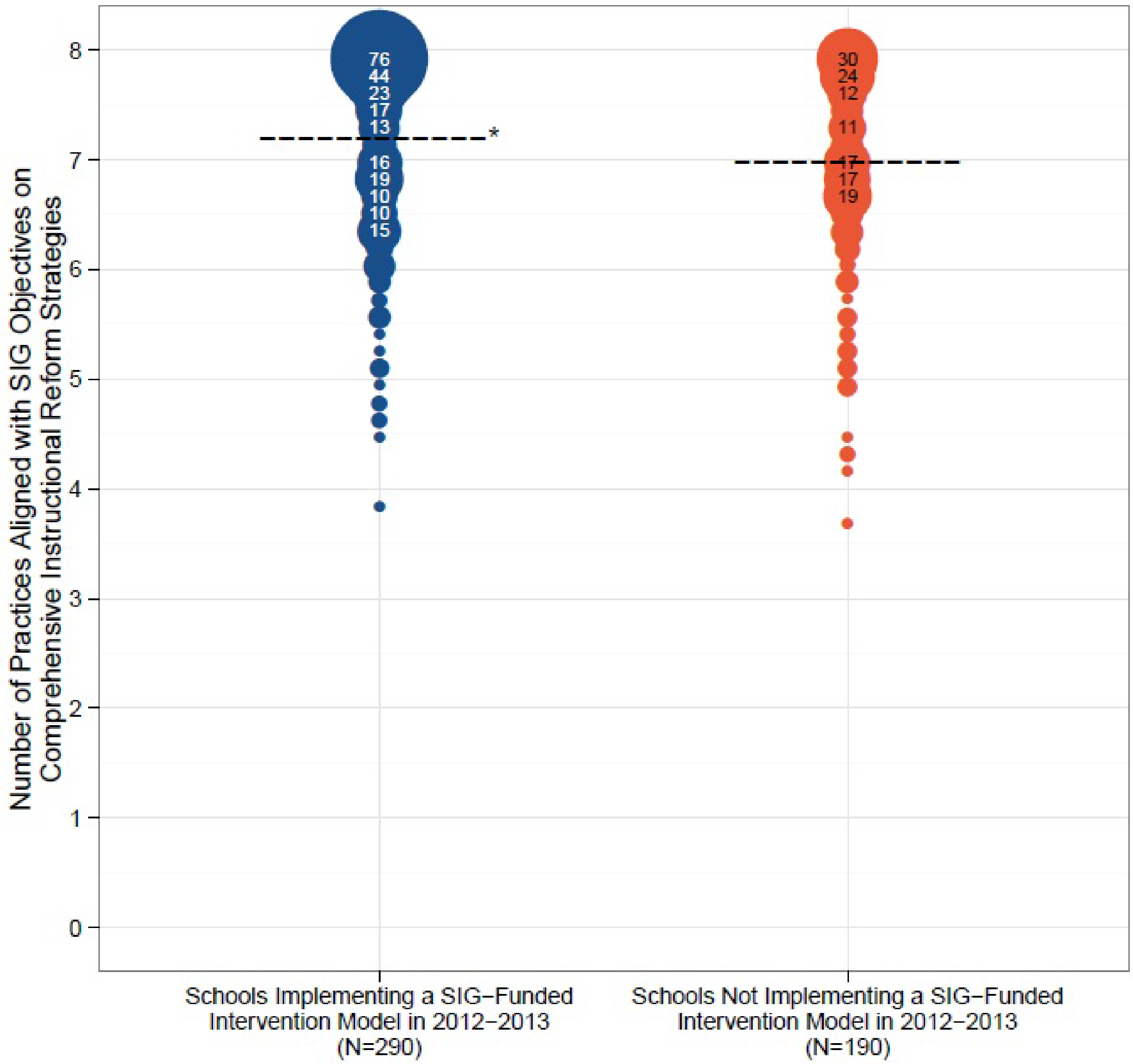
Note: See Appendix F for a list of survey questions that aligned with the SIG practices in this table.

[a] As we described in Appendix C, for a school (in this case, an elementary school) that might be missing values for a practice, we multiplied the mean of the nonmissing practices by the total number of practices for the overall topic area. For example, for the comprehensive instructional reform strategies topic area, which has eight practices, if a school had data available for five practices and reported using two of those, the number of the school's reported practices would be equal to (2/5)*8.

ELLs = English language learners.

In spring 2013, schools implementing a SIG-funded model reported using more SIG-promoted practices in the comprehensive instructional reform area than schools not implementing such a model. Schools in the first group reported using an average of 7.2 of 8 practices in this area, compared to 7.0 for schools in the second group (Figure IV.1). Although this difference was statistically significant, it was small (just 0.2 practices), so it is unclear whether it is substantively important. Many schools in each group reported using nearly all of these practices, and few reported using less than half of the practices.

Figure IV.1. Use of practices aligned with SIG objectives on comprehensive instructional reform strategies



Source: Surveys of school administrators in spring 2013.

Note: Table IV.1 lists the practices summarized in this figure. Each dot in this figure represents the schools that reported using a particular number of practices (out of eight examined) that aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For example, 16 schools implementing a SIG-funded intervention model reported using 7 of the 8 comprehensive instructional reform practices aligned with the SIG application criteria. For two of the practices, a "yes" response received one point. In the other six cases, a school could receive a fraction of one point. See Appendix C for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

In spring 2013, for one of the five comprehensive instructional reform subtopics, schools implementing a SIG-funded model reported using more SIG-promoted practices than schools not implementing one. For the subtopic focused on using and integrating technology-based supports, 90 percent of schools implementing a SIG-funded model reported using the practice in this subtopic, compared to 78 percent of schools not implementing such a model (Appendix C, Figure C.4).

The practice in this area with the largest difference between the two groups of schools was tracking student progress toward (and readiness for) high school graduation. On average, schools implementing a SIG-funded model used 89 percent of the components of this practice and schools not implementing one used 77 percent (Figure IV.2).[21,22]

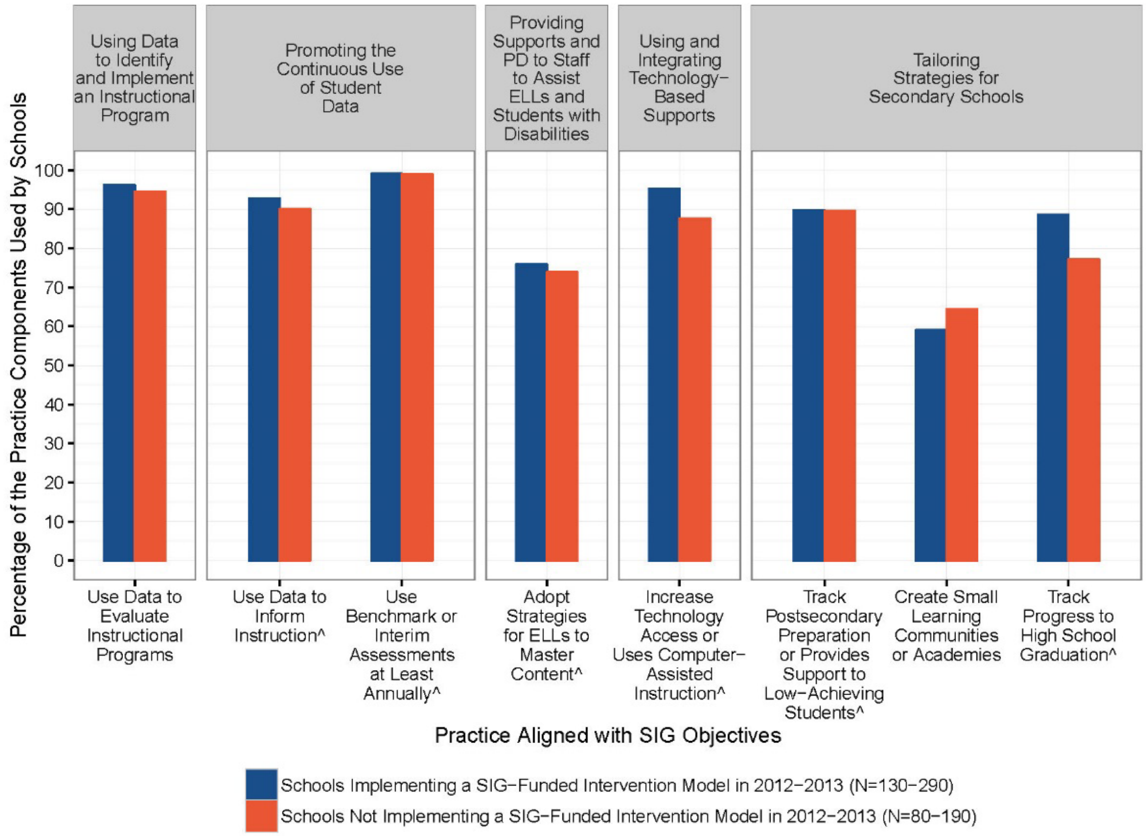
Most practices in this area were used by nearly all study schools, including (1) using data to evaluate instructional programs, (2) using data to inform instruction, (3) using benchmark or interim assessments at least once per year, (4) increasing technology use or using computer-assisted instruction, (5) tracking postsecondary preparation or providing support to low-achieving students, and (6) tracking progress toward (and readiness for) high school graduation (Figure IV.2).

---

[21] For readers interested in specific practices where large differences were reported, we show descriptive findings on individual practices with the largest differences between schools implementing a SIG-funded intervention model and schools not implementing one. We did not conduct statistical tests to assess whether differences in individual practices between schools implementing a SIG-funded model and schools not implementing one were statistically significant.

[22] As described in Appendix C, for each practice in the SIG application for which we identified one or more relevant survey questions, we used those questions to calculate the percentage of questions to which each school responded "yes." This variable measures the percentage of the components of the practice that each school used. We then calculated the average percentage across all schools. Some practices were based on multiple survey questions (rather than a single survey question), so each school could use less than 100 percent of the components of these practices. For example, if eight separate questions constituted a particular practice, a school received 12.5 percent, or one-eighth of 100 percent, for each "yes" response. If half of the schools responded "yes" to all eight questions and the other half responded "yes" to none of the eight questions, we would say that schools, on average, used 50 percent of the components of the practice. For practices that were addressed by a single survey question, we indicate the *percentage of schools* that used each practice. For practices that were addressed by multiple survey questions, we indicate the average *percentage of the components* of each practice that schools used.

Figure IV.2. Use of individual practices aligned with SIG objectives on comprehensive instructional reform strategies



Source: Surveys of school administrators in spring 2013.

Note: This figure has a separate panel for each subtopic. As described in Appendix C, for each practice in the SIG application criteria for which we identified one or more survey questions aligned with the practice, we calculated the percentage of survey questions with a "yes" response as a measure of the percentage of components each school used. The height of each bar represents the average percentage of the components of the practice that each group of schools used. A range is provided for the sample sizes because nonresponse varied across items.

^Multiple survey questions were used to assess whether schools used all of the components of this practice.

ELLs = English language learners; PD = professional development.

B. Use of practices in the teacher and principal effectiveness area was higher among schools implementing a SIG-funded model

The SIG program encouraged schools and districts receiving grants to increase the capacity and quality of their teachers and principals to improve student outcomes. Specifically, the SIG application criteria focused on practices to develop and increase teacher and principal effectiveness in four subtopics: (1) Using Rigorous, Transparent, and Equitable Evaluation Systems; (2) Identifying and Rewarding Effective Teachers and Principals and Removing Ineffective Ones; (3) Providing High-Quality, Job-Embedded Professional Development or Supports; and (4) Implementing Strategies to Recruit, Place, and Retain Staff. The evaluation's school surveys asked about 20 practices aligned with SIG objectives in this topic area (Table IV.2).

Table IV.2. Practices aligned with SIG objectives on teacher and principal effectiveness, by subtopic

Teacher effectiveness
Using rigorous, transparent, and equitable evaluation systems
Student achievement growth was a required component of teacher evaluations, and the extent to which student achievement growth must factor into teacher evaluations, or state test scores were used to assess student growth for teacher evaluations was specified
Using multiple performance measures for teacher evaluations
Identifying and rewarding effective teachers and removing ineffective ones
Using teacher evaluation results to inform decisions about compensation
Reviewing the strengths and competencies of instructional staff for the purposes of hiring or removing staff
Providing high-quality, job-embedded professional development or supports
Providing instructional staff with PD that consisted mostly or entirely of multiple-session events
Providing instructional staff with PD focused on transitioning to Common Core State Standards, aligning instruction to state standards, or strategies for turning around a low-performing school
Providing staff with PD that involved educators working collaboratively or was facilitated by school leaders or coaches
Providing staff with PD that was focused on understanding and addressing student learning needs (including reviewing student work and achievement data, and collaboratively planning, testing, and adjusting instructional strategies based on data)
Providing staff with PD designed with input from school staff
Using data to evaluate the success of PD offerings
Implementing strategies to recruit, place, and retain staff
Implementing strategies, such as financial incentives or more flexible work conditions, designed to recruit, place, and retain staff
Using teacher evaluation results as the primary consideration in reductions in force and excessing decisions, or having teacher assignment policies that allow for principal discretion in which staff to hire for the school
Principal effectiveness
Using rigorous, transparent, and equitable evaluation systems
Measures of student achievement growth were used for principal evaluations and the extent to which student achievement growth must factor into principal evaluations was specified
Using multiple performance measures for principal evaluations
Identifying and rewarding effective principals and removing ineffective ones
Using principal evaluation results to inform decisions about compensation
School has a new principal
Providing high-quality, job-embedded professional development or supports
State or district provides the principal or other school leaders with PD on analyzing and revising budgets or strategies for turning around a low-performing school
State or district provides the principal or other school leaders with PD on identifying effective instructional staff for leadership positions and supporting them in these positions
State or district uses principal evaluation results to develop the principal's PD or provides the principal with PD on aligning teachers' PD with evaluation results
Implementing strategies to recruit, place, and retain staff
Principals have the opportunity to receive financial incentives designed to recruit, place, and retain staff

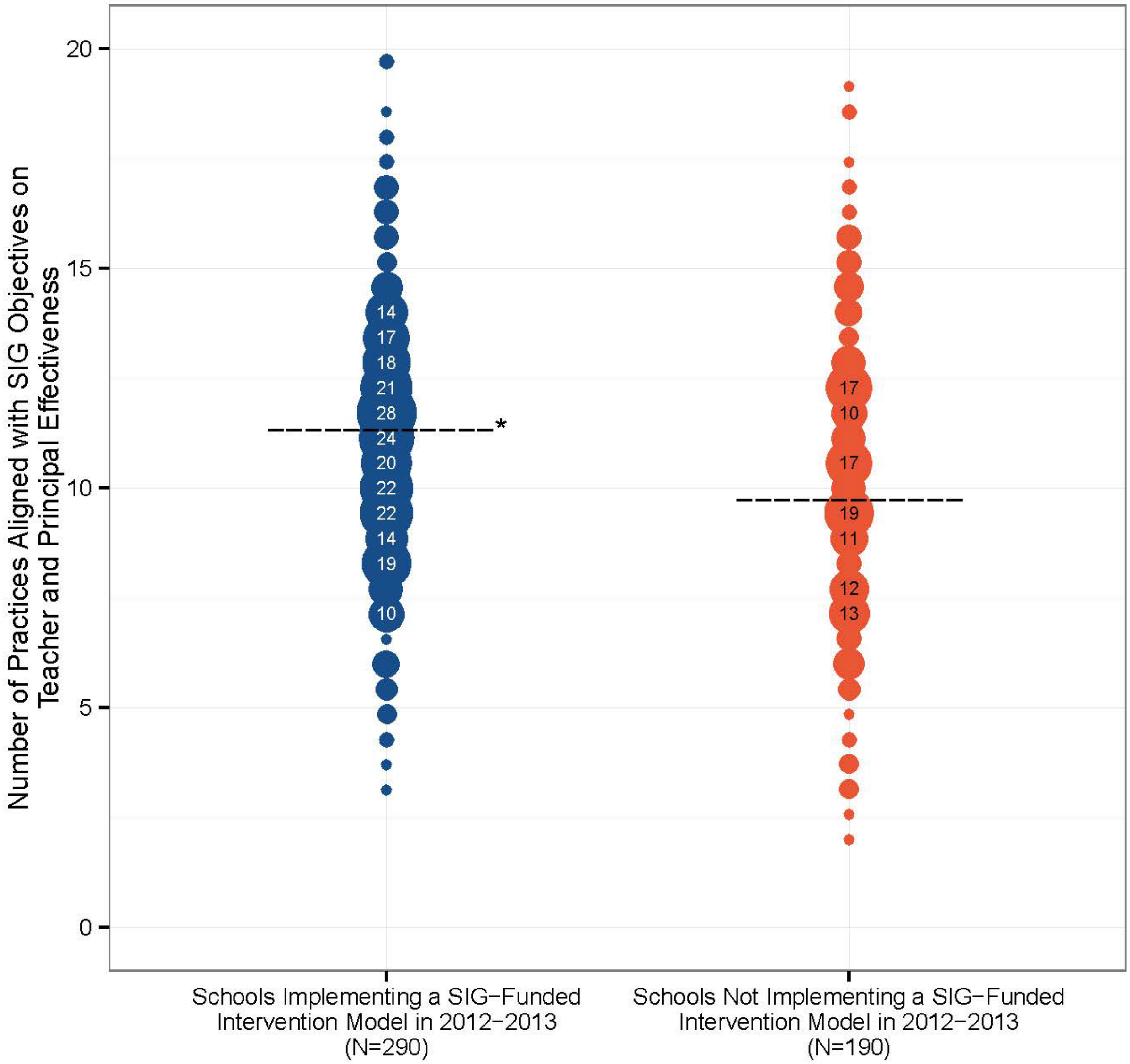
Source: SIG application; surveys of school administrators in spring 2012 and spring 2013.
Note: See Appendix F for a list of survey questions that aligned with the SIG practices in this table.
PD = professional development.

Schools implementing a SIG-funded model reported in spring 2013 that they used more SIG-promoted practices in the teacher and principal effectiveness area than schools not implementing one. Schools in the first group reported using an average of 11.3 of 20 practices in this area, compared to 9.7 for schools in the second group, a difference of 1.6 practices (Figure IV.3).

For all four teacher and principal effectiveness subtopics, we found significant differences in spring 2013 between the two groups of schools:

- **Using Rigorous, Transparent, and Equitable Evaluation Systems.** Schools implementing a SIG-funded model reported using an average of 2.3 of 4 practices in this subtopic, compared to 2.1 for schools not implementing such a model (see Appendix C, Figure C.6).
- **Identifying and Rewarding Effective Teachers and Principals and Removing Ineffective Ones.** Similarly, schools in the first group reported using an average of 1.5 of 4 practices in this subtopic, compared to 1.2 for schools in the second group (see Appendix C, Figure C.7).
- **Providing High-Quality, Job-Embedded Professional Development or Supports**. In this subtopic, schools implementing a SIG-funded model indicated that they used an average of 6.5 of 9 practices, relative to 5.7 practices for schools not implementing a SIG-funded model (see Appendix C, Figure C.8).
- **Implementing Strategies to Recruit, Place, and Retain Staff.** Schools implementing a SIG-funded model reported that they used 1.0 of 3 practices on average, compared to schools not implementing a SIG-funded model, which reported using 0.8 practices (see Appendix C, Figure C.9). Despite the relatively small (0.2) difference in this subtopic overall, it included the practice with the largest difference between the two groups of schools across all four topic areas examined—reviewing competencies of staff or replacing instructional staff. On average, schools implementing a SIG-funded model used 43 percent of the components of this practice, and schools not implementing one used 27 percent (Figure IV.4).

Figure IV.3. Use of practices aligned with SIG objectives on teacher and principal effectiveness



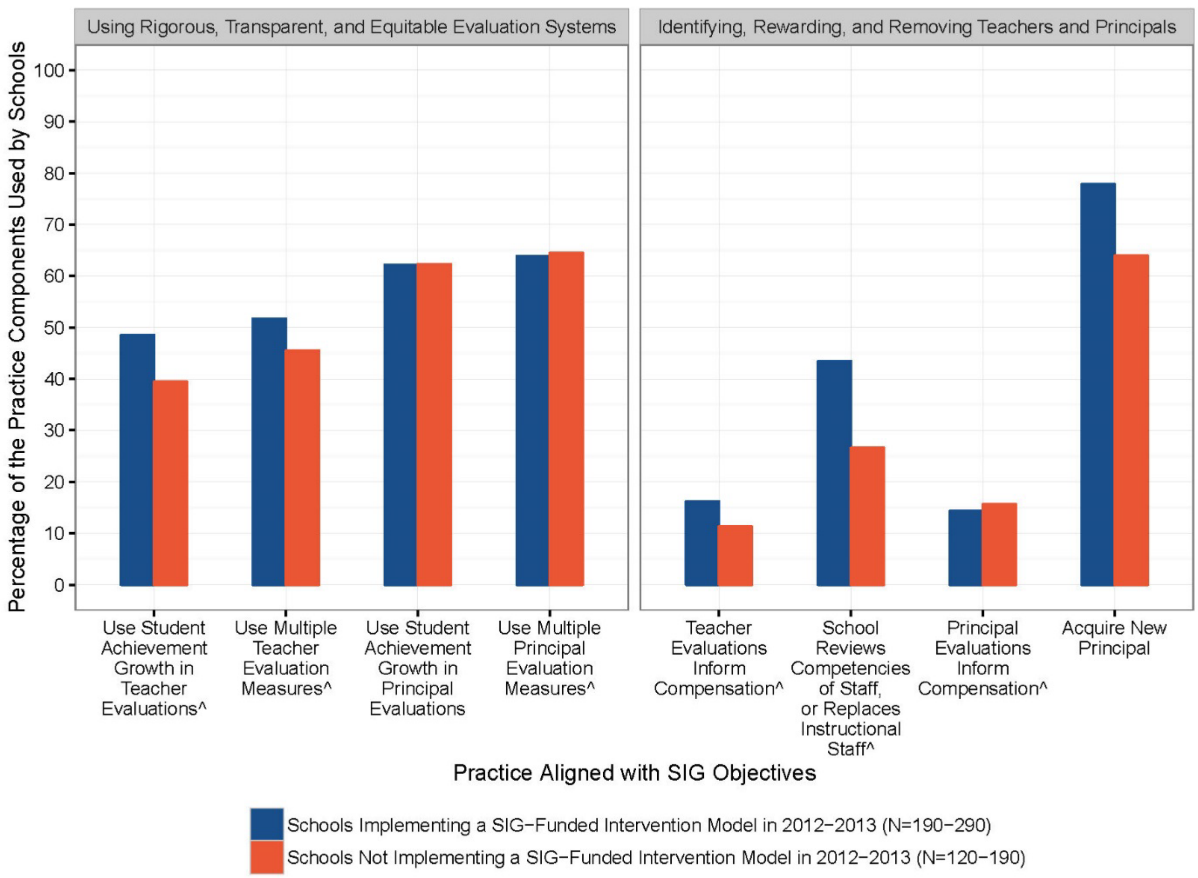
Source: Surveys of school administrators in spring 2013.

Note: Table IV.2 presents the practices summarized in this figure. Each dot in this figure represents the number of schools that reported using a particular number of practices (out of 20 examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For example, 22 schools implementing a SIG-funded intervention model reported using 10 of the 20 teacher and principal effectiveness practices aligned with the SIG application criteria. For 15 of the practices, a “yes” response received one point. In the other 5 cases, a school could receive a fraction of one point. See Appendix C for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

Most study schools reported using two practices in this area: (1) providing professional development focused on understanding and addressing student learning needs and (2) using data to evaluate professional development (Figure IV.5). In contrast, few study schools reported (1) using teacher evaluation results to inform decisions about compensation, (2) using principal evaluation results to inform decisions about compensation, or (3) using financial incentives to recruit and retain effective principals (Figures IV.4 and IV.6).

Figure IV.4. Use of individual practices aligned with SIG, using rigorous, transparent, and equitable evaluation systems subtopic and identifying and rewarding effective teachers and principals and removing ineffective ones subtopic

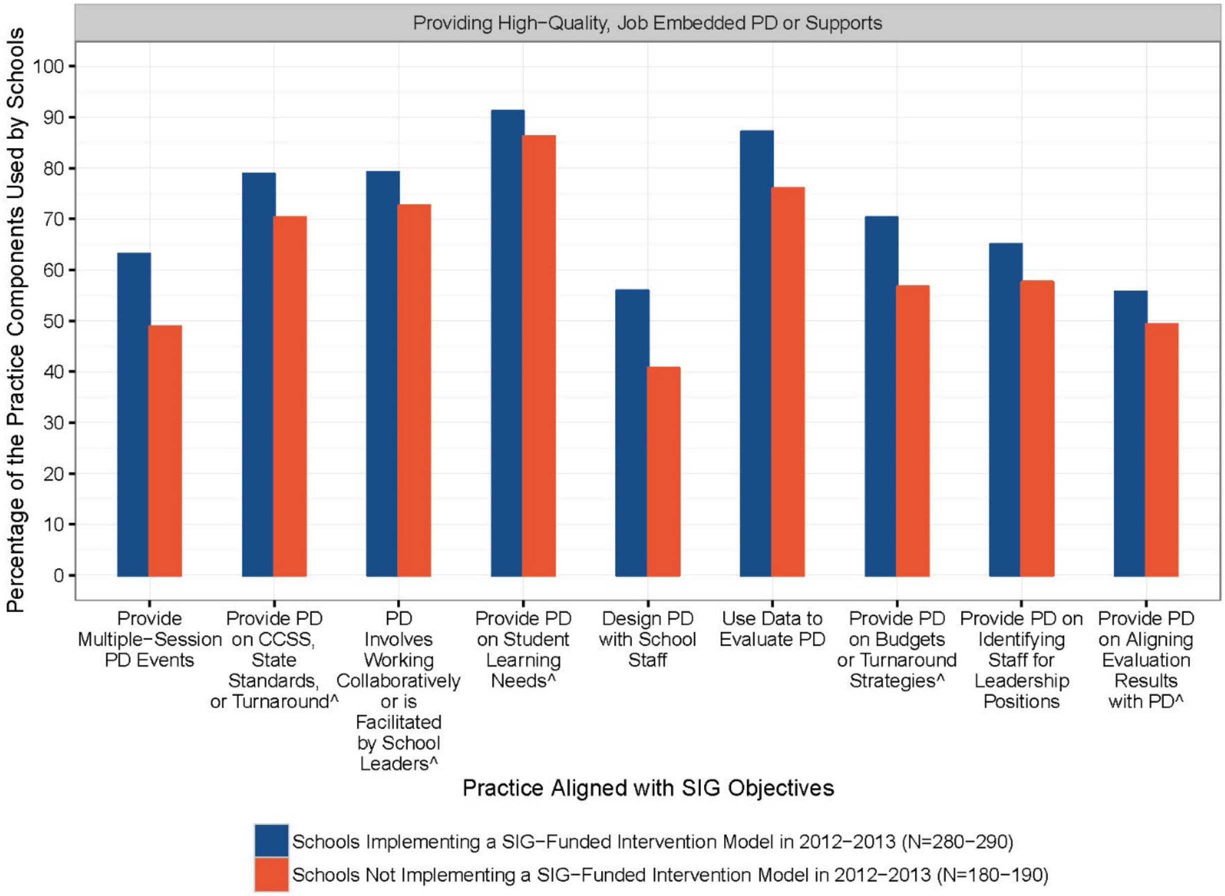


Source: Surveys of school administrators in spring 2013.

Note: This figure has a separate panel for each subtopic. As described in Appendix C, for each practice in the SIG application criteria for which we identified one or more survey questions aligned with the practice, we calculated the percentage of survey questions with a "yes" response as a measure of the percentage of components each school used. The height of each bar represents the average percentage of the components of the practice that each group of schools used. A range is provided for the sample sizes because nonresponse varied across items.

^Multiple survey questions were used to assess whether schools used all of the components of this practice.

Figure IV.5. Use of individual practices aligned with SIG, providing high-quality, job embedded professional development or supports subtopic



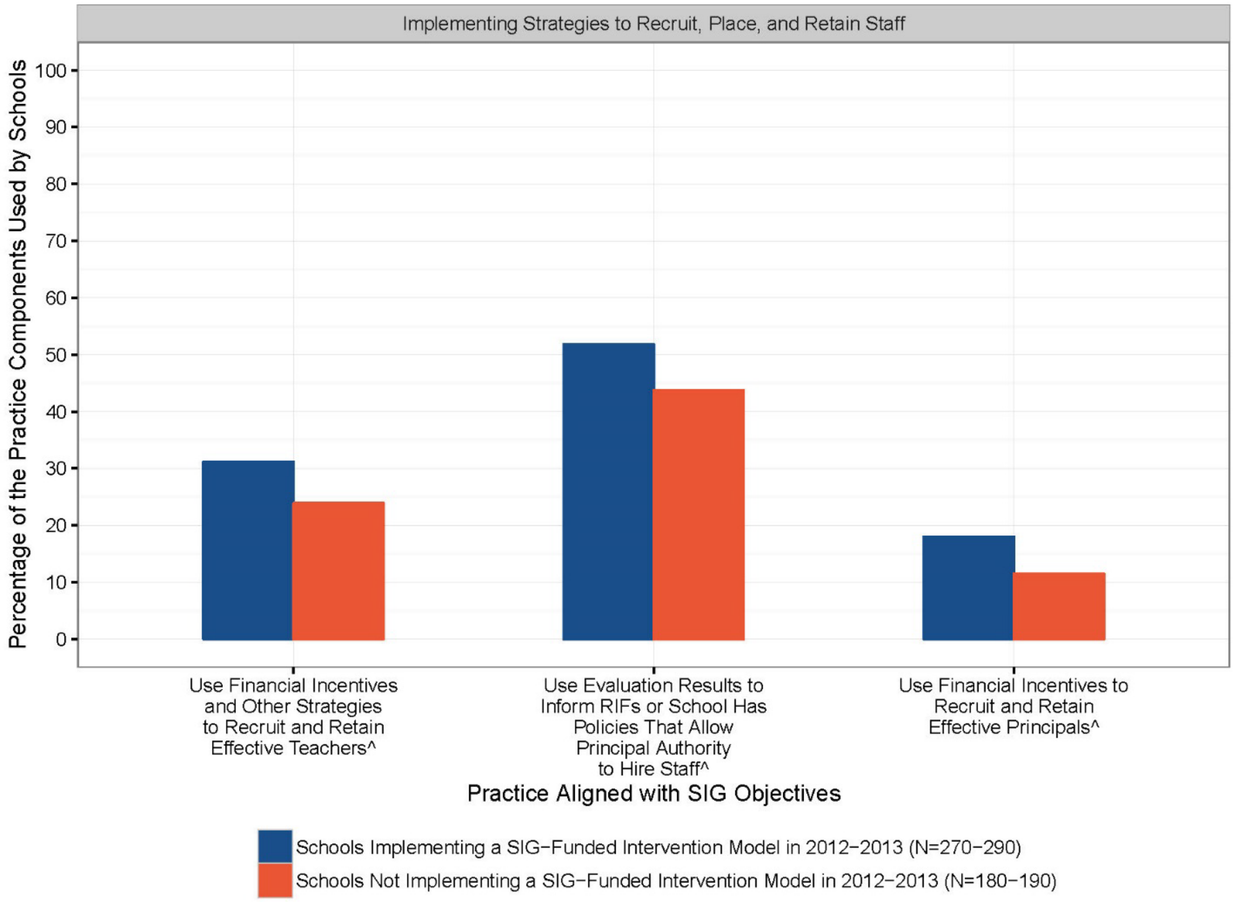
Source: Surveys of school administrators in spring 2013.

Note: As described in Appendix C, for each practice in the SIG application criteria for which we identified one or more survey questions aligned with the practice, we calculated the percentage of survey questions with a “yes” response as a measure of the percentage of components each school used. The height of each bar represents the average percentage of the components of the practice that each group of schools used. A range is provided for the sample sizes because nonresponse varied across items.

^Multiple survey questions were used to assess whether schools used all of the components of this practice.

CCSS = Common Core State Standards; PD = professional development.

Figure IV.6. Use of individual practices aligned with SIG, implementing strategies to recruit, place, and retain staff subtopic



Source: Surveys of school administrators in spring 2013.

Note: As described in Appendix C, for each practice in the SIG application criteria for which we identified one or more survey questions aligned with the practice, we calculated the percentage of survey questions with a “yes” response as a measure of the percentage of components each school used. The height of each bar represents the average percentage of the components of the practice that each group of schools used. A range is provided for the sample sizes because nonresponse varied across items.

^Multiple survey questions were used to assess whether schools used all of the components of this practice.

RIF = reductions in force.

C. Schools implementing a SIG-funded model used more learning time and community-oriented schools practices than schools not implementing such a model

To ensure that SIG schools have sufficient time for instruction and a supportive environment in which to implement policies, the application criteria for SIG focused on practices in two subtopics: (1) Increasing Learning Time and (2) Engaging Families and Communities and Providing a Safe School Environment That Meets Students' Social, Emotional, and Health Needs. The evaluation's school surveys asked about five practices aligned with SIG objectives in this topic area (Table IV.3).

Table IV.3. Practices aligned with SIG objectives on learning time and community-oriented schools, by subtopic

Increasing learning time
Using schedules and strategies that provide increased learning time or increasing the number of hours per year that school was in session
Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs
Changing policies or strategies related to parent or community engagement
State or district provided professional development on working with parents or creating a safe school environment
Changing discipline policies
Guiding the development and implementation of, or making changes to, nonacademic supports or enrichment programs for students

Source: SIG application; surveys of school administrators in spring 2012 and 2013.

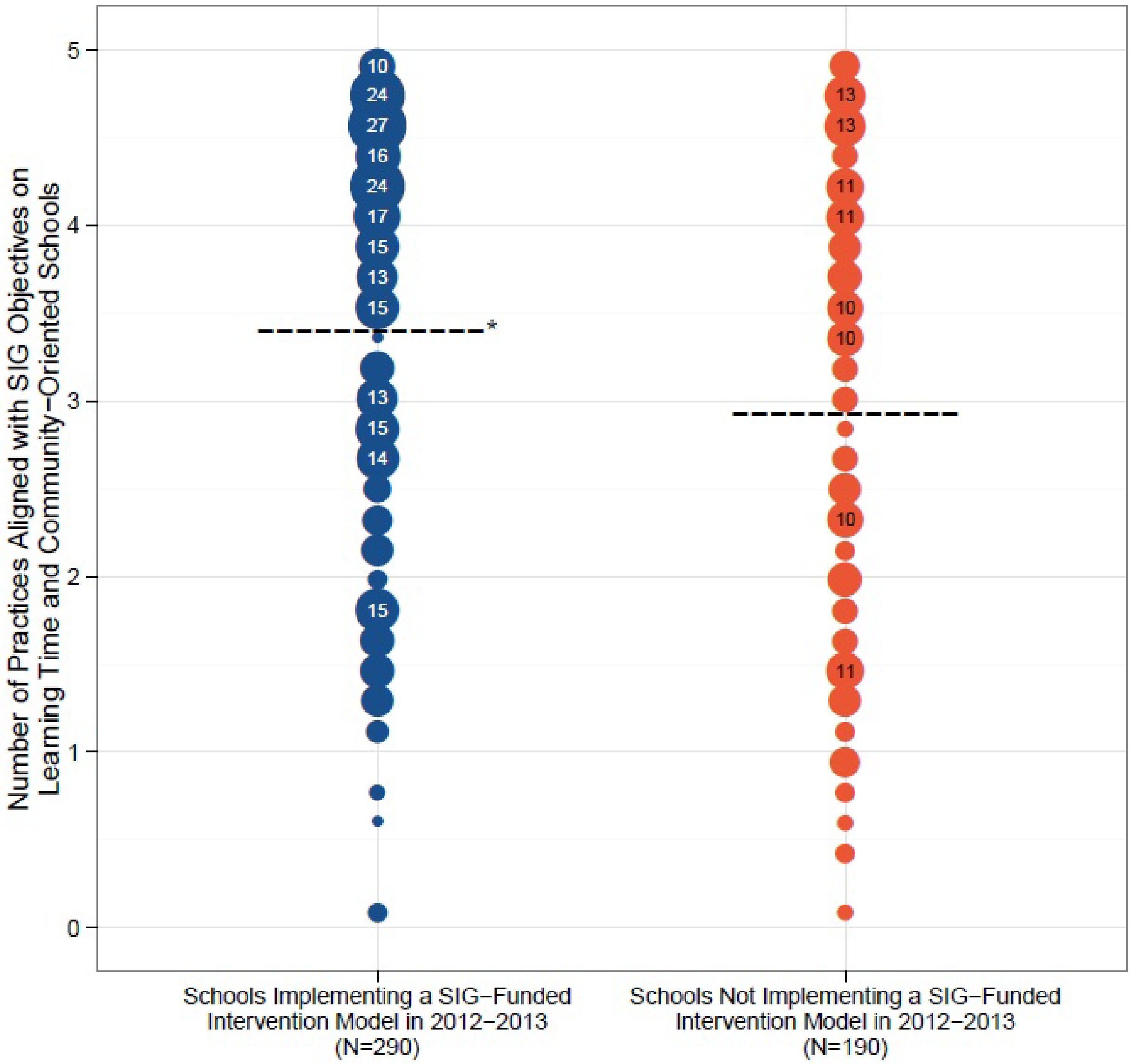
Note: See Appendix F for a list of survey questions that aligned with the SIG practices in this table.

Compared to schools that were not implementing a SIG-funded model, schools that were implementing such a model indicated in spring 2013 that they used more SIG-promoted practices in the increasing learning time and creating community-oriented schools area. Schools implementing a SIG-funded model reported using an average of 3.4 of 5 practices in this area, compared to 2.9 practices for schools not implementing one (Figure IV.7), a difference of 0.5 practices.

In spring 2013, for both subtopics in this area, schools implementing a SIG-funded model reported using more SIG-promoted practices than schools not implementing one:

- **Increasing Learning Time.** On average, schools in the first group reported using 66 percent of the components of the sole practice in this subtopic (using schedules or strategies to increase learning time), compared to 53 percent of the components for schools in the second group (Appendix C, Figure C.10).
- **Engaging Families and Communities and Providing a Safe School Environment That Meets Students' Social, Emotional, and Health Needs.** For this subtopic, schools in the first group reported using 2.7 of the 4 practices on average, compared to 2.4 for schools in the second group (Appendix C, Figure C.11).

Figure IV.7. Use of practices aligned with SIG objectives on learning time and community-oriented schools



Source: Surveys of school administrators in spring 2013.

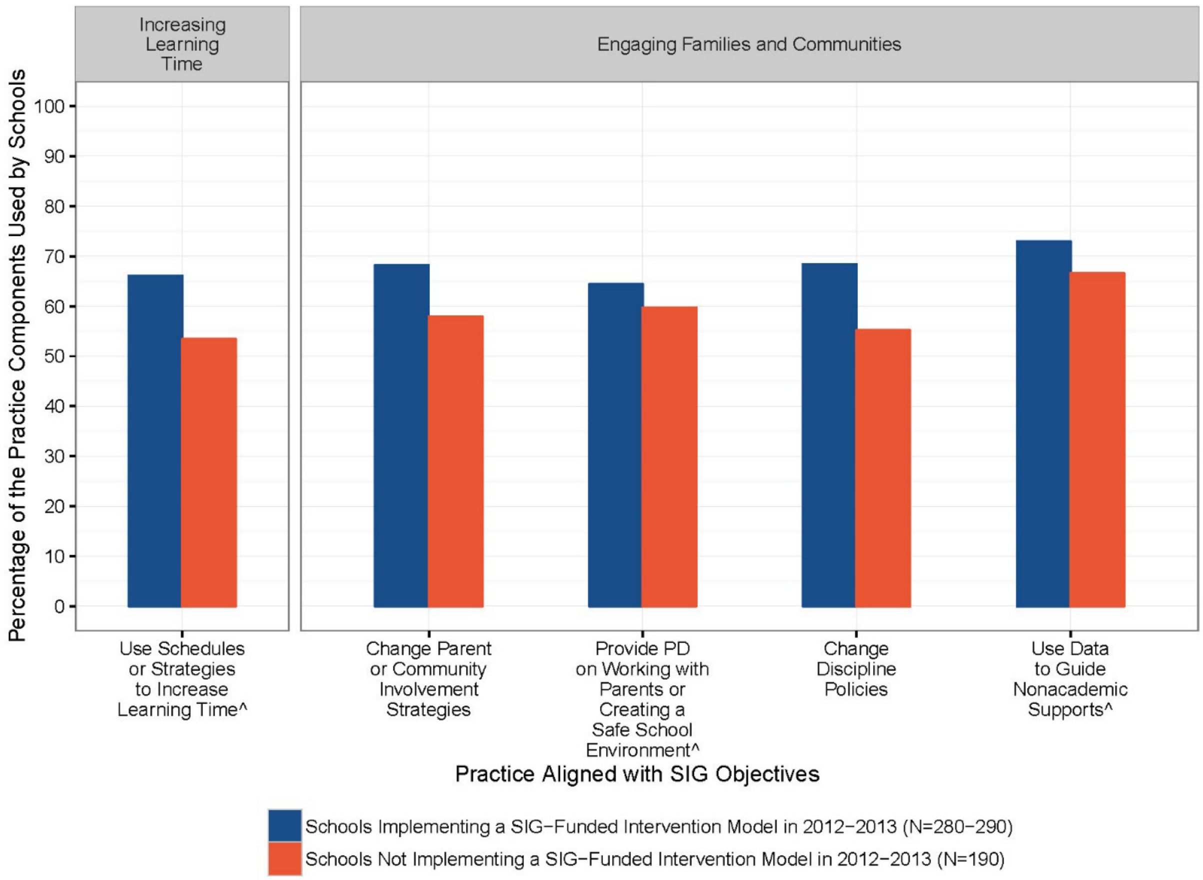
Note: Table IV.3 lists the practices summarized in this figure. Each dot in this figure represents the number of schools that reported using a particular number of practices (out of five examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For example, 13 schools implementing a SIG-funded intervention model reported using three of the five learning time and community-oriented schools practices aligned with the SIG application criteria. For four of the practices, a "yes" response received one point. In the other case, the school could receive a fraction of one point. See Appendix C for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

In this area, the practice with the largest difference between the two groups of schools was changing discipline policies. Sixty-eight percent of schools implementing a SIG-funded model and 55 percent of schools not implementing one used this practice (Figure IV.8).

In general, use of different practices in this area was fairly similar (Figure IV.8). Across all study schools, schools used 70 percent of the components of the most-used practice (using data to guide nonacademic supports), while schools used 61 percent of the components of the least-used practice (using schedules or strategies to increase learning time).

Figure IV.8. Use of individual practices aligned with SIG objectives on learning time and community-oriented schools



Source: Surveys of school administrators in spring 2013.

Note: This figure has a separate panel for each subtopic. As described in Appendix C, for each practice in the SIG application criteria for which we identified one or more survey questions aligned with the practice, we calculated the percentage of survey questions with a "yes" response as a measure of the percentage of components each school used. The height of each bar represents the average percentage of the components of the practice that each group of schools used. A range is provided for the sample sizes because nonresponse varied across items.

^Multiple survey questions were used to assess whether schools used all of the components of this practice.

PD = professional development.

D. Use of operational flexibility and support practices was higher in schools implementing SIG-funded models

To facilitate the implementation of turnaround efforts and ensure that schools receive the support needed to implement policies, the SIG application criteria focused on practices for states and districts to give schools implementing SIG models (1) operational flexibility and (2) technical assistance and support. Our school surveys asked about two practices aligned with SIG objectives in this area (Table IV.4).

Table IV.4. Practices aligned with SIG objectives on operational flexibility and support, by subtopic

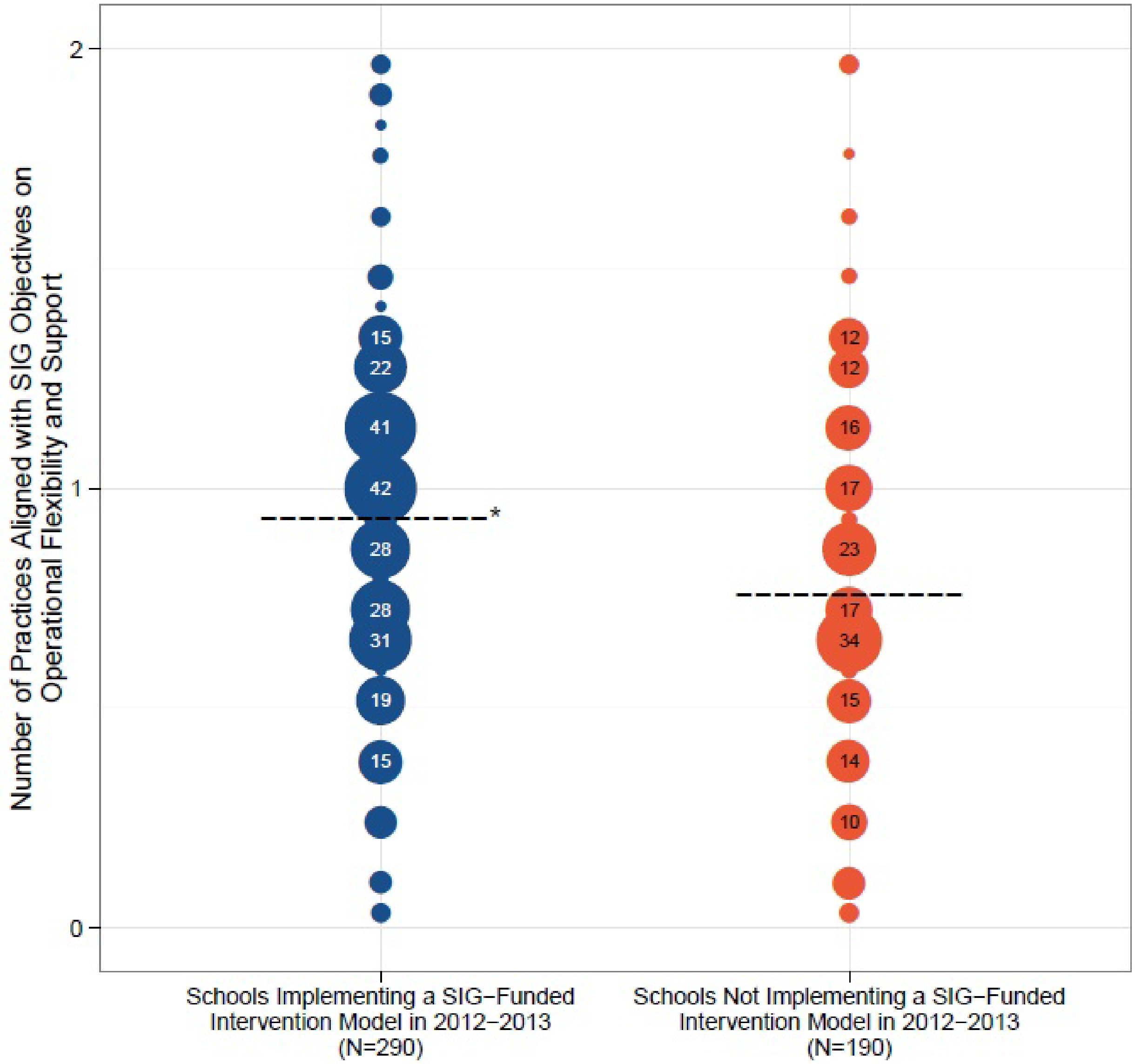
Having operational flexibility
School has primary responsibility for decisions on budget, hiring, discipline, or school year length
Receiving technical assistance and support
State, district, or an external support provider sponsored by the state or district provided training or technical assistance to support school improvement efforts, or the school received support to help administrators and teachers use data to improve instruction

Source: SIG application; surveys of school administrators in spring 2012 and spring 2013.

Note: See Appendix F for a list of survey questions that aligned with the SIG practices in this table.

Use of practices in this area was higher among schools implementing a SIG-funded model relative to schools not implementing one. In spring 2013, schools in the first group used 0.9 of 2 practices in this area, on average, compared to 0.8 for schools in the second group (Figure IV.9). Although this difference was statistically significant, it was small (just 0.1 practice), so it is unclear whether it is substantively important.

Figure IV.9. Use of practices aligned with SIG objectives on operational flexibility and support



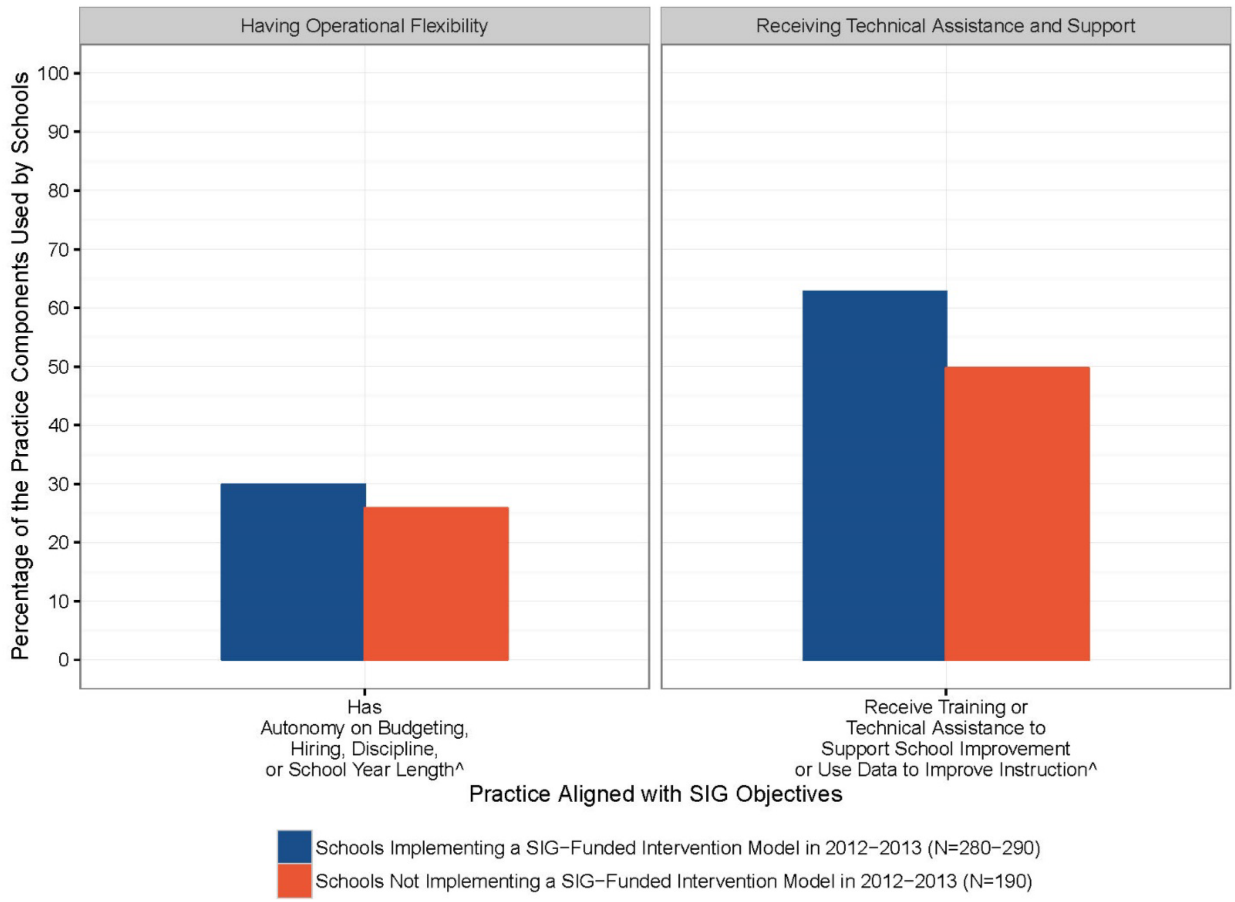
Source: Surveys of school administrators in spring 2013.

Note: Table IV.4 lists the practices summarized in this figure. Each dot in this figure represents the number of schools that reported using a particular number of practices (out of two examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For example, 42 schools implementing a SIG-funded intervention model reported using one of the two operational flexibility and support practices aligned with the SIG application criteria. For one practice, a "yes" response received one point. In the other case, it was possible for a school to receive a fraction of one point. See Appendix C for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

In spring 2013, receipt of training or support on school improvement or using data to improve instruction was higher among schools implementing a SIG-funded model than schools not implementing one. Schools in the first group used an average of 63 percent of the components of this practice, compared to 50 percent of the components for schools in the second group (Figure IV.10). This practice had a larger difference between the two groups of schools and was used to a greater extent, compared to the other practice in this area (having autonomy on budgeting, hiring, discipline, or school year length).

Figure IV.10. Use of individual practices aligned with SIG objectives on operational flexibility and support



Source: Surveys of school administrators in spring 2013.

Note: This figure has a separate panel for each subtopic. As described in Appendix C, for each practice in the SIG application criteria for which we identified one or more survey questions aligned with the practice, we calculated the percentage of survey questions with a "yes" response as a measure of the percentage of components each school used. The height of each bar represents the average percentage of the components of the practice that each group of schools used. A range is provided for the sample sizes because nonresponse varied across items.

^Multiple survey questions were used to assess whether schools used all of the components of this practice.

E. Across all four topics areas, use of SIG-promoted practices was higher among schools implementing a SIG-funded model

On average, schools implementing a SIG-funded model reported using more SIG-promoted practices than schools not implementing one in each of the four areas in spring 2013. The magnitude of these differences ranged from 0.1 to 1.6 practices per area (Figure IV.11). For all four areas, use of SIG-promoted practices varied substantially within each group of schools. Therefore, even though on average schools implementing a SIG-funded model reported using more practices than schools not implementing one, there was overlap between the number of practices used by the two groups.

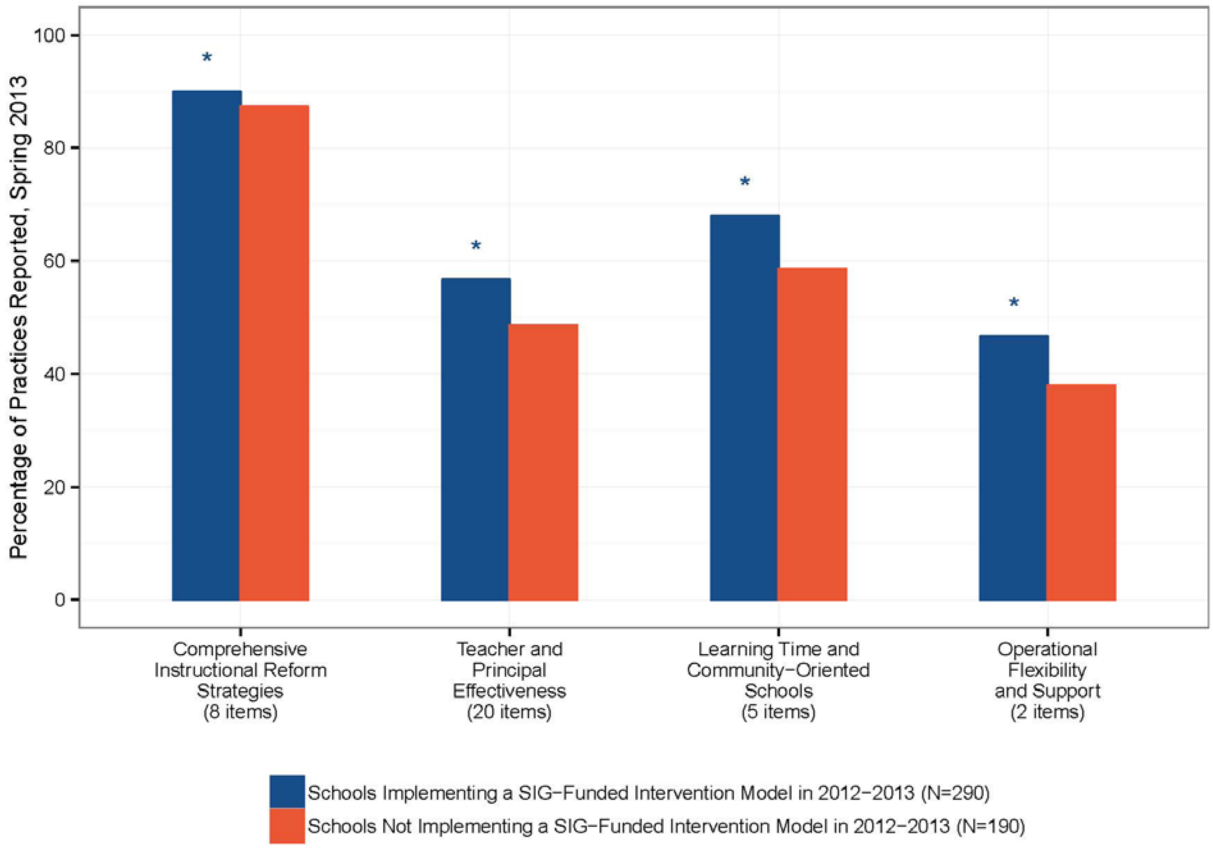
Adding the differences across the four areas, schools implementing a SIG-funded model reported using more SIG-promoted practices overall than schools not implementing one. Schools implementing a SIG-funded model reported using 22.8 out of 35 total practices (65 percent), compared to 20.3 practices (58 percent) for schools not implementing such a model, a difference of 2.5 practices. It is not clear whether a difference of this size would be meaningful in its overall influence on improvement practices and school outcomes.

Across all schools, use of SIG-promoted practices was highest in the comprehensive instructional reform strategies area and lowest in the operational flexibility and support area. Study schools reported using, on average, 89 percent of the SIG-promoted practices in the comprehensive instructional reform strategies area (7.1 of the 8 practices examined) compared to just 43 percent of the SIG-promoted practices in the operational flexibility and support area (0.87 of the 2 practices examined; Figure IV.11).

Across all practices, the practice with the largest difference between schools implementing a SIG-funded model and schools not implementing one—reviewing competencies of staff or replacing instructional staff—was in the area of teacher and principal effectiveness.

Given that use of SIG-promoted practices was highest in the comprehensive instructional reform strategies area, it may not be surprising that across all schools, the individual practice that was used the most fell in this area. Nearly all schools reported using benchmark or interim assessments at least once per year (Figure IV.2). The three practices used the least—(1) using teacher evaluation results to inform decisions about compensation, (2) using principal evaluation results to inform decisions about compensation, and (3) using financial incentives to recruit and retain effective principals—were all in the area of teacher and principal effectiveness (Figure IV.4).

Figure IV.11. Use of practices promoted by SIG, by topic area



Source: Surveys of school administrators in spring 2013.

Note: The total number of practices (shown in parentheses below each set of bars) differs by topic area. This figure reads as follows (using the first bar on the left as an example): schools implementing a SIG-funded intervention model reported using 90 percent of the practices in the comprehensive instructional reform strategies area, or 7.2 out of 8 practices examined in that area.

*Significantly different from schools not implementing a SIG-funded intervention model in 2012–2013 at the 0.05 level, two-tailed test.

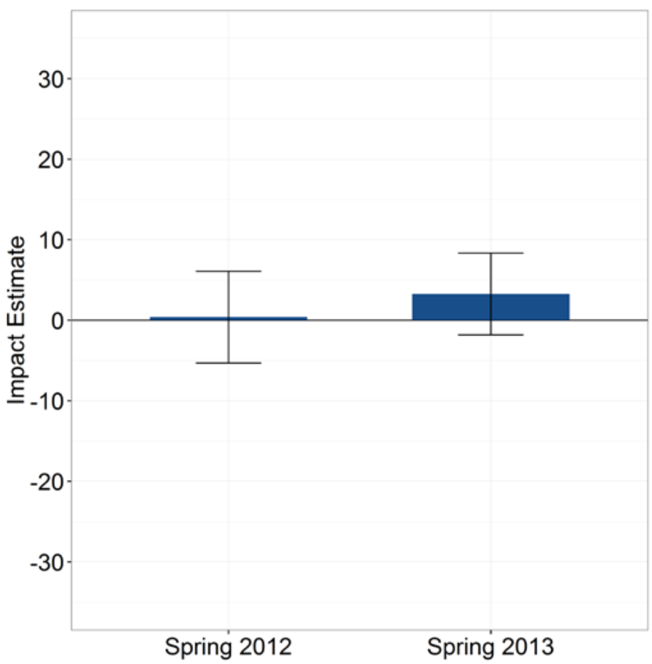
Overall, these spring 2013 findings were the same as the spring 2012 findings presented in an earlier report from this evaluation (Dragoset et al. 2015). In particular, in both spring 2012 and spring 2013, we found that:

- Schools implementing a SIG-funded model reported using statistically significantly more SIG-promoted practices than schools not implementing one in all four areas.
- There was substantial variation in use of SIG-promoted practices within each group of schools (those implementing a SIG-funded model and those not implementing one).
- Across all schools, use of SIG-promoted practices was highest in the comprehensive instructional reform strategies area and lowest in the operational flexibility and support area.

F. SIG-funded models had no significant impact on the total number of SIG-promoted practices used

When we used more rigorous RDD methods than those just presented, we found no evidence of an impact of SIG-funded models on the number of SIG-promoted practices used. In particular, we found no statistically significant impact on the number of practices used by schools close to the SIG eligibility cutoff in either spring 2012 or spring 2013 (Figure IV.12). The differences between schools that just met the eligibility criteria and those that just missed the criteria were 0.4 practices (1 percent of the 35 total practices examined) in spring 2012 and 3.3 practices (9 percent of the 35 total practices examined) in spring 2013 (Figure IV.12). SIG-funded models also had no impact on the number of SIG-promoted practices used in any of the four topic areas and little impact on the number of SIG-promoted practices used in any of the subtopic areas (see Appendix A for detailed findings). In particular, the number of significant impacts across subtopics was about what would be expected by chance: one out of 25 estimates. There were a total of 13 subtopics and two years of data (26 estimates), but we were unable to calculate one of those estimates due to insufficient sample size.

Figure IV.12. Impacts of SIG-funded models on the number of SIG-promoted practices used



Source: State and district administrative records; surveys of school administrators in spring 2012 and 2013.

Note: Units are the number of practices used, out of a total of 35 practices examined. Black lines show 95 percent confidence intervals. This figure reads as follows (using the first bar on the left as an example): in spring 2012, schools that implemented a SIG-funded model used 0.4 more practices than schools that did not implement such a model, but this difference was not statistically significant. The results shown in this figure were calculated using the regression discontinuity design methods described in Appendix A.

To help interpret these findings, it is useful to compare the findings from the descriptive analysis and RDD analysis. The less rigorous descriptive analysis presented in Sections A through E found that schools implementing a SIG-funded model in 2012–2013 used 2.5 more

SIG-promoted practices in that year than schools not implementing such a model. In contrast, the more rigorous RDD analysis (presented in Section F) found that the implementation of a SIG-funded model had no significant impact on the total number of SIG-promoted practices used by schools near the SIG eligibility cutoff in spring 2013. The difference between schools that just met the SIG eligibility criteria and those that just missed the criteria was 3.3 practices. Although this difference is similar in size to the difference we observed in the descriptive analysis, it was not statistically significant. A likely explanation for this discrepancy in statistical significance is that the RDD analysis was not able to detect the size of the difference in practices that we observed between the two groups of schools; the minimum detectable difference for the RDD analysis was 5.2 practices in spring 2013 compared to 0.8 practices for the descriptive analysis. Therefore, although our analyses show that schools implementing SIG-funded models used more SIG-promoted practices than other schools, we cannot conclude that SIG *caused* those observed differences.

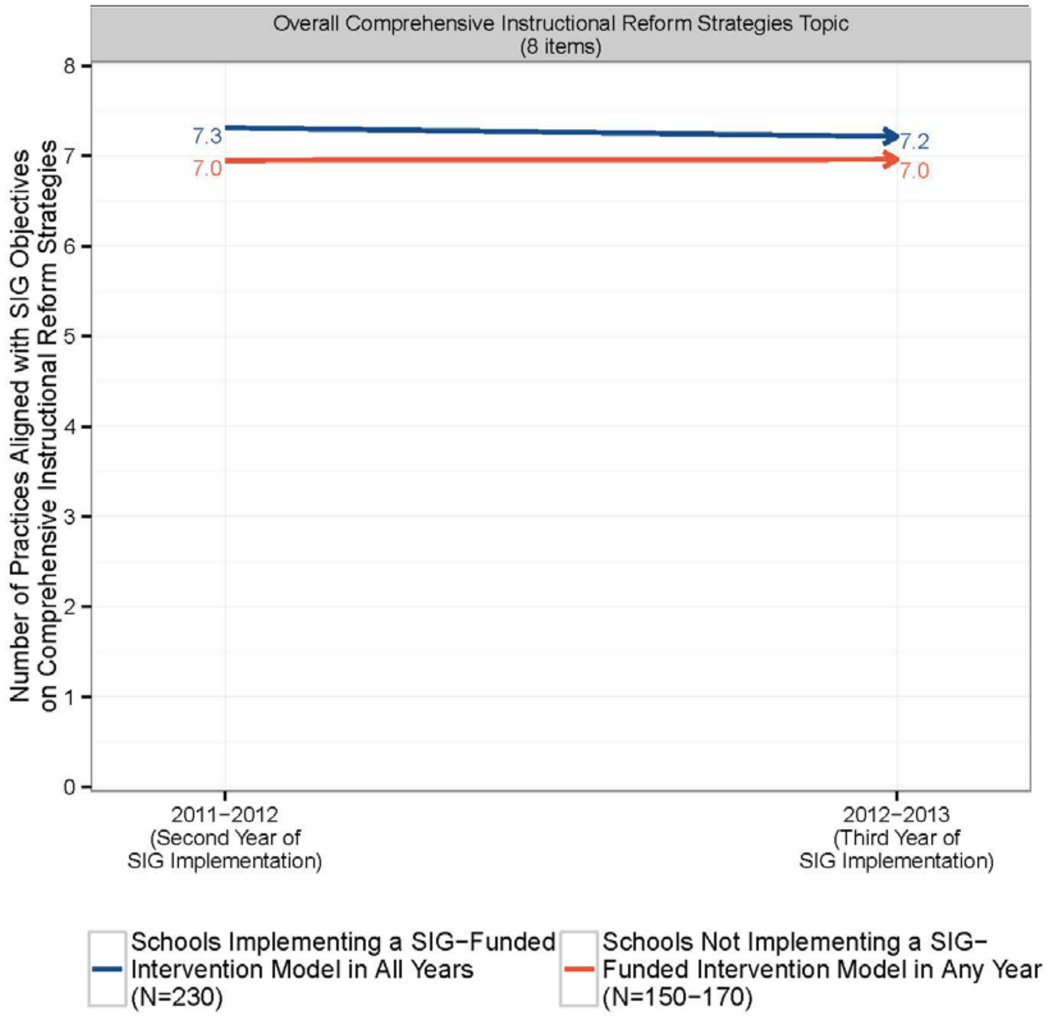
V. CHANGE OVER TIME IN SCHOOLS' USE OF PRACTICES PROMOTED BY SCHOOL IMPROVEMENT GRANTS

The process of changing education practices can be complex and require substantial time to implement. Schools might choose to prioritize certain practices over others in their early years of implementing a SIG-funded model, thereby staggering their use of SIG-promoted practices across multiple years. They may also have existing capabilities that facilitate using certain practices more quickly than others. In this chapter, we examine changes between 2011–2012 and 2012–2013 in the extent to which study schools reported using SIG-promoted practices. We focus on the same four topic areas and SIG-promoted practices that were the focus of Chapter IV.

As described in Chapter II, the analyses in this chapter focus on a different set of schools than those in Chapter IV. In this chapter, we compare schools that received cohort 1 SIG grants in 2010 and continued to implement a SIG-funded model for three years (2010–2011, 2011–2012, and 2012–2013) to schools that did not implement a SIG-funded model in any of those three years. Schools that implemented a SIG-funded intervention model in all three years would have been in their second year of implementation in spring 2012 and third year of implementation in spring 2013. In contrast, the schools examined in Chapter IV—schools that implemented a SIG-funded model in 2012–2013—may have been in their first, second, or third year of implementation in spring 2013. Schools that received cohort 1 SIG awards in 2010 would have been in their third year of implementation, schools that received cohort 2 SIG awards in 2011 would have been in their second year of implementation, and schools that received cohort 3 SIG awards in 2012 would have been in their first year of implementation in spring 2013. The comparison in this chapter is intended to provide a clearer picture of how schools' use of practices changes over time among schools at the same stage of implementation.

For three of the four areas examined, the changes over time in use of SIG-promoted practices did not significantly differ between schools implementing a SIG-funded model and schools not implementing one. In the comprehensive instructional reform area, both groups of schools used about seven out of eight practices in the two years (Figure V.1), so there may not have been much opportunity for either group to increase its use of practices from one year to the next. For the teacher and principal effectiveness area, schools implementing a SIG-funded model used 11.3 out of 20 SIG-promoted practices in this area in both years (Figure V.2). Schools not implementing a SIG-funded model used 9.1 practices in the first year and 9.5 practices in the second year. Finally, in the operational flexibility and support area, both groups of schools reported using about one practice in each of the two years (Figure V.3).

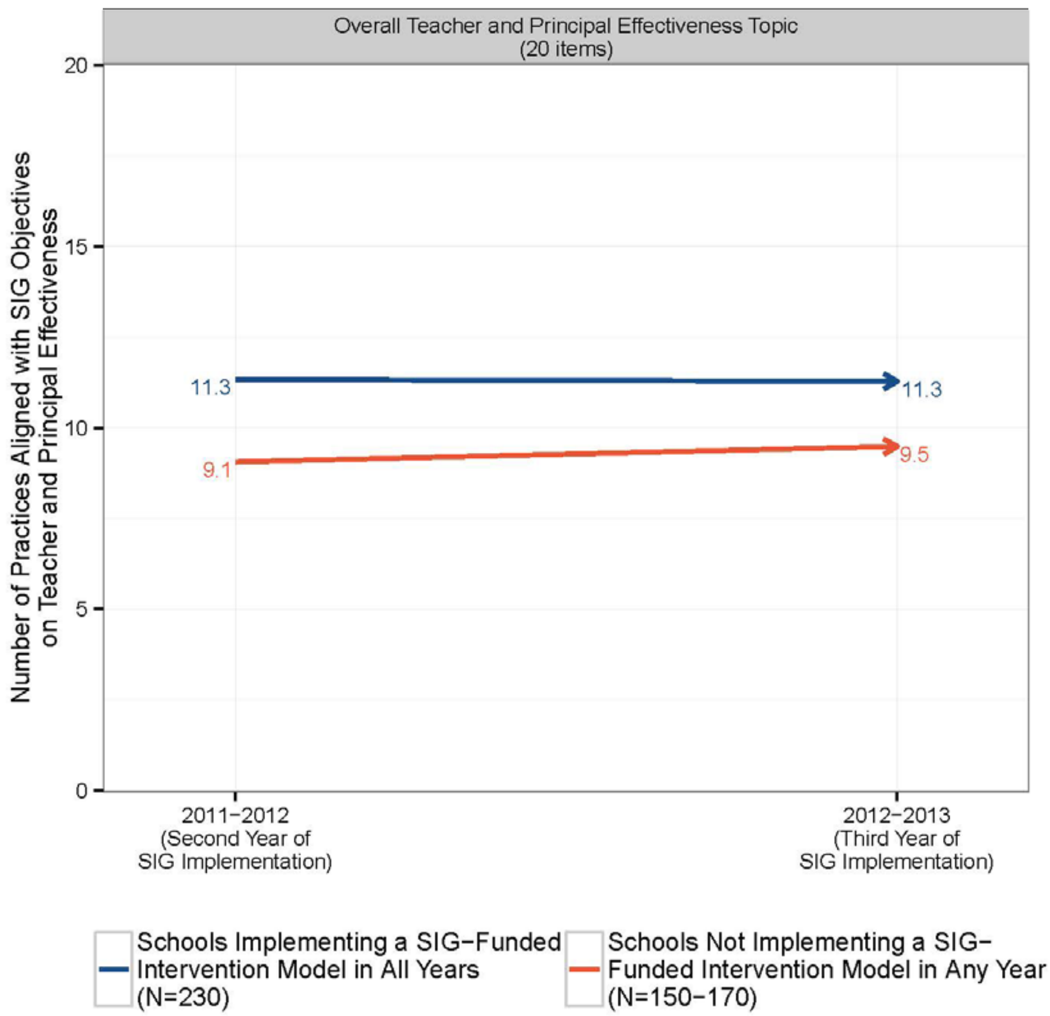
Figure V.1. Change in use of practices aligned with SIG objectives on comprehensive instructional reform strategies



Source: Surveys with school administrators in spring 2012 and spring 2013.

Note: This figure shows change over time for schools implementing a SIG-funded model and schools not implementing one in the use of SIG-promoted practices in the comprehensive instructional reform area. The arrow for each group of schools starts at the average number of reported practices used in spring 2012 and ends at the average number of reported practices used in spring 2013. For example, on average, schools implementing a SIG-funded model reported using 7.3 of the 8 practices aligned with the overall comprehensive instructional reform area in spring 2012 and 7.2 of these practices in spring 2013. There were no statistically significant differences between schools implementing a SIG-funded model and schools not implementing one with respect to changes between 2011–2012 and 2012–2013 in the number of practices used, at the 0.05 level using a two-tailed test. A range is provided for the sample sizes because nonresponse varied across years.

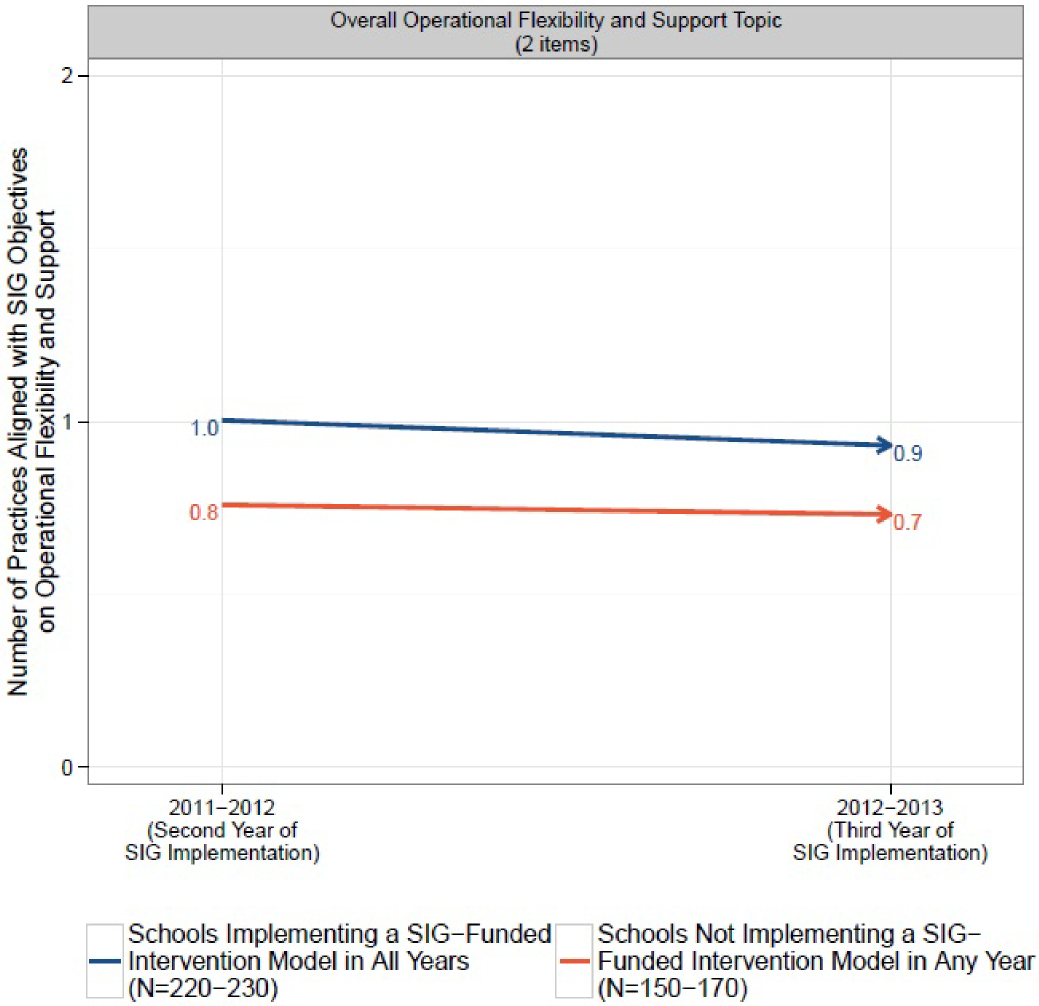
Figure V.2. Change in use of practices aligned with SIG objectives on teacher and principal effectiveness



Source: Surveys with school administrators in spring 2012 and spring 2013.

Note: This figure shows change over time for schools implementing a SIG-funded model and schools not implementing one in the use of SIG-promoted practices in the teacher and principal effectiveness area. The arrow for each group of schools starts at the average number of reported practices used in spring 2012 and ends at the average number of reported practices used in spring 2013. For example, on average, schools not implementing a SIG-funded model reported using 9.1 of the 20 practices aligned with the teacher and principal effectiveness area in spring 2012 and 9.5 of these practices in spring 2013. There were no statistically significant differences between schools implementing a SIG-funded model and schools not implementing one with respect to changes between 2011–2012 and 2012–2013 in the number of practices used, at the 0.05 level using a two-tailed test. A range is provided for the sample sizes because nonresponse varied across years.

Figure V.3. Change in use of practices aligned with SIG objectives on operational flexibility and support

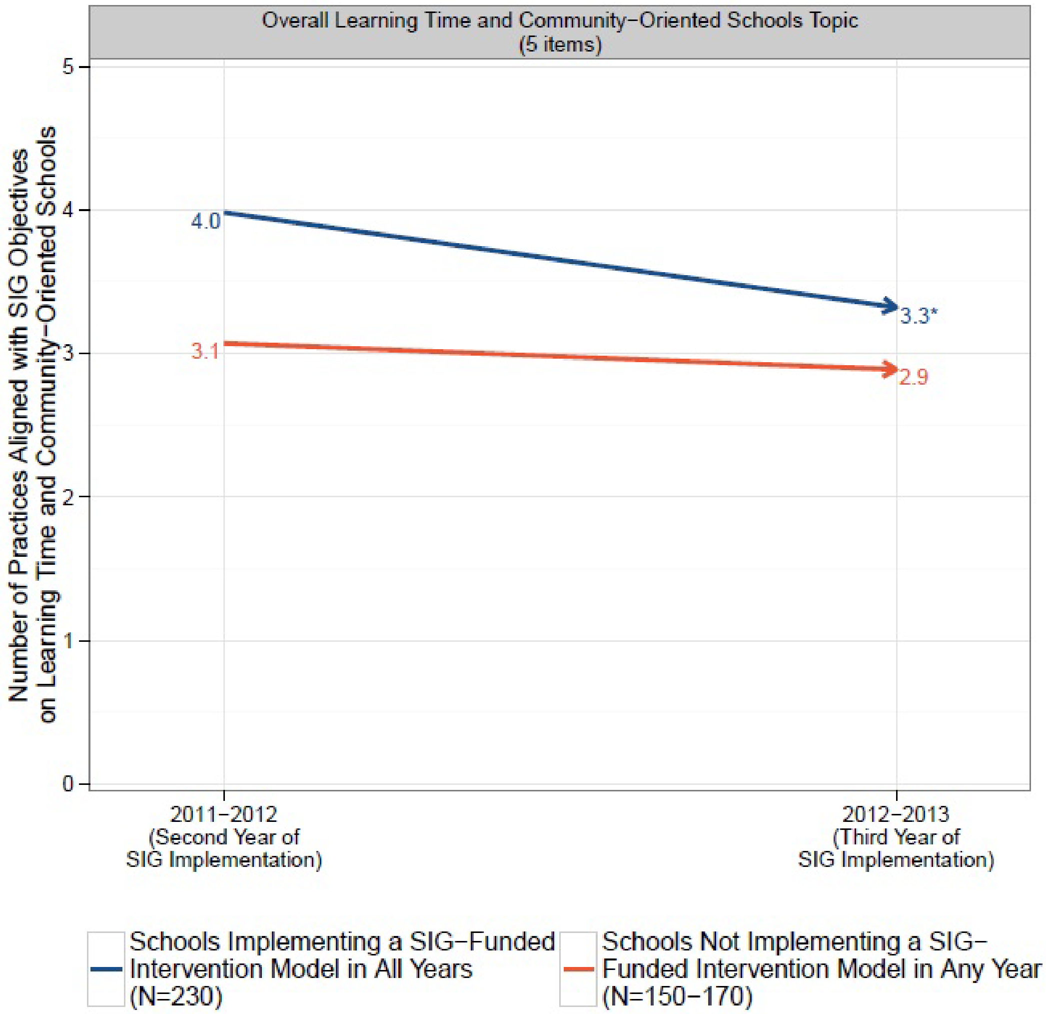


Source: Surveys with school administrators in spring 2012 and spring 2013.

Note: This figure shows change over time for schools implementing a SIG-funded model and schools not implementing one in the use of SIG-promoted practices in the operational flexibility and support area. The arrow for each group of schools starts at the average number of reported practices used in spring 2012 and ends at the average number of reported practices used in spring 2013. For example, on average, schools implementing a SIG-funded model reported using 1.0 of the 2 practices aligned with the operational flexibility and support area in spring 2012 and 0.9 of these practices in spring 2013. There were no statistically significant differences between schools implementing a SIG-funded model and schools not implementing one with respect to changes between 2011–2012 and 2012–2013 in the number of practices used, at the 0.05 level using a two-tailed test. A range is provided for the sample sizes because nonresponse varied across years.

In the remaining area (learning time and community-oriented schools), we observed decreases over time in use of SIG-promoted practices for both groups of schools and the decrease was larger for schools implementing a SIG-funded model than for schools not implementing one. The decrease between 2011–2012 and 2012–2013 was 0.7 practices for schools implementing a SIG-funded model, compared to a decrease of 0.2 for schools not implementing one (Figure V.4). However, in both years, schools implementing a SIG-funded model reported using more SIG-promoted practices than schools not implementing one; the difference between the groups was just smaller in 2012–2013. Chapter VIII provides some potential explanations for this finding.

Figure V.4. Change in use of practices aligned with SIG objectives on learning time and community-oriented schools



Source: Surveys with school administrators in spring 2012 and spring 2013.

Note: This figure shows change over time for schools implementing a SIG-funded model and schools not implementing one in the use of SIG-promoted practices in the learning time and community-oriented schools area. The arrow for each group of schools starts at the average number of reported practices used in spring 2012 and ends at the average number of reported practices used in spring 2013. For example, on average, schools implementing a SIG-funded model reported using 4.0 of the 5 practices aligned with the learning time and community-oriented schools area in spring 2012 and 3.3 of these practices in spring 2013. A range is provided for the sample sizes because nonresponse varied across years.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

Summary

This chapter examined changes over time in use of SIG-promoted practices reported by schools implementing a SIG-funded model and schools not implementing one. In three areas—comprehensive instructional reform strategies, teacher and principal effectiveness, and operational flexibility and support—there were no significant differences between the two groups of schools with respect to changes over time in practices used (Figures V.1, V.2, and V.3). In the area of learning time and community-oriented schools, however, schools implementing a SIG-funded model experienced larger decreases over time in the number of practices used than schools not implementing such a model (Figure V.4).

Thus, there was no evidence that, over time, schools implementing a SIG-funded model increased their use of SIG-promoted practices more than other schools. However, this analysis focused on use of practices in the final two years of the grant, so it is possible that schools increased their use of practices prior to that.

VI. EXAMINING WHETHER SIG-FUNDED INTERVENTION MODELS IMPROVED STUDENT OUTCOMES

In this chapter, we examine the relationship between SIG-funded models and student outcomes. We first examine whether SIG-funded models had an impact on outcomes for low-performing schools. As described in Chapter II, we used a rigorous regression discontinuity design (RDD) to conduct this analysis. This design enables us to determine whether SIG-funded models *caused* changes in outcomes for low-performing schools. The RDD impacts apply to schools closest to the RDD cutoff value—that is, schools near the lowest 5 percent of achievement in the state. Therefore, the impacts presented in this chapter do not necessarily apply to schools far from this cutoff value.

We then examine whether the type of school intervention model implemented was related to changes in outcomes for low-performing schools. As described in Chapter II, we were unable to use a rigorous method such as an RCT or RDD to conduct this analysis. Therefore, we conducted a less rigorous correlational analysis. Although this approach can provide suggestive evidence about the relative effectiveness of different models, it cannot conclusively establish which models are most effective at improving student achievement. This limitation exists because factors other than the model implemented, such as baseline differences between schools implementing different models, may explain any differences in achievement gains. Therefore, interpreting the results requires caution: the type of model implemented might not have caused any observed changes in outcomes.

In Section A, we present the impacts of SIG-funded models on student outcomes. In Section B, we present findings for subgroups of students and schools. (Appendix A provides a detailed description of the method used to conduct this analysis and presents additional findings.) Sections C and D present results from the correlational analyses conducted to examine the link between the intervention model implemented and student outcomes. Section C focuses on these findings for elementary grades (2 through 5) and Section D focuses on higher grades (6 through 12). Appendix B includes a detailed description of the methods used to conduct this analysis, as well as additional findings.

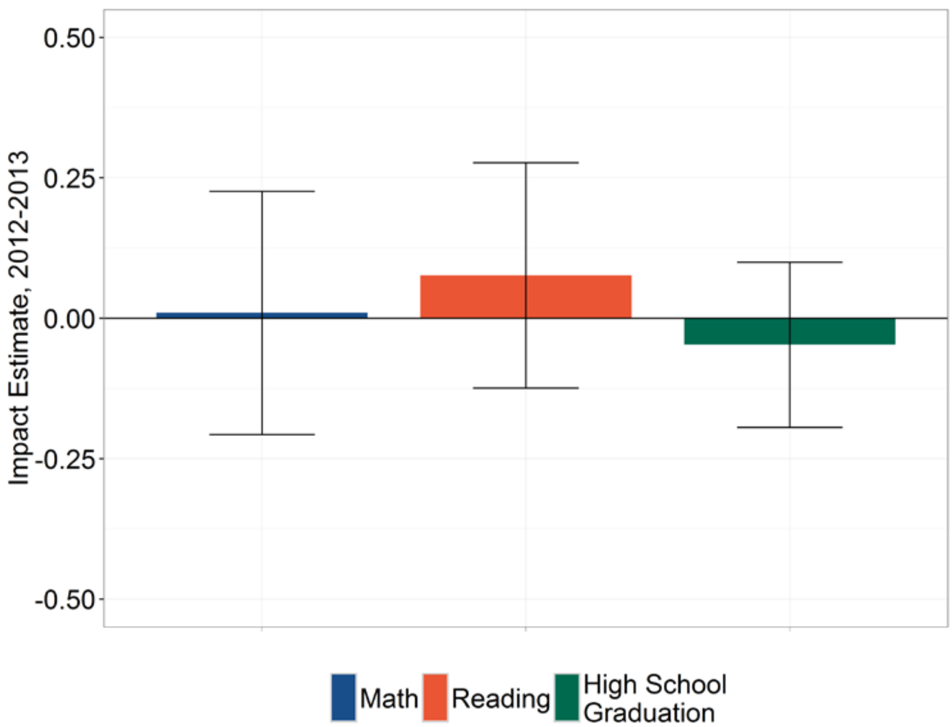
A. SIG-funded models had no statistically significant impact on test scores, high school graduation, or college enrollment

There were no significant impacts of SIG-funded models on math or reading test scores, high school graduation, or college enrollment of students in schools at the SIG eligibility cutoff (Figure VI.1 shows results for 2012–2013; Appendix A, Figure A.1 shows results for earlier years [2010–2011 and 2011–2012]). For 2012–2013, the impact on math test scores was 0.01 standard deviations, the impact on reading test scores was 0.08 standard deviations, and the impact on high school graduation was -5 percentage points, but these impacts were not statistically significant. We were unable to calculate an impact on college enrollment for 2012–2013 due to insufficient sample sizes, but we found no significant impacts on college enrollment for the other two school years (the impacts for 2010-2011 and 2011-2012 were -11 and 2 percentage points; see Appendix A for more details on this analysis).

These impact findings were robust to changes in our analytic approach. We estimated impacts using different RDD analysis methods and found no statistically significant impacts in any of these sensitivity analyses.

When we separately examined different grade levels, we found no significant impacts of SIG-funded models on math or reading test scores for any grade (Figure VI.2 shows results for 2012–2013; Appendix A, Tables A.11 and A.12 show results for the two earlier years).

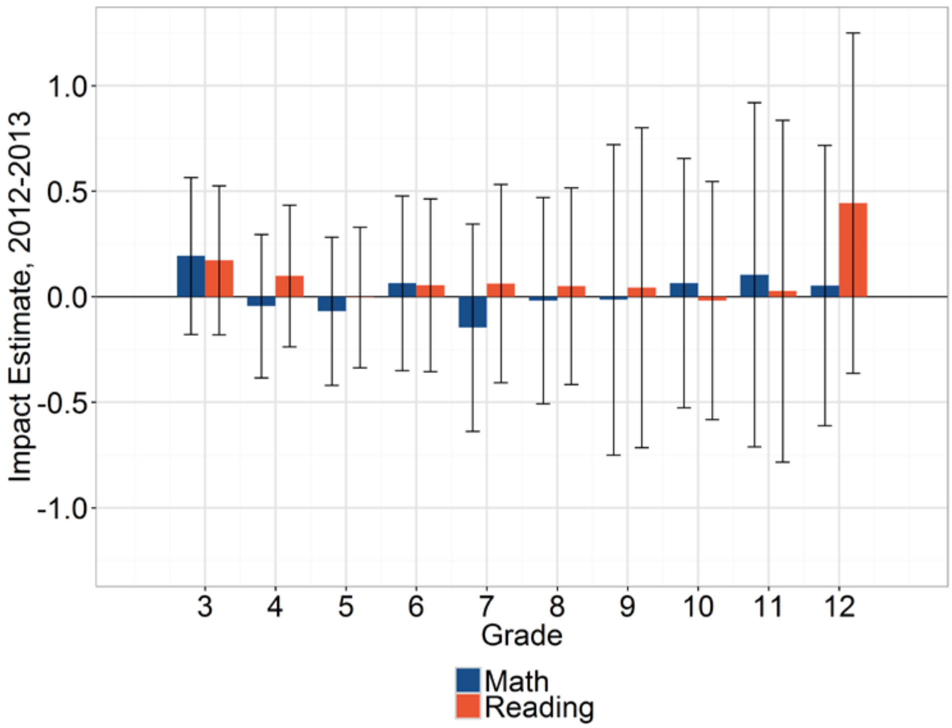
Figure VI.1. Impacts of SIG-funded models on student outcomes



Source: State and district administrative records.

Note: Units for test scores are effect sizes (test scores were standardized to have a standard deviation of 1). Units for high school graduation are percentage points/100. For example, an impact of 0.1 indicates an increase of 10 percentage points. We were unable to calculate an impact on college enrollment for 2012–2013 due to insufficient sample sizes. Black bars show 95 percent confidence intervals. The results shown in this figure were calculated using the regression discontinuity design methods described in Chapter II and Appendix A.

Figure VI.2. Impacts of SIG-funded models on student test scores, by grade



Source: State and district administrative records.

Note: Units are effect sizes (test scores were standardized to have a standard deviation of 1). Black bars show 95 percent confidence intervals. The results shown in this figure were calculated using the regression discontinuity design methods described in Chapter II and Appendix A.

SIG-funded models might have had no impact on student outcomes if SIG supplanted, rather than supplemented, other funds. Per-pupil spending would be similar in SIG and non-SIG schools if districts reallocated other funds away from SIG grantees and toward non-grantees. We investigated the hypothesis that SIG supplanted other funds using two analyses. The first analysis (using RDD methods) was uninformative because it lacked statistical power. More specifically, the smallest impacts on per-pupil spending that we were able to detect ($9,202 in 2011–2012 and $4,231 in 2012–2013) were larger than the average per-pupil grant amount ($1,600). In the second analysis, we calculated the correlation between the annual per-pupil SIG award and per-pupil spending. If SIG supplemented, rather than supplanted, other funds, this correlation should be statistically significant and positive. The correlation was 0.3 (and not statistically significant) for the 2011–2012 school year, and was 0.8 (and statistically significant) for the 2012–2013 school year. This finding is consistent with the hypothesis that SIG funds supplemented other funds in the 2012–2013 year. However, because this analysis was correlational, we cannot rule out other explanations for this observed relationship (for example, perhaps more disadvantaged schools had both larger SIG amounts and greater resources from other non-SIG sources). Therefore, we were unable to determine whether SIG funds supplemented or supplanted other existing funds.

B. There were no statistically significant impacts on student outcomes within student and school subgroups

We examined impacts separately for each of the following policy-relevant subgroups:

- English language learners (ELLs) and non-ELLs
- Elementary and secondary schools
- Title I-receiving secondary schools in improvement, corrective action, or restructuring and secondary schools that were eligible for, but did not receive, Title I funds[23]
- Schools in early RTT, later RTT, and non-RTT states[24]

We found no significant impacts of SIG-funded models for any of these subgroups.

Although SIG-funded models had no impacts on student outcomes for schools near the eligibility cutoff on average, impacts could have differed across the four types of intervention models that schools implemented. We were not able to examine these differences in impacts using the RDD, so we instead used a less rigorous correlational analysis. Our correlational analysis examined the associations between each intervention model and student achievement, separately for elementary and middle/high school grades.

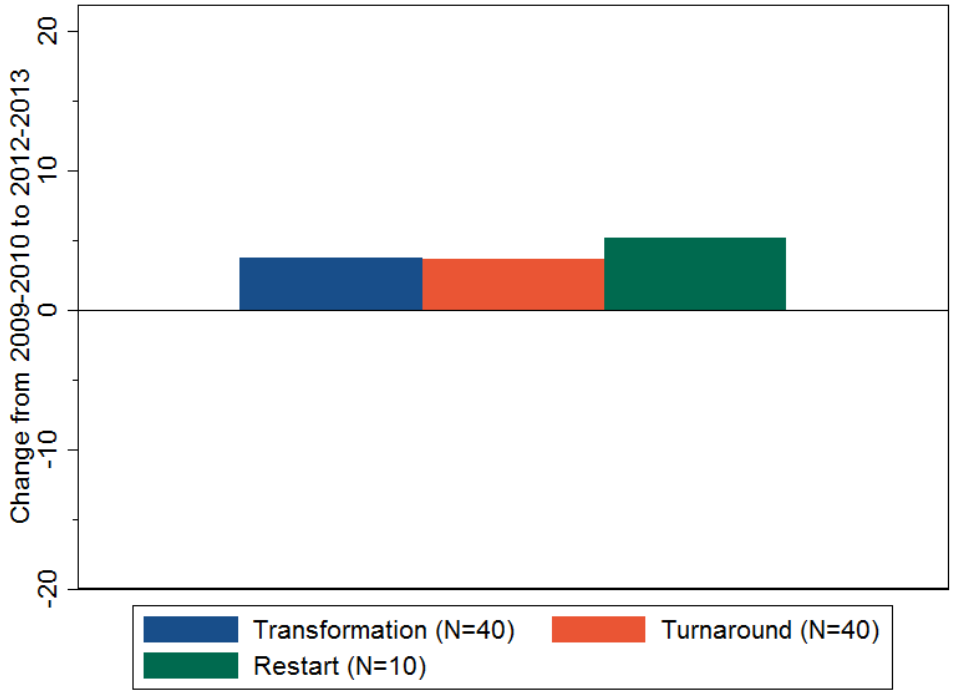
C. In elementary grades, there was no evidence that one model was more effective at improving student achievement than another

For elementary grades (2nd through 5th), we found no evidence that one intervention model was more effective than another at improving student achievement. Between 2009–2010 (the year prior to SIG implementation) and 2012–2013, there were no significant differences in math or reading gains between schools implementing different models (Figure VI.3 presents math results; see Appendix B, Figure B.2 for reading results). This finding was also true for the two other outcome years we examined (2010–2011 and 2011–2012) and across all sensitivity analyses (Appendix B, Figures B.1, B.2, B.16, and B.17).

[23] These subgroups were of interest because each one represented a separate RDD opportunity with distinct intervention and comparison groups under the original tier definitions. The first group—Title I-receiving secondary schools in improvement, corrective action, or restructuring—were divided into Tier I (the intervention group) and Tier III (the comparison group). The second subgroup—secondary schools that were eligible for, but did not receive, Title I funds—were divided into Tier II (the intervention group) and SIG-ineligible schools (the comparison group).

[24] The RDD analysis sample included 7 of the 12 early RTT states, 4 of the 7 later RTT states, and 10 of the 32 non-RTT states, so the analysis of RTT versus other states was not fully representative of these groups of states.

Figure VI.3. Changes in math test scores in elementary grades, by model



Source: State administrative data.

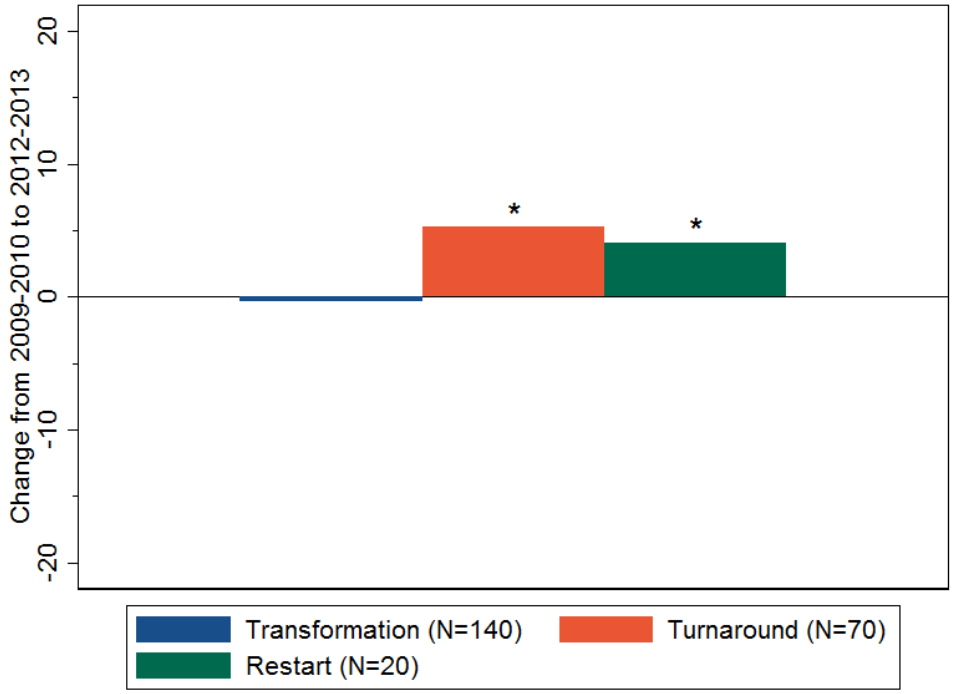
Notes: This figure depicts regression-adjusted changes in math test scores between the baseline year (2009–2010) and 2012–2013 in grades 2 through 5. Changes in math test scores were regression-adjusted for state and grade using a linear model. Units are normal curve equivalents (NCEs). The NCEs reported in this figure correspond to the following effect sizes (ESs): transformation ES = 0.18, turnaround ES = 0.17, restart ES = 0.25. There were no statistically significant differences between schools implementing different models.

D. In higher grades, we found larger student achievement gains in math for the turnaround model than the transformation model, but factors other than the model implemented may explain these differences

For higher grades (6th through 12th), between 2009–2010 and 2012–2013, schools implementing the turnaround model experienced larger gains in math than schools implementing the transformation model (Figure VI.4 shows math results; see Appendix B, Figure B.4 for reading results). These gains do not appear to be fully explained by changes in the student body composition of turnaround schools.

The difference in gains that we observed for turnaround versus transformation schools was substantively important—it was equal to approximately one year of typical growth in math for middle and early high school students. The gain for turnaround schools was equivalent to approximately 0.25 standard deviations, compared to a gain of essentially zero for transformation schools (the small loss depicted in Figure VI.4 was not significantly different from zero). The difference between the two translated into an effect size of 0.25, which is in line with a typical year of growth in math for middle and early high school students that Hill et al. (2008) estimated to be 0.30 for 6th graders, 0.32 for 7th graders, 0.22 for 8th graders, and 0.25 for 9th graders.

Figure VI.4. Changes in math test scores in higher grades, by model



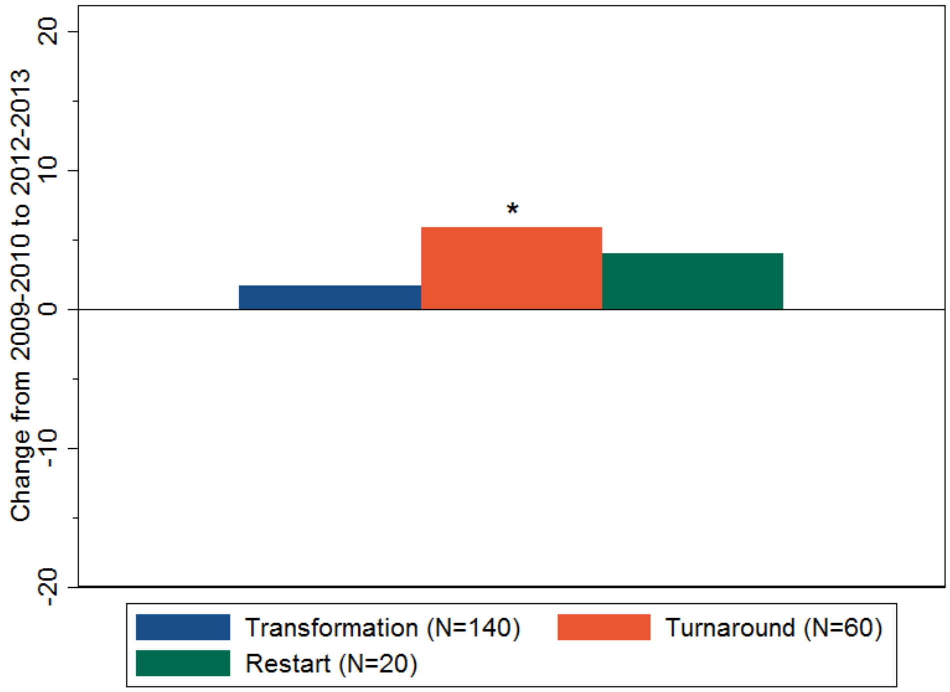
Source: State administrative data.

Notes: This figure depicts regression-adjusted changes in math test scores between the baseline year (2009–2010) and 2012–2013 in grades 6 through 12. Changes in math test scores were regression-adjusted for state and grade using a linear model. Units are normal curve equivalents (NCEs). The NCEs reported in this figure correspond to the following effect sizes (ESs): transformation ES = -0.02, turnaround ES = 0.25, restart ES = 0.19.

*Significantly different from transformation model.

While restart schools also experienced larger gains in math and reading than transformation schools, the gains for restart schools appear to be largely explained by changes in the composition of students attending restart schools. In the sensitivity analyses that accounted for student mobility, we calculated outcome gains using test scores of students who were *slated to attend* a particular school, as opposed to students who *actually attended* the school (which was the approach used in the benchmark analysis). We identified the school a student was slated to attend based on the school they attended in the baseline year and typical school feeder patterns in the district. In these analyses, the difference between the restart and transformation models was no longer significant (Figure VI.5 shows math results; see Appendix B for more details on this sensitivity analysis and Appendix B, Figure B.15 for reading results).

Figure VI.5. Changes in math test scores in higher grades, accounting for student mobility, by model



Source: State administrative data.

Notes: This figure depicts regression-adjusted changes in math test scores between the baseline year (2009–2010) and 2012–2013 in grades 6 through 12, using changes calculated in a way that accounted for student mobility. Changes in math test scores were regression-adjusted for state and grade using a linear model. Units are normal curve equivalents (NCEs). The NCEs reported in this figure correspond to the following effect sizes (ESs): transformation ES = 0.08, turnaround ES = 0.28, restart ES = 0.19.

*Significantly different from transformation model.

Consistent with these findings, we found that the composition of students attending restart schools serving higher grades changed over time, relative to transformation and turnaround schools:

1. Between 2009–2010 and 2012–2013, restart schools serving students in higher grades lost more economically disadvantaged students (that is, students eligible to receive free or reduced-price lunch) than transformation and turnaround schools (Figure VI.6).
2. Students attending restart schools in 2012–2013 were higher achieving (as measured by pre-SIG math test scores from 2009–2010) than students attending restart schools in 2009–2010 (Figure VI.7) and this change over time differed from the change over time for transformation and turnaround schools for both math (Figure VI.7) and reading (Appendix B, Figure B.13).

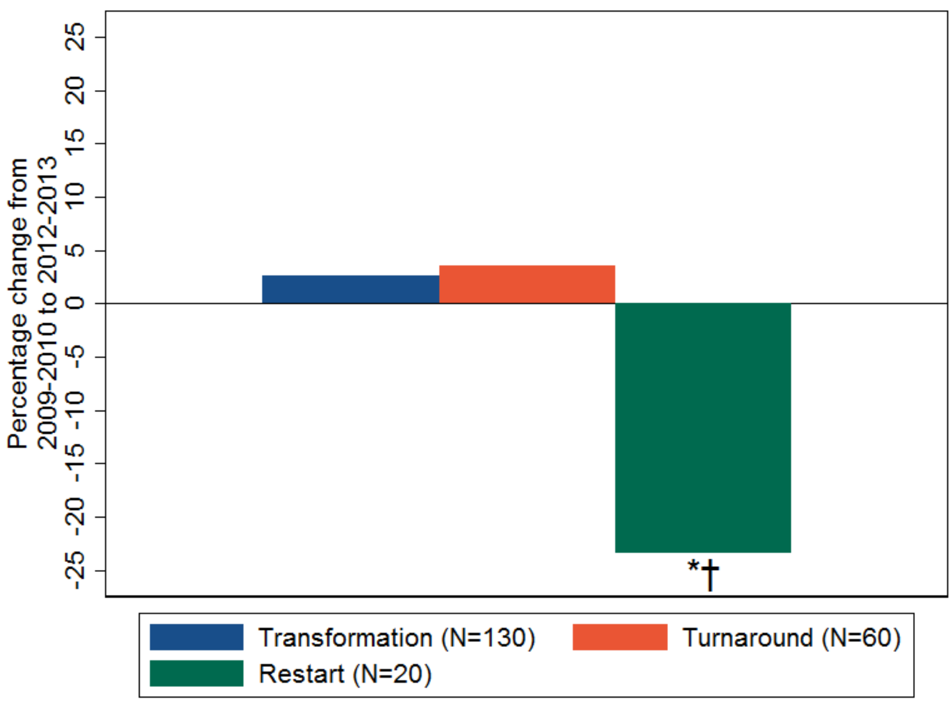
These compositional changes could account for the statistically significantly larger gains in test scores that we observed for restart versus transformation schools in our benchmark analysis. In the sensitivity analysis that accounts for student mobility, we found that the test score gains for restart schools were no longer significantly larger than the gains for transformation schools.

Relative to the benchmark analysis, accounting for student mobility increased gains for transformation schools, reducing the difference between their gains and those of restart schools. For example, in math, accounting for student mobility increased the gains for transformation schools from -0.02 to 0.08 standard deviations but did not change the gains for restart schools, which stayed at 0.19 standard deviations in both the benchmark and sensitivity analyses. This pattern of findings suggests that the larger gains for the restart model appear to be due to changes in student composition rather than the restart model being more effective than the transformation model.

The sensitivity analyses that account for student mobility did not change the finding of larger student achievement gains in schools implementing the turnaround model relative to schools implementing the transformation model. The difference in math gains between these two types of schools remained significant in the sensitivity analyses.

This pattern was consistent with findings related to changes over time in the composition of students attending turnaround and transformation model schools. In particular, we found that the composition of students attending turnaround schools serving higher grades did not change over time, relative to transformation schools serving those grades, which suggests that the larger test score gains we observed for turnaround schools were not entirely due to student mobility.

Figure VI.6. Percentage change in students eligible for free and reduced-price lunch in schools serving students in higher grades, by model



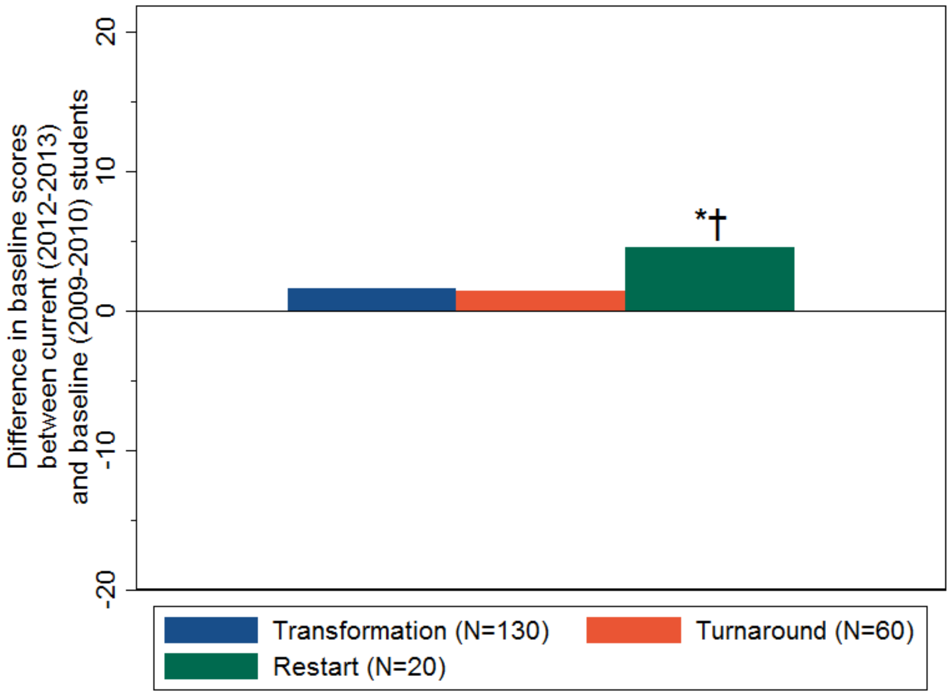
Source: State administrative data.

Notes: This figure depicts regression-adjusted changes in the percentage of students eligible for free or reduced-price lunch relative to the baseline year (2009–2010) in schools serving grade 6 through 12. Percentage changes were regression-adjusted for state using a linear model.

*Significantly different from transformation model.

†Significantly different from turnaround model.

Figure VI.7. Difference in baseline math scores between current and baseline students in schools serving students in higher grades, by model



Source: State administrative data.

Notes: This figure depicts regression-adjusted average differences in baseline math scores between students attending schools serving grade 6 through 12 in 2012–2013 and students attending those same schools in the baseline year (2009–2010). For each study school serving grade 6 through 12, we calculated the average 2009–2010 math test score for students attending that school in 2009–2010 (baseline students). We also calculated the average 2009–2010 math test score for students attending that school in 2012–2013 (current students). To compute the difference, we subtracted the average 2009–2010 math test score for baseline students from the average 2009–2010 math test score for current students. Positive differences indicate that the students attending a school in 2012–2013 had higher scores at baseline than the students attending that school in the baseline year (2009–2010). Differences in baseline test scores were regression-adjusted for state using a linear model. Units are normal curve equivalents (NCEs). The NCEs reported in this figure correspond to the following z-scores (in other words, differences in standard deviation units): transformation z-score = 0.08, turnaround z-score = 0.07, restart z-score = 0.22.

*Significantly different from transformation model.

†Significantly different from turnaround model.

However, factors other than the SIG model implemented, such as baseline differences between schools implementing different models, may explain the differences in math achievement gains between transformation and turnaround schools. Schools implementing different models served different populations of students at baseline. Among schools serving students in higher grades, we found the following:

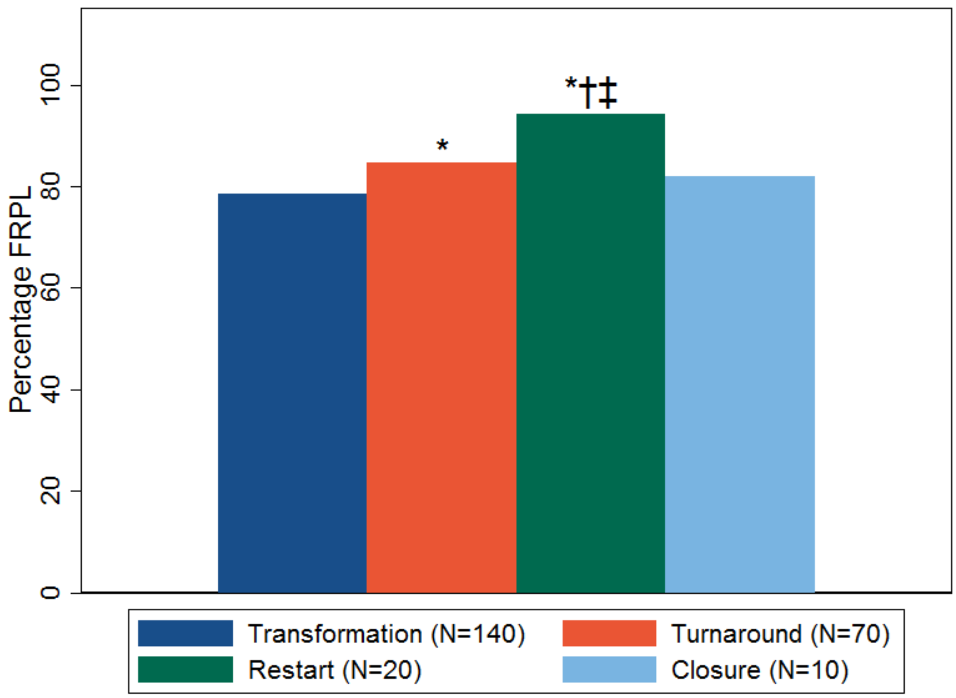
- Turnaround schools served more economically disadvantaged and lower-achieving students at baseline than transformation schools (Figures VI.8 and VI.9).

- Restart schools served more economically disadvantaged and lower-achieving students at baseline than transformation and turnaround schools.

The pattern for reading test scores was similar to the pattern for math (see Appendix B, Figure B.23).

Based on these findings, we cannot rule out the possibility that differences between schools implementing different models might account for the significant differences in math gains that we observed between transformation and turnaround schools for higher grades. In sensitivity analyses that included baseline school characteristics as control variables, the math gains for turnaround schools remained significantly larger than the gains for transformation schools (Appendix B, Figure B.18). However, significant differences between models in terms of measurable characteristics (such as the percentage of students eligible for free or reduced-price lunch and baseline test scores) could indicate that schools implementing different models also differed in unmeasurable ways that our analyses could not capture.

Figure VI.8. Baseline percentages of students eligible for free and reduced-price lunch in schools serving students in higher grades, by model



Source: State administrative data.

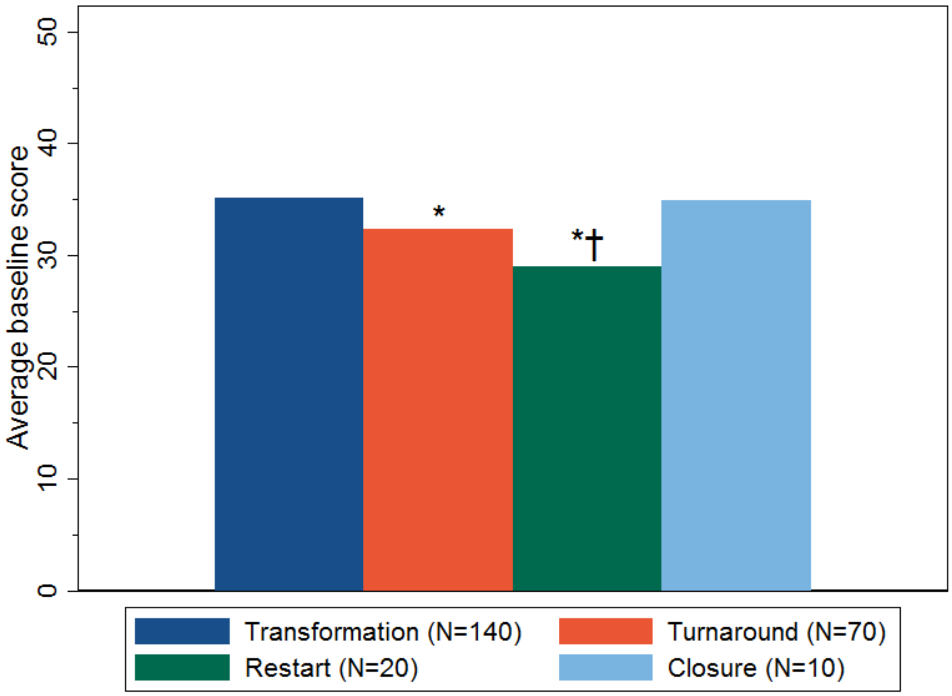
Notes: This figure depicts regression-adjusted percentages of students eligible for free or reduced-price lunch at baseline (2009–2010) in schools serving grade 6 through 12. Percentages were regression-adjusted for state using a linear model.

*Significantly different from transformation model.

†Significantly different from turnaround model.

‡Significantly different from closure model.

Figure VI.9. Average baseline math test scores in schools serving students in higher grades, by model



Source: State administrative data.

Notes: This figure depicts regression-adjusted average math scores at baseline (2009–2010) in schools serving grade 6 through 12. Math scores were regression-adjusted for state using a linear model. Units are normal curve equivalents (NCEs). An NCE of 50 represents the statewide mean score. The NCEs reported in this figure correspond to the following z-scores (in other words, differences in standard deviation units): transformation z-score = -0.70, turnaround z-score= -0.84, restart z-score = -1.0, closure z-score = -0.72.

*Significantly different from transformation model.

†Significantly different from turnaround model.

E. Summary

We found that the implementation of SIG-funded models had no significant impact on math or reading test scores, high school graduation, or college enrollment for schools near the SIG eligibility cutoff. In addition, there were no significant impacts within student and school subgroups.

For elementary grades, we found no evidence suggesting that one model was more effective at improving student achievement than another. In higher grades, the turnaround and restart models were associated with larger student achievement gains than the transformation model. However, factors other than the SIG model implemented, such as changes in student composition and baseline differences between schools implementing different models, may explain these differences in gains.

VII. EXTENT TO WHICH SCHOOLS FOCUS ON ENGLISH LANGUAGE LEARNERS IN THEIR USE OF PRACTICES PROMOTED BY SCHOOL IMPROVEMENT GRANTS

English language learners (ELLs) are of particular interest to this evaluation for two reasons. First, they have historically had lower academic achievement than other students. Since 2002, ELLs have scored lower than other students on the National Assessment of Education Progress reading exam.[25] Second, the SIG program emphasized prioritizing the academic achievement of high-needs students, including ELLs, as a way to address this achievement gap. In particular, the SIG application criteria called on districts and schools to provide supports and professional development to teachers and principals to ensure that ELLs acquire language skills to master academic content (U.S. Department of Education 2010a). For these reasons, ED's Office of English Language Acquisition requested that part of this evaluation focus on how schools used the practices promoted by SIG to address the needs of ELLs.

In this chapter, we assess the extent to which schools implementing a SIG-funded model and those not implementing one reported focusing on ELLs in their use of practices promoted by SIG. We present results from four types of analyses:

1. We compared use of ELL-focused practices aligned with the SIG application criteria for schools implementing a SIG-funded model and schools not implementing one.

It is possible that schools with higher ELL populations and higher achievement gaps between ELLs and other students prioritized ELL-focused education reforms more than other schools. Because of this, the first analysis might obscure important differences between schools implementing a SIG-funded model and schools not implementing one that exist only in schools with higher ELL populations or ELL achievement gaps. For this reason, we conducted the following analyses:

2. We compared use of ELL-focused practices for schools implementing a SIG-funded model and schools not implementing one within each of the following four groups:
   a. Schools with higher ELL populations (defined as schools with percentages of ELLs above the median for our study sample). For example, within the group of schools with higher ELL populations, we compared schools implementing a SIG-funded model to schools not implementing one.
   b. Schools with lower ELL populations (defined as schools with percentages of ELLs below the median).
   c. Schools with higher ELL achievement gaps (defined as schools with achievement gaps above the median for our study sample).

[25] National Center for Education Statistics. *The Condition of Education*. Accessed February 17, 2014 at https://nces.ed.gov/programs/coe/indicator_cgf.asp.

d. Schools with lower ELL achievement gaps (defined as schools with achievement gaps below the median).

To examine whether schools with higher ELL populations and higher ELL achievement gaps prioritized ELL-focused education reforms more than other schools, we conducted the following analyses:

3. We compared use of ELL-focused practices for schools that had higher and lower ELL populations within each of the following two groups of schools—those implementing a SIG-funded model and those not implementing one. For example, within the group of schools implementing a SIG-funded model, we compared schools with higher ELL populations to schools with lower ELL populations.
4. We compared use of ELL-focused practices for schools that had higher and lower ELL achievement gaps within each of the following two groups of schools—those implementing a SIG-funded model and those not implementing one.

Details on how these groups of schools were formed were presented in Chapter II.

Table VII.1 shows descriptive statistics on the ELL population and ELL achievement gap for each group of schools (those implementing a SIG-funded model and those not implementing one). On average, across all study schools, ELLs made up 17.4 percent of the student body and scored 0.33 standard deviations lower than other students on the state math assessment. Schools implementing a SIG-funded model had similar ELL populations to schools not implementing one (18.6 percent versus 15.8 percent of the student body). There were also no significant differences between schools implementing a SIG-funded model and schools not implementing one in the average ELL achievement gap.

Table VII.1. Distribution of ELL population and ELL achievement gap

	All study schools	Study schools implementing a SIG-funded intervention model in 2012–2013	Study Schools Not Implementing a SIG-Funded Intervention Model in 2012–2013
Distribution of ELL population			
10th percentile	0.5	0.3	1.1
50th percentile	12.5	13.9	11.2
90th percentile	42.2	44.2	36.1
Mean	17.4	18.6	15.8
Percentage of schools that had higher and lower ELL populations			
Higher	50.0	52.7	46.1
Lower	50.0	47.3	53.9
Distribution of ELL achievement gap			
10th percentile[a]	-0.23	-0.22	-0.28
50th percentile	0.34	0.33	0.38
90th percentile	0.80	0.82	0.79
Mean	0.33	0.31	0.36
Percentage of schools that had higher and lower ELL achievement gaps			
Higher	50.1	46.4	55.1
Lower	49.9	53.6	44.9
Number of schools	**370–400**	**210–240**	**160–170**

Source: State and district administrative records.

Note: We calculated the ELL population using student-level administrative data from 2009–2010 that contained indicators for whether each student participated in a program for ELLs. We calculated ELL achievement gaps as the average standardized score for non-ELLs minus the average standardized score for ELLs on the 2009–2010 state math assessment. All scores were standardized to have a standard deviation of 1, so results are reported in effect size units. Schools were classified into higher and lower groups based on whether their value (for either the ELL population or the ELL achievement gap) was above or below the median value across all study schools. See Chapter II for more details on how we classified schools into groups. Schools that had no ELLs were not included in the analysis. A range is provided for the sample size because missing data varied across items. There were no statistically significant differences between schools implementing a SIG-funded intervention model in 2012–2013 and schools not implementing one at the 0.05 level using a two-tailed test.

[a] The negative numbers in this row indicate that there were some schools in which ELLs scored higher than non-ELLs on average.

ELL = English language learner.

The evaluation's school administrator surveys asked about six ELL-focused practices aligned with SIG objectives (Table VII.2). More details about the individual ELL-focused practices are provided in Appendix G. Appendix G also displays findings on the extent to which *districts* reported using practices promoted by SIG that focused on ELLs.

Table VII.2. ELL-focused practices aligned with SIG objectives

Teachers have the opportunity to receive financial incentives designed to increase the number of staff with ELL expertise
Principals have the opportunity to receive financial incentives designed to increase the number of staff with ELL expertise
Using data on ELLs to inform and differentiate instruction
Implementing strategies (including additional supports or PD) to ensure that ELL students acquire language skills to master academic content
Providing additional services for ELLs (such as tutors, bilingual aides, or an after-school program)
Receiving supports from the state education agency or local education agency to use data on ELLs to improve or differentiate instruction

Source: SIG application; surveys of school administrators in spring 2012 and spring 2013.

Note: See Appendix G for a list of survey questions that were aligned with the ELL-focused practices in this table. All the practices listed in this table were included in the main analyses described in Chapter IV, but some of them are not listed in the Chapter IV tables because they were included in a broader practice that is listed in those tables.

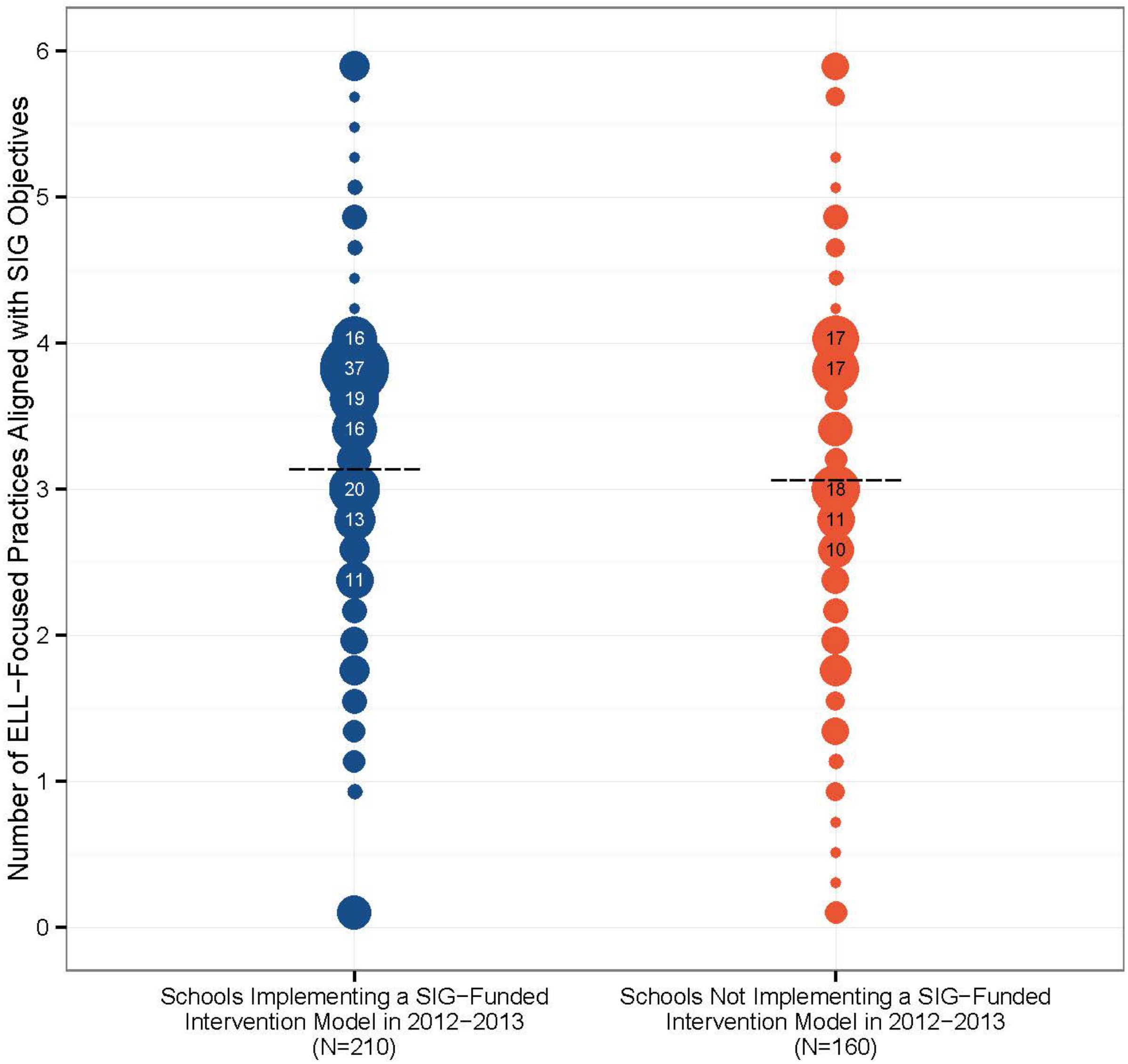
ELL = English language learner; PD = professional development.

In spring 2013, we found no differences in use of ELL-focused practices promoted by SIG based on schools' SIG model status. Both groups of schools (those implementing a SIG-funded model and those not implementing one) reported using an average of 3.1 of 6 ELL-focused practices (Figure VII.1). The RDD analysis also showed no evidence of an impact of implementing a SIG-funded model on ELL-focused practices (see Appendix A, Table A.8 for details).

There were also no significant differences by schools' SIG model status in use of ELL-focused practices within any of the ELL population or achievement gap subgroups. Among schools with higher ELL populations, schools implementing a SIG-funded model reported using 3.3 out of 6 ELL-focused practices promoted by SIG, compared to 3.2 practices for schools not implementing a SIG-funded model (Figure VII.2). Among schools with lower ELL populations, both groups of schools reported using 2.9 practices. Findings for schools with higher ELL achievement gaps and schools with lower achievement gaps were similar (Figure VII.3).

Thus, there is no evidence that schools implementing a SIG-funded model used more ELL-focused practices than schools not implementing such a model, even when we focused on subgroups of schools (with higher ELL populations and higher ELL achievement gaps) that might have prioritized ELL-related education reforms more than other groups.

Figure VII.1. Use of ELL-focused practices aligned with SIG objectives



Source: Surveys of school administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table VII.2. Each dot in this figure represents the number of schools that reported using a particular number of ELL-focused practices (out of six examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For example, 20 schools implementing a SIG-funded intervention model reported using three of six ELL-focused practices aligned with the SIG application criteria. For three of the ELL-focused practices, a "yes" response received one point. In the other three cases, it was possible for a school to receive a fraction of one point. See

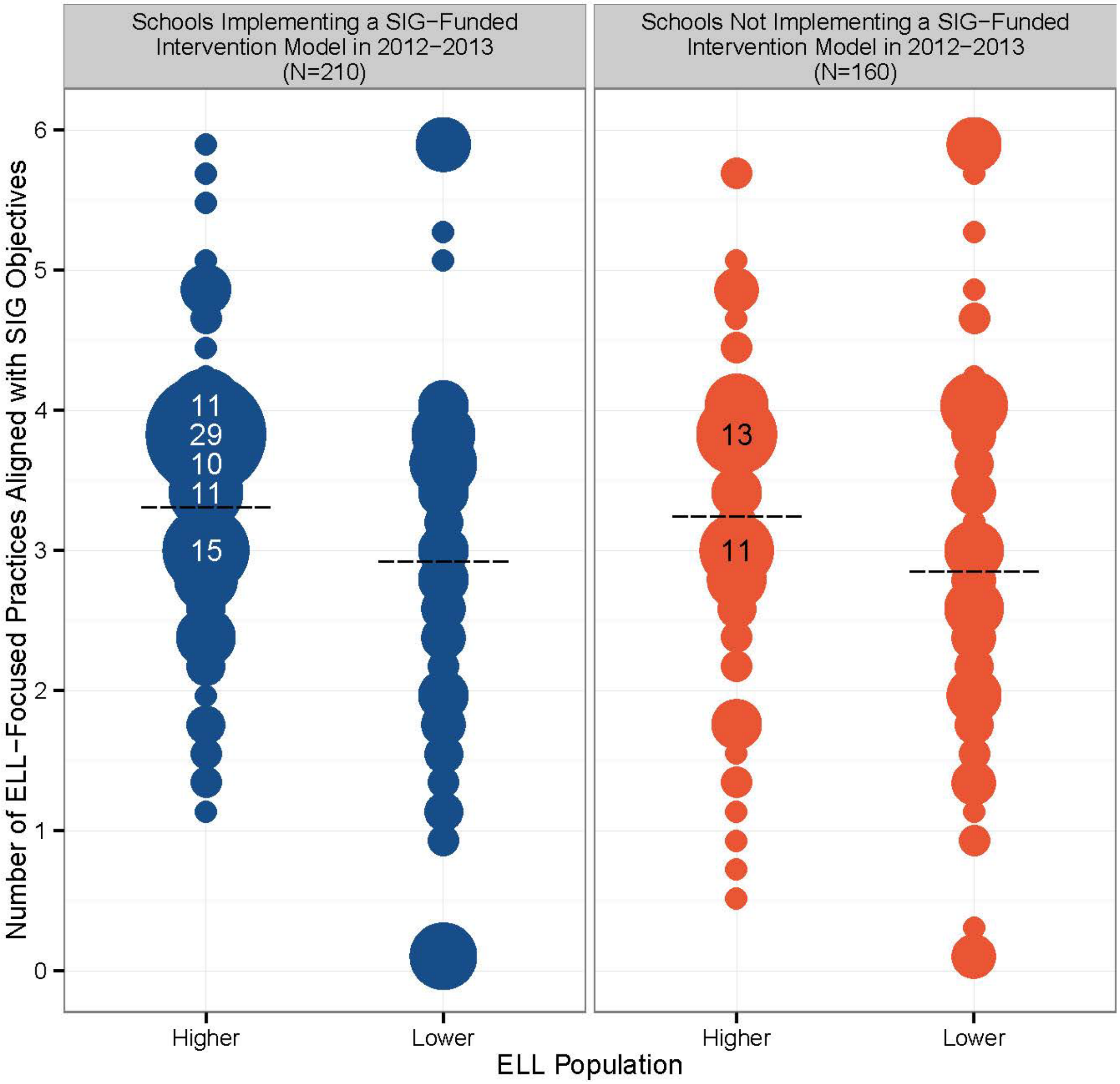
Appendix C for details on how we determined the number of ELL-focused practices for each school. The dashed line denotes the average number of ELL-focused practices for each group of schools. Schools that had no ELLs were not included in the analysis. The sample sizes in this figure are smaller than those in Table VII.1 because some schools had a missing value for all six ELL-focused practices and were therefore excluded from the analysis for this figure. There were no statistically significant differences between schools implementing a SIG-funded intervention model in 2012–2013 and schools not implementing one at the 0.05 level using a two-tailed test.

ELL = English language learner.

We found no evidence that schools with higher ELL populations prioritized the use of ELL-focused practices more than those with lower ELL populations. When focusing separately on schools by SIG-funded model status, there was no significant difference in use of ELL-focused practices based on schools' ELL populations. The differences between schools with higher and lower ELL populations were similar for each of the SIG status subgroups (0.4 practices among schools implementing a SIG-funded model and 0.3 practices among schools not implementing one; Figure VII.2).

However, we found that schools with higher ELL achievement gaps prioritized the use of ELL-focused practices promoted by SIG more than schools with lower gaps. Among schools implementing a SIG-funded model, schools with higher ELL achievement gaps reported using 0.3 more ELL-focused practices than schools with lower ELL achievement gaps (Figure VII.3). Among schools not implementing a SIG-funded model, schools with higher ELL achievement gaps also reported using 0.3 more ELL-focused practices than schools with lower ELL achievement gaps, but this difference was not statistically significant (the p-value was 0.08).

Figure VII.2. Use of ELL-focused practices aligned with SIG objectives, by ELL population

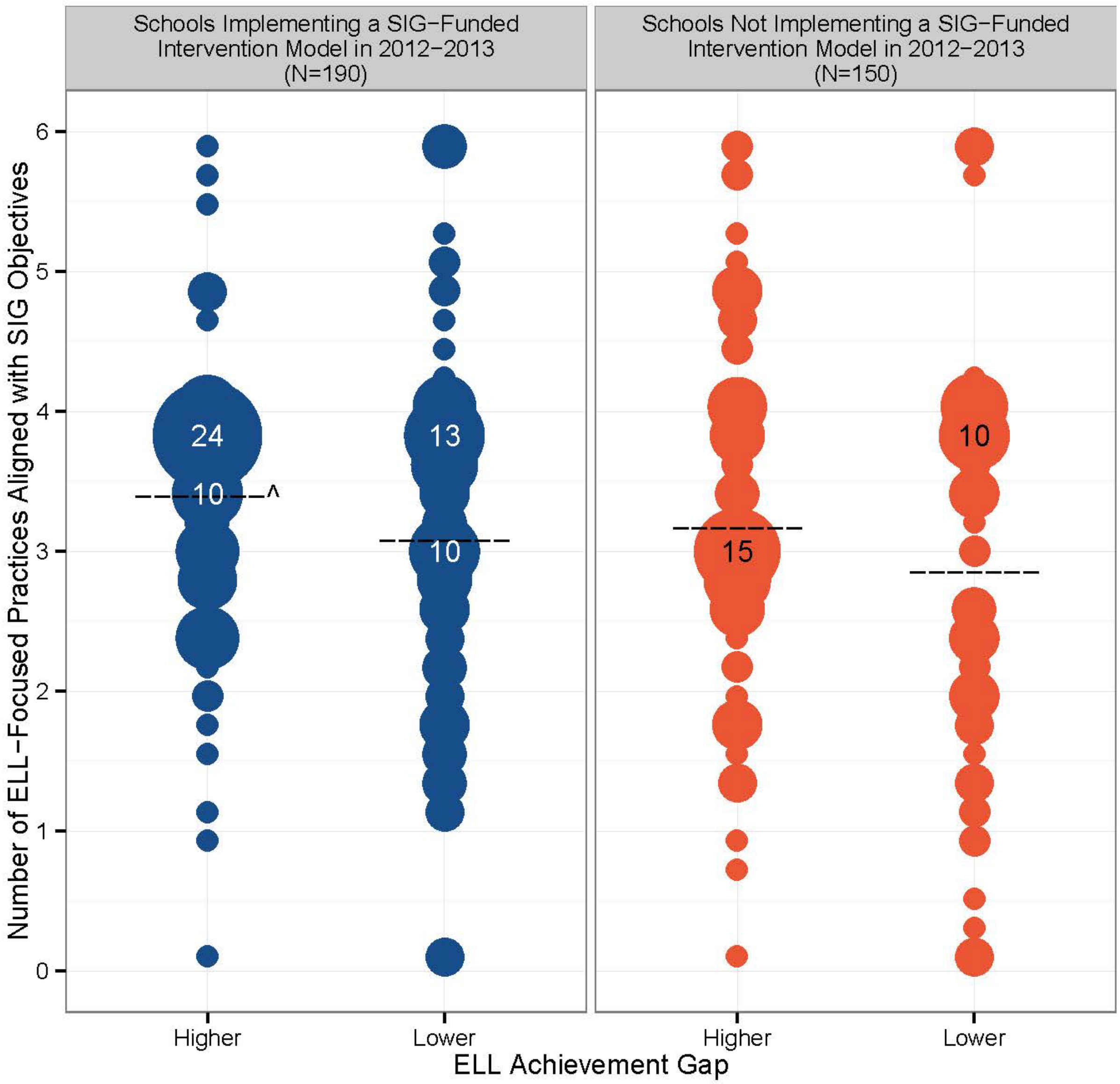


Source: Surveys of school administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table VII.2. Each column in the figure shows the number of ELL-focused practices that schools in each group reported using, by schools that had higher and lower ELL populations. Each dot in this figure represents the number of schools that reported using a particular number of ELL-focused practices (out of six examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For example, 15 schools implementing a SIG-funded intervention model that had a higher ELL population reported using three of six ELL-focused practices aligned with the SIG application criteria. For three of the practices, a "yes" response received one point. In the other three cases, it was possible for a school to receive a fraction of one point. See Appendix C for details on how we determined the number of ELL-focused practices for each school. The dashed line denotes the average number of ELL-focused practices for each group of schools. Schools that had no ELLs were not included in the analysis. The sample sizes in this figure are smaller than those in Table V.1 because some schools had a missing value for all six ELL-focused practices and were therefore excluded from the analysis for this figure. There were no statistically significant differences at the 0.05 level using a two-tailed test (1) between schools implementing and not implementing a SIG-funded intervention model with the same ELL population sizes, and (2) between higher and lower ELL population schools with the same SIG-funded intervention model implementation status.

ELL = English language learner.

Figure VII.3. Use of ELL-focused practices aligned with SIG objectives, by ELL achievement gap



Source: Surveys of school administrators in spring 2013.

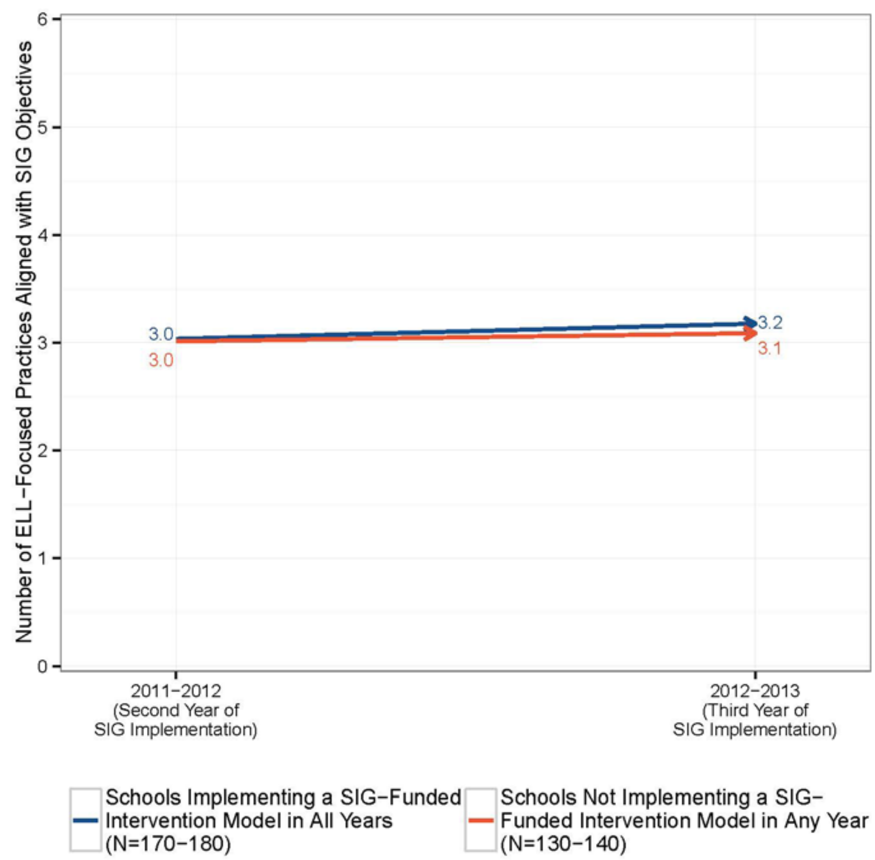
Note: The practices summarized in this figure are presented in Table VII.2. Each column in the figure shows the number of ELL-focused practices that schools in each group reported using, by schools that had higher and lower ELL achievement gaps. Each dot in this figure represents the number of schools that reported using a particular number of ELL-focused practices (out of six examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For example, 10 schools implementing a SIG-funded intervention model that had a lower ELL achievement gap reported using three of six ELL-focused practices aligned with the SIG application criteria. For three of the ELL-focused practices, a “yes” response received one point. In the other three cases, it was possible for a school to receive a fraction of one point. See Appendix C for details on how we determined the number of ELL-focused practices for each school. The dashed line denotes the average number of ELL-focused practices for each group of schools. Schools that had no ELLs were not included in the analysis. The sample sizes in this figure are smaller than those in Figures VII.1 and VII.2 because some schools with lower ELL populations did not have test score data available for ELLs, so the ELL achievement gap could not be calculated. There were no statistically significant differences at the 0.05 level using a two-tailed test between schools implementing and not implementing a SIG-funded intervention model with the same ELL achievement gap classification.

^Significantly different at the 0.05 level, two-tailed test, from schools with the same SIG-funded intervention model status but lower ELL achievement gaps. For example, among schools implementing a SIG-funded intervention model, schools with higher ELL achievement gaps reported using statistically significantly more ELL-focused practices than schools with lower ELL achievement gaps.

ELL = English language learner.

Changes over time in use of ELL-focused practices did not significantly differ by schools' SIG model status. In 2011–2012, both groups of schools reported using the same number of ELL-focused practices. Use of these practices was similar in 2012–2013 (Figure VII.4).

Figure VII.4. Change in use of ELL-focused practices aligned with SIG objectives



Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: This figure shows change over time for each group of schools in the use of ELL-focused practices aligned with the SIG application criteria. The arrow for each group of schools starts at the average number of reported practices aligned with the SIG application criteria in spring 2012 and ends at the average number of reported practices aligned with the SIG application criteria in spring 2013. For example, on average, schools implementing a SIG-funded model reported using 3.0 of the 6 ELL-focused practices aligned with the SIG application criteria in spring 2012 and 3.2 of these practices in spring 2013. The sample sizes in this figure differ from those in Figures VII.1, VII.2, and VII.3 because the analysis of change in practices between 2011–2012 and 2012–2013 focused on a slightly different sample of schools than the analysis of practices used in spring 2013. In particular, the analysis of change over time focused on schools that received grants in 2010 (cohort 1 schools) and continued to implement SIG-funded intervention models for all three years (2010–2011, 2011–2012, and 2012–2013), compared to non-SIG schools that didn't implement a SIG-funded intervention model in any of those three years. There were no statistically significant differences between schools implementing and not implementing a SIG-funded model with respect to changes between 2011–2012 and 2012–2013 in the number of ELL-focused practices used, at the 0.05 level using a two-tailed test. A range is provided for the sample sizes because nonresponse varied across years.

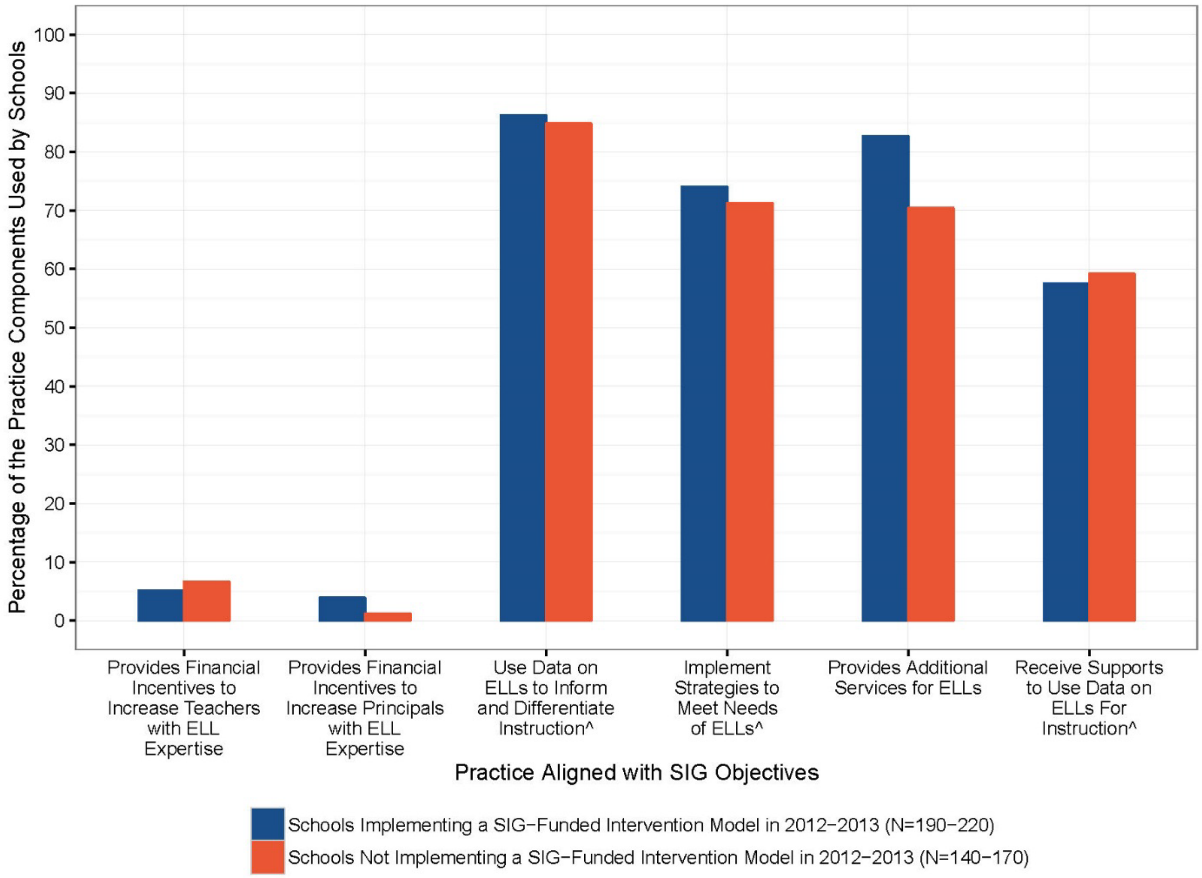
ELL = English language learner.

The ELL-focused practice with the largest difference between schools implementing a SIG-funded model and schools not implementing one was providing additional services for ELLs (such as tutors, bilingual aides, or an after-school program). Eighty-three percent of schools implementing a SIG-funded model and 70 percent of schools not implementing one used this practice (Figure VII.5).

The ELL-focused practices also differed in how widely they were used. Most study schools reported using data on ELLs to inform and differentiate instruction (Figure VII.5). However, very few study schools reported that teachers or principals had the opportunity to receive financial incentives designed to increase the number of staff with ELL expertise.

For the most part, the spring 2013 findings presented in this chapter were similar to the spring 2012 findings presented in an earlier report from this evaluation (Dragoset et al. 2015). In a few cases, the findings from one year were statistically significant, whereas the findings from the other year were not, but in both years, the differences between groups of schools with respect to the number of ELL-focused practices used were small (less than or equal to one practice).

Figure VII.5. Use of individual ELL-focused practices aligned with SIG objectives



Source: Surveys of school administrators in spring 2013.

Note: As described in Appendix C, for each ELL-focused practice aligned with the SIG application criteria for which we identified one or more survey questions aligned with the practice, we calculated the percentage of survey questions with a "yes" response as a measure of the percentage of components each school used. The height of each bar represents the average percentage of the components of the ELL-focused practice that each group of schools used. A range is provided for the sample sizes because nonresponse varied across items.

^Multiple survey questions were used to assess whether schools used all of the components of this practice.

ELL = English language learner

VIII. DISCUSSION OF MAIN FINDINGS

Despite the substantial resources devoted to SIG and its importance as an initiative seeking to turn around the nation's lowest-performing schools, few studies have examined the impact of SIG on student outcomes. In addition, few studies have examined the implementation of SIG-promoted practices in all of the topic areas described in the SIG application, and whether schools implementing a SIG-funded model used different practices than those that did not implement such a model. This evaluation sought to address these gaps in the existing literature by examining: (1) the extent to which schools implementing a SIG-funded model and schools not implementing one reported using SIG-promoted practices in all the topic areas described in the SIG application, (2) whether SIG-funded models affected the number of SIG-promoted practices used and student outcomes, and (3) whether the type of model implemented was related to changes in outcomes for low-performing schools. We also assessed the extent to which schools reported focusing on English language learners (ELLs) in their use of SIG-promoted practices.

In this chapter, we summarize our findings on SIG implementation (Section A) and impacts (Section B). We then lay out several questions of possible interest and potential explanations for these findings (Section C).

A. Findings on use of SIG-promoted practices

We examined the use of SIG-promoted practices in two ways. First, we conducted a descriptive analysis that compared the use of SIG-promoted practices in spring 2013 for schools that implemented a SIG-funded model in 2012–2013 and schools that did not. This analysis cannot conclusively establish whether receipt of SIG to implement a school intervention model caused schools to use SIG-promoted practices. It is possible that factors other than implementing a SIG-funded model, such as unobservable differences between the two groups of schools, may explain any observed differences in use of SIG-promoted practices. However, this analysis does shed light on the extent to which a large group of low-performing schools used SIG-promoted practices during the 2012–2013 school year. We also examined changes in use of these practices between 2011–2012 and 2012–2013.

Second, we examined whether implementation of a SIG-funded model in 2010–2011 affected the use of SIG-promoted practices in spring 2012 and spring 2013 for low-performing schools, using a regression discontinuity design (RDD). The advantage of the RDD analysis is that it provides rigorous evidence on whether SIG-funded models *caused* schools to use SIG-promoted practices. However, in contrast to the descriptive analysis, the RDD impacts apply only to the schools near the SIG eligibility cutoff—that is, schools near the lowest 5 percent of achievement in the state—so we do not know whether the impacts of SIG that we present in this report also apply to schools far away from these cutoff values.

Overall, we found that schools implementing SIG-funded models reported using more SIG-promoted practices than other schools, but we found no evidence that SIG *caused* those schools to implement more practices. Our descriptive analysis showed that schools implementing a SIG-funded model used significantly more SIG-promoted practices than other schools, though the difference was small. Our rigorous RDD findings showed a difference of similar size between these groups of schools, but the difference was not significant. Therefore, we cannot conclude that SIG *caused* the observed difference in use of practices. A possible explanation for this pattern of findings (similar differences but a discrepancy in statistical significance) is that the rigorous analysis was not able to detect the small difference in number of practices that we observed between groups of schools. Below, we summarize these two sets of findings.

1. Findings from the descriptive analysis

In spring 2013, schools implementing a SIG-funded model reported using more SIG-promoted practices than schools not implementing one in all four topic areas considered. This finding is consistent with prior studies that showed SIG schools reported implementing school improvement practices more intensively than non-SIG schools (Center on Education Policy 2012a; Dragoset et al. 2015). Although this finding focuses on averages for each group of schools, there was substantial variation within each group in the number of SIG-promoted practices used. We found that the differences between the two groups of schools in the average number of SIG-promoted practices used were statistically significant in all four areas.

Adding up the differences across the four areas, schools implementing a SIG-funded model reported using more SIG-promoted practices overall (22.8 out of 35 total practices, or 65 percent) than schools not implementing one (20.3 practices, or 58 percent), a difference of 2.5 practices. It is not clear whether a difference of this size would be meaningful in its overall influence on improvement practices and school outcomes.

In spring 2013, across both groups of schools, use of SIG-promoted practices was highest in the comprehensive instructional reform strategies area—which is consistent with findings from an earlier study (Dragoset et al. 2015)—and lowest in the operational flexibility and support area. Study schools reported using, on average, 7.1 of the 8 SIG-promoted practices in the comprehensive instructional reform strategies area (89 percent) and 0.87 of the 2 SIG-promoted practices in the operational flexibility and support area (43 percent).

Across all topic areas, the use of individual practices varied widely. Nearly all study schools reported using benchmark or interim assessments at least once per year (a practice in the comprehensive instructional reform strategies area). In contrast, very few study schools reported (1) using teacher evaluation results to inform decisions about compensation, (2) using principal evaluation results to inform decisions about compensation, or (3) using financial incentives to recruit and retain effective principals (practices in the teacher and principal effectiveness topic area). The practice with the largest difference between schools implementing a SIG-funded model and schools not implementing one was reviewing competencies of staff or replacing instructional staff, which is a practice in the teacher and principal effectiveness topic area (with schools implementing a SIG-funded model using 43 percent of the components of this practice, on average, and schools not implementing one using 27 percent).

When we focused on changes over time in schools' use of SIG-promoted practices, we found a significant difference between the two groups of schools in one of the four areas—increasing learning time and creating community-oriented schools. In this area, there was a larger decline between 2011–2012 and 2012–2013 for schools implementing a SIG-funded model than for schools not implementing one.

We found no significant difference in use of ELL-focused practices promoted by SIG between schools implementing a SIG-funded model and schools not implementing one.

Overall, the spring 2013 findings presented in this report were the same as the spring 2012 findings presented in an earlier report from this evaluation (Dragoset et al. 2015).

2. Findings from the RDD analysis

We found no evidence that SIG-funded models affected the number of SIG-promoted practices used. In particular, implementation of a SIG-funded model in 2010–2011 had no statistically significant impact on the total number of SIG-promoted practices used in either spring 2012 or spring 2013, for schools near the SIG eligibility cutoff. The difference in number of practices used between schools that just met the eligibility cutoff and schools that just missed the cutoff was 0.4 practices (1 percent of the 35 total practices examined) in 2012 and 3.3 practices (9 percent of the 35 total practices examined) in 2013, but neither difference was significant. SIG-funded models also had no impact on the number of SIG-promoted practices used in any of the four topic areas.

B. Findings on whether SIG-funded models improved student outcomes

We also found no evidence that SIG-funded models affected student outcomes. Specifically, using a rigorous RDD analysis, we found no significant impacts of SIG-funded models overall on math or reading test scores, high school graduation, or college enrollment, for schools near the SIG eligibility cutoff. In addition, there were no significant impacts of SIG-funded models on student outcomes within student and school subgroups. This finding is perhaps not surprising, given that we found no significant impacts of SIG-funded models on the number of SIG-promoted practices used.

In a correlational analysis that examined specific SIG models, we found that the turnaround model was associated with larger student achievement gains than the transformation model but only for some grades. Specifically, the turnaround model was associated with larger student achievement gains than the transformation model in grades 6 through 12. In contrast, in grades 2 through 5, we found no evidence that one model was associated with larger improvements in student achievement than another. However, it is worth noting that these associations may be due to factors other than the SIG model implemented, such as unobservable differences between schools implementing different models.

C. Questions of interest and potential explanations for findings

Readers may have questions about the findings reported here. Below, we lay out some questions of possible interest and potential explanations for the findings.

Why were there no significant RDD impacts of SIG-funded models on the use of SIG-promoted practices, even though the descriptive analysis showed significant differences in use of practices between schools that did and did not implement a SIG-funded model? One possible explanation for this pattern of findings is that our statistical power (that is, our ability to detect differences between the two groups of schools) was lower for the RDD analysis than for the descriptive analysis. The smallest difference that the RDD could detect in spring 2013 was 5.2 practices, so the RDD could not have detected the difference of 3.3 practices that we found in the RDD analysis in that year. In contrast, we were able to detect the difference of 2.5 practices that we observed in the descriptive analysis.

Why were there no significant impacts of SIG-funded models on student outcomes? SIG provides funding to help schools use practices that it believes will improve student outcomes. Therefore, SIG would not be expected to improve student outcomes if it did not substantially increase the use of SIG-promoted practices or the funds available to schools. Indeed, we found no evidence that SIG-funded models substantially increased the use of SIG-promoted practices. (Because of the low statistical power for the RDD analysis, discussed previously, we cannot rule out the possibility that SIG-funded models led to small increases in the number of practices used.) Even if implementing the models *had* increased the use of these practices, the models could have had no impact on student achievement if the practices were not well-implemented or ineffective. For example, previous literature provides mixed evidence on the effectiveness of some of these practices at raising student achievement. SIG-funded models could also have no impact on per-pupil spending if districts reallocated other existing funds away from SIG grantees and toward non-grantees. We investigated this hypothesis, and the findings from that investigation were inconclusive (see Chapter VI).

How do our impact findings on student outcomes compare to other rigorous studies of SIG? Whereas we found no significant impacts of SIG-funded models on student outcomes using an RDD, three other studies found positive impacts of SIG on student achievement (Dee 2012; LiCalsi et al. 2015; and Gold et al. 2012). Our estimates, which were not significant, equaled student-level effect sizes of 0.01 standard deviations in math and 0.08 standard deviations in reading. In contrast, the significant student-level effect sizes were 0.10 standard deviations for an academic performance index in Dee (2012) and 0.22 standard deviations in math and reading in LiCalsi et al. (2015). Gold et al. (2012) reported grade-level effect sizes of 1.11 standard deviations in math and 0.83 standard deviations in reading. Differences between our study samples and analysis methods may explain these differences in findings.

There are several benefits of our study relative to Dee (2012), LiCalsi et al. (2015), and Gold et al. (2012). In particular, the RDD component of our study included more than 20 states, compared with just one state for the other three studies. In addition, our methods were more rigorous than those used by LiCalsi et al. (2015) and Gold et al. (2012). Finally, unlike Dee (2012) and Gold et al. (2012), we had student-level data, so we were able to estimate impacts that accounted for student mobility.

Dee (2012) used an RDD and found a positive impact of SIG on a school-level academic performance index that was based on student tests in English, math, social studies, and science. Our study also used an RDD, but it differs from Dee's 2012 study in several ways. First, Dee (2012) focused on a single state (California), whereas our RDD analysis included more than 20 states.[26] Second, our study used student-level data while Dee (2012) used school-level data. Student-level data allowed us to calculate impacts based on where students actually attended

[26] Dee (2012) used all school districts in California and had a sample of nearly 3,000 schools. We restricted our sample to districts where large numbers of schools were eligible for inclusion in our RDD analysis and where high proportions of schools eligible for SIG actually received SIG funds. Our RDD sample of about 10 to 200 schools within each state included schools within and outside the bandwidth, for a total of 850 schools.

school in each outcome year and impacts that accounted for student mobility based on the schools students were slated to attend in each outcome year. Third, unlike Dee (2012), our analysis method accounted for bandwidth selection when determining the significance of our findings and aggregated impacts that we had estimated separately by grade. Finally, Dee (2012) focused on an academic performance index that combined scores across multiple subjects, whereas we examined math and reading outcomes separately.

LiCalsi et al. (2015) used a less rigorous, comparative interrupted time series approach and found positive impacts of SIG on student achievement. They focused on schools in eight districts in a single state (Massachusetts). Gold et al. (2012) used multilevel multivariate regression methods, also less rigorous than an RDD analysis, to assess school-level impacts on student achievement for 11 SIG schools and 72 comparison schools in the Philadelphia area.

Why did schools implementing a SIG-funded model in 2012–2013 not report using more SIG-promoted practices? In two areas (developing and increasing teacher and principal effectiveness, and having operational flexibility and receiving support), schools implementing a SIG-funded model reported using only about half of the SIG-promoted practices. One possible explanation for these results is that schools encountered barriers in using practices in these areas. For example, schools implementing a SIG-funded model reported using about one-third of the practices in the following two subtopics in the area of developing and increasing teacher and principal effectiveness: (1) identifying and rewarding effective teachers and principals and removing ineffective ones, and (2) implementing strategies to recruit, place, and retain staff. Prior literature on SIG implementation identified challenges to using practices in these subtopics such as requirements in collective bargaining agreements, cumbersome district hiring and placement procedures, and shortages of qualified teachers and leaders (Center on Education Policy 2012a; Yatsko et al. 2012). A second possibility is that schools chose to focus their efforts on a select group of practices in each area rather than using all of them, perhaps because they had technical capacity constraints or because they considered certain practices more important than others to their particular reform agenda and needs.

Why did schools implementing a SIG-funded model in 2012–2013 report a larger decrease in the number of learning time and community-oriented schools practices used than schools not implementing a SIG-funded model? One possible explanation is that schools initially used practices in this area in response to SIG requirements and funding but ceased using these practices by spring 2013 after experiencing implementation challenges. For example, case study research found that extended learning time can be costly to implement and can cause fatigue among teachers and students (McMurrer et al. 2015). Another possible explanation is that schools implementing a SIG-funded model used more practices in this area prior to SIG or in the 2011–2012 school year, but chose to focus on practices in other areas by spring 2013. A third possibility is that some of the practices in this area (such as changing discipline policies) were one-time events that occurred in the 2011–2012 school year and did not need to occur again in the following year.

Why did we find that the turnaround model was associated with larger student achievement gains than the transformation model for grades 6 through 12? In contrast to the RDD analysis, which examined the overall impact of SIG-funded models without regard to the type of model implemented, our correlational analysis compared the various models to each other. In that analysis, we found that, for higher grades, the turnaround model might be more effective at improving student achievement in math than the transformation model. This finding is consistent with findings from a study of the effectiveness of SIG in California, which found that achievement gains among SIG schools were concentrated in schools that implemented the turnaround model (Dee 2012). One possible explanation for this finding is that the turnaround model is more effective than the transformation model at improving student achievement in higher grades, perhaps because the turnaround model involved more intensive reforms than the transformation model. In particular, although both models required schools to replace the principal, turnaround schools were also required to screen all existing staff and rehire no more than 50 percent (U.S. Department of Education 2010a). The turnaround schools in our sample were more likely to report using this practice than the transformation schools. Another possible explanation is that the difference in achievement gains between turnaround and transformation schools was due to factors other than the model implemented; for example, schools that had the motivation to select the more intensive turnaround model may have had the drive to make large gains even in the absence of the model.

Although we cannot definitively accept or reject any of these possible explanations for these findings, we offer them as starting points for future investigations into the implementation and impact of SIG.

REFERENCES

Abdulkadiroglu, A., J. D. Angrist, S. M. Dynarski, T. J. Kane, and P. A. Pathak. "Accountability and Flexibility in Public Schools: Evidence from Boston's Charters and Pilots." *Quarterly Journal of Economics,* vol. 126, no. 2, May 2011, pp. 699–748.

Allen, J. P., R. C. Pianta, A. Gregory, A. Y. Mikami, and J. Lun. "An Interaction-Based Approach to Enhancing Secondary School Instruction and Student Achievement." *Science,* vol. 333, no. 6045, August 2011, pp. 1034–1037.

Angrist, J. D., P. A. Pathak, and C. R. Walters. "Explaining Charter School Effectiveness." NBER working paper series no. 17332. Cambridge, MA: National Bureau of Economic Research, August 2011 (revised April 2012).

Betts, J., A. Zau, and K. King. "From Blueprint to Reality: San Diego's Education Reforms." San Francisco, CA: Public Policy Institute of California, 2005.

Bifulco, R., W. Duncombe, and J. Yinger. "Does Whole-School Reform Boost Student Performance? The Case of New York City." Paper 108. Syracuse, NY: Center for Policy Research, 2003. Available at http://surface.syr.edu/cpr/108. Accessed February 13, 2013.

Black, A. R., M. A. Somers, F. Doolittle, R. Unterman, and J. B. Grossman. "The Evaluation of Enhanced Academic Instruction in After-School Programs: Final Report." Publication no. NCEE 2009-4077. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, September 2009.

Booker, K., B. T. Gross, and D. Goldhaber. "Boosting Student Achievement: The Effect of Comprehensive School Reform on Student Achievement." *Educational Evaluation and Policy Analysis,* vol. 31, no. 2, 2009, pp. 111–126.

Borman, G. D., G. M. Hewes, L. T. Overman, and S. Brown. "Comprehensive School Reform and Achievement: A Meta-Analysis." *Review of Educational Research,* vol. 73, no. 2, 2003, pp. 125–230.

Calonico, S., M. D. Cattaneo, and R. Titiunik. "Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs." *Econometrica*, vol. 82, no. 6, 2014, pp. 2295–2326.

Carlson, D., G. D. Borman, and M. Robinson. "A Multistate District-Level Cluster Randomized Trial of the Impact of Data-Driven Reform on Reading and Mathematics Achievement." *Educational Evaluation and Policy Analysis,* vol. 33, no. 3, 2011, pp. 378–398.

Center on Education Policy. "Early State Implementation of Title I School Improvement Grants Under the Recovery Act." Washington, DC: Center on Education Policy, February 2011a.

Center on Education Policy. "Federal Efforts to Improve the Lowest-Performing Schools: District Views on School Improvement Grant Requirements." Washington, DC: Center on Education Policy, November 2011b.

Center on Education Policy. "Opportunities and Obstacles: Implementing Stimulus-Funded School Improvement Grants in Maryland, Michigan, and Idaho." Washington, DC: Center on Education Policy, March 2012a.

Center on Education Policy. "State Implementation and Perceptions of Title I School Improvement Grants Under the Recovery Act: One Year Later." Washington, DC: Center on Education Policy, March 2012b.

Clark, D. "The Performance and Competitive Effects of School Autonomy." *The Journal of Political Economy,* vol. 117, no. 4, 2009, pp. 745–783.

Clark, M. A., H. S. Chiang, T. Silva, S. McConnell, K. Sonnenfeld, A. Erbe, and M. Puma. "The Effectiveness of Secondary Math Teachers from Teach For America and the Teaching Fellows Programs." NCEE 2013-4015. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2013.

Constantine, J., D. Player, T. Silva, K. Hallgren, M. Grider, and J. Deke. "An Evaluation of Teachers Trained Through Different Routes to Certification. Final Report." Princeton, NJ: Mathematica Policy Research, 2009.

Cortes, K. E., J. Goodman, and T. Nomi. "Doubling Up: Intensive Math Instruction and Educational Attainment." Harvard University working paper. Cambridge, MA: Harvard University, 2012.

Council of the Great City Schools. "School Improvement Grants: Progress Report from America's Great City Schools." Washington, DC: Council of the Great City Schools, 2015.

Decker, P. T., D. P. Mayer, and S. Glazerman. "The Effect of Teach For America on Students: Findings from a National Evaluation." Princeton, NJ: Mathematica Policy Research, 2004.

Dee, T. "School Turnarounds: Evidence from the 2009 Stimulus." Cambridge, MA: National Bureau of Economic Research, 2012. Available at http://www.nber.org/papers/w17990.pdf. Accessed February 13, 2013.

de la Torre, M., and J. Gwynne. "When Schools Close: Effects on Displaced Students in Chicago Public Schools." Chicago, IL: Consortium on Chicago School Research at the University of Chicago, 2009.

Dobbie, W., and R. G. Fryer, Jr. "Are High-Quality Schools Enough to Increase Achievement Among the Poor? Evidence from the Harlem Children's Zone." *American Economic Journal: Applied Economics,* vol. 3, no. 3, 2011, pp. 158–187.

Dragoset, L., S. James-Burdumy, K. Hallgren, I. Perez-Johnson, M. Herrmann, C. Tuttle, M. H. Angus, R. Herman, M. Murray, C. Tanenbaum, and C. Graczewski. "Usage of Practices Promoted by School Improvement Grants." NCEE 2015-4019. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2015.

Fryer, Jr., R. G. "Injecting Charter School Best Practices into Traditional Public Schools: Evidence from Field Experiments." Cambridge, MA: Harvard University, 2014. Available at http://scholar.harvard.edu/files/fryer/files/2014_injecting_charter_school_best_practices_into_traditional_public_schools.pdf. Accessed November 20, 2014.

Furgeson, J., B. Gill, J. Haimson, A. Killewald, M. McCullough, I. Nichols-Barrer, B. Teh, N. Verbitsky-Savitz, M. Bowen, A. Demerrit, P. Hill, and R. Lake. "Charter School Management Organizations: Diverse Strategies and Diverse Student Impacts." Cambridge, MA: Mathematica Policy Research, 2012.

Garet, M. S., A. J. Wayne, F. Stancavage, J. Taylor, K. Walters, M. Song, S. Brown, S. Hurlburt, P. Zhu, S. Sepanik, and F. Doolittle. "Middle School Mathematics Professional Development Impact Study: Findings After the First Year of Implementation." NCEE 2010-4009. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2010.

Gelman, A., and G. Imbens. "Why High-order Polynomials Should not be Used in Regression Discontinuity Designs." NBER working paper series no. 20405. Cambridge, MA: National Bureau of Economic Research, 2014.

Glazerman, S., D. Mayer, and P. Decker. "Alternative Routes to Teaching: The Impacts of Teach For America on Student Achievement and Other Outcomes." *Journal of Policy Analysis and Management,* vol. 25, no. 1, 2006, pp. 75–96.

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

Gold, E., M. H. Norton, D. Good, and S. Levin. "Philadelphia's Renaissance Schools Initiative: 18 Month Interim Report." Philadelphia, PA: Research for Action, 2012.

Gwet, K. L. Handbook of Inter-Rater Reliability, Fourth Edition. Gaithersburg, MD: Advanced Analytics, LLC, 2014.

Heissel, J. A., and H. F. Ladd. "School Turnaround in North Carolina: A Regression Discontinuity Analysis." CALDER working paper series no. 156. Washington, DC: National Center for Analysis of Longitudinal Data in Education Research, March 2016.

Henderson, S., A. Petrosino, S. Guckenburg, and S. Hamilton. "Measuring How Benchmark Assessments Affect Student Achievement." REL technical brief, REL 2007–No. 039. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands, 2007. Available at http://ies.ed.gov/ncee/edlabs. Accessed February 13, 2013.

Henderson, S., A. Petrosino, S. Guckenburg, and S. Hamilton. "A Second Follow-Up Year for Measuring How Benchmark Assessments Affect Student Achievement." REL technical brief REL 2008–No. 002. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, Regional Educational Laboratory Northeast and Islands, April 2008. Available at http://ies.ed.gov/ncee/edlabs/regions/northeast/pdf/techbrief/tr_00208.pdf. Accessed February 13, 2013.

Herman, R., C. Graczewski, S. James-Burdumy, M. Murray, I. Perez-Johnson, and C. Tanenbaum. "Operational Authority, Support, and Monitoring of School Turnaround." NCEE 2014-4008. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2014.

Herrmann, M., L. Dragoset, and S. James-Burdumy. "Are Low-Performing Schools Adopting Practices Promoted by School Improvement Grants?" NCEE 2015-4001. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2014.

Hill, C. J., H. S. Bloom, A. R. Black, and M. W. Lipsey. "Empirical benchmarks for interpreting effect sizes in research." *Child Development Perspectives*, vol. 2, no. 3, 2008, pp. 172-177.

Hoxby, C. M., and J. E. Rockoff. "Findings from the City of Big Shoulders." *Education Next,* vol. 5, no. 4, 2005, pp. 52–58.

Hoxby, C. M., S. Murarka, and J. Kang. "How New York City's Charter Schools Affect Student Achievement: August 2009 Report." Cambridge, MA: New York City Charter Schools Evaluation Project, September 2009.

Hurlburt, S., K. C. Le Floch, S. B. Therriault, and S. Cole. "Baseline Analyses of SIG Applications and SIG-Eligible and SIG-Awarded Schools." NCEE 2011-4019. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2011.

Hurlburt, S., S. B. Therriault, and K. C. Le Floch. "School Improvement Grants: Analyses of State Applications and Eligible and Awarded Schools." NCEE 2012-4060. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2012.

Imbens, G., and J. Wooldridge. "Lecture Notes 3." Cambridge, MA: National Bureau of Economic Research, 2007.

Imbens, G., and T. Lemieux. "Regression Discontinuity Designs: A Guide to Practice." *Journal of Econometrics*, vol. 142, no. 2, 2008, pp. 615–635.

Imbens, G., and K. Kalyanaraman. "Optimal Bandwidth Choice for the Regression Discontinuity Estimator." *Review of Economic Studies*, vol. 79, no. 3, 2012, pp. 933–959.

James-Burdumy, S., M. Dynarski, M. Moore, J. Deke, W. Mansfield, and C. Pistorino. "When Schools Stay Open Late: The National Evaluation of the 21st Century Community Learning Centers Program: Final Report." Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, April 2005.

Kemple, J. J. "High School Closures in New York City: Impacts on Students' Academic Outcomes, Attendance, and Mobility." New York, NY: Research Alliance for New York City Schools, 2015.

Klein, Alyson. "ESEA Reauthorization: The Every Student Succeeds Act Explained." *Education Week*, November 30, 2015.

Lee, D. S., and D. Card. "Regression Discontinuity Inference with Specification Error." *Journal of Econometrics*, vol. 142, 2008, pp. 655–674.

LiCalsi, C., M. Citkowicz, L. B. Friedman, and M. Brown. “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, June 2015.

May, H., and M. A. Robinson. “A Randomized Evaluation of Ohio’s Personalized Assessment Reporting System (PARS).” Philadelphia, PA: Consortium for Policy Research in Education, 2007.

McCrary, J. “Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test.” *Journal of Econometrics*, vol. 142, no. 2, 2008, pp. 698–714.

McMurrer, J., M. Frizzell, N. Yoshioka, S. Scott, and N. Ostler. “Expanded Learning Time: A Summary of Findings from Case Studies in Four States.” Washington, DC: Center on Education Policy, 2015.

Mertler, C. A. “Using Standardized Test Data to Guide Instruction and Intervention.” ERIC Digest, 2002.

National Alliance for Public Charter Schools. “Dashboard: A Comprehensive Data Resource from the National Alliance for Public Charter Schools.” Washington, DC: National Alliance for Public Charter Schools, 2013. Available at http://dashboard2.publiccharters.org/Home/. Accessed June 3, 2013.

Player, D., and V. Katz. “School Improvement in Ohio and Missouri: An Evaluation of the School Turnaround Specialist Program.” Working Paper. Charlottesville, VA: Center on Education Policy and Workforce Competitiveness, 2013.

Quint, J. C., S. Sepanik, and J. K. Smith. “Using Student Data to Improve Teaching and Learning: Findings from an Evaluation of the Formative Assessments of Student Thinking in Reading (FAST-R) Program in Boston Elementary Schools.” New York, NY: MDRC, 2008. Available at http://www.mdrc.org/sites/default/files/full_601.pdf. Accessed February 13, 2013.

Reardon, S. F., and J. P. Robinson. “Regression Discontinuity Designs with Multiple Rating-Score Variables.” Working paper. Palo Alto, CA: Stanford University, 2010.

Slavin, R., A. Cheung, G. Holmes, N. A. Madden, and A. Chamberlain. “Effects of a Data-Driven District Reform Model.” Baltimore, MD: Success for All Foundation, 2011.

Steinberg, M. “Does Greater Autonomy Improve School Performance? Evidence from a Regression Discontinuity Analysis in Chicago.” *Education Finance and Policy,* vol. 9, no. 1, 2014, pp. 1-35.

Stock J., and M. Yogo. “Testing for Weak Instruments in Linear IV Regression.” In: *Identification and Inference for Econometric Models*. New York: Cambridge University Press, 2005.

Strunk, K. O., J. A. Marsh, A. Hashim, S. Bush, and T. Weinstein. "The Efficacy of the Los Angeles Unified School District Public School Choice Initiative for Student Achievement Outcomes: Early Evidence from the First Year." Los Angeles: University of Southern California, 2012. Available at http://www.officeforeducationpolicy.org/wp-content/uploads/2012/08/strunk_et_al-the_efficacy_of_the_lausd_psci_for_student_achievement_outcomes-early_evidence_from_the_first_year.pdf Accessed November 20, 2014.

Tanenbaum, C., A. Boyle, C. Graczewski, S. James-Burdumy, L. Dragoset, and K. Hallgren. "State Capacity to Support School Turnaround." NCEE 2015-4012. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2015.

Troppe, P., A. Milanowski, R. Garrison-Mogren, A. Webber, B. Gutmann, F. Reisner, and M. Goertz, "State, District, and School Implementation of Reforms Promoted Under the Recovery Act: 2009-10 through 2011-12." NCEE 2015-4016. Washington, DC: U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance, 2015.

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

U.S. Congress. "School Improvement Grants—American Recovery and Reinvestment Act of 2009; Title I of the Elementary and Secondary Education Act of 1965; Notice of Proposed Requirements." *Federal Register,* vol. 74, no. 164, August 26, 2009, pp. 43101–43114. Available at http://www.gpo.gov/fdsys/pkg/FR-2009-08-26/html/E9-20612.htm. Accessed February 13, 2013.

U.S. Department of Education. "School Improvement Grants Application, Section 1003(g) of the Elementary and Secondary Education Act." Washington, DC: U.S. Department of Education, 2010a.

U.S. Department of Education. Standards for Regression Discontinuity Designs. *What Works Clearinghouse,* 2010b.

U.S. Department of Education. "California: Targeted Monitoring Review of School Improvement Grants, March 7–11, 2011." Washington, DC: U.S. Department of Education, 2011a. Available at http://www2.ed.gov/admins/lead/account/monitoring/reports11/index.html. Accessed February 13, 2013.

U.S. Department of Education. "Indiana: Targeted Monitoring Review of School Improvement Grants, March 14–18, 2011." Washington, DC: U.S. Department of Education, 2011b. Available at http://www2.ed.gov/admins/lead/account/monitoring/reports11/index.html. Accessed February 13, 2013.

U.S. Department of Education. "Maine: Targeted Monitoring Review of School Improvement Grants, March 21–25, 2011." Washington, DC: U.S. Department of Education, 2011c. Available at http://www2.ed.gov/admins/lead/account/monitoring/reports11/index.html. Accessed February 13, 2013.

U.S. Department of Education. “Michigan: Targeted Monitoring Review of School Improvement Grants, April 4–8, 2011.” Washington, DC: U.S. Department of Education, 2011d. Available at http://www2.ed.gov/admins/lead/account/monitoring/reports11/index.html. Accessed February 13, 2013.

U.S. Department of Education. “Minnesota: Targeted Monitoring Review of School Improvement Grants, May 2–6, 2011.” Washington, DC: U.S. Department of Education, 2011e. Available at http://www2.ed.gov/admins/lead/account/monitoring/reports11/index.html. Accessed February 13, 2013.

U.S. Department of Education. “Mississippi: Targeted Monitoring Review of School Improvement Grants, May 2–6, 2011.” Washington, DC: U.S. Department of Education, 2011f. Available at http://www2.ed.gov/admins/lead/account/monitoring/reports11/index.html. Accessed February 13, 2013.

U.S. Department of Education. “Nebraska: Targeted Monitoring Review of School Improvement Grants, May 9–13, 2011.” Washington, DC: U.S. Department of Education, 2011g. Available at http://www2.ed.gov/admins/lead/account/monitoring/reports11/index.html. Accessed February 13, 2013.

U.S. Department of Education. “Nevada: Targeted Monitoring Review of School Improvement Grants, February 14–18, 2011.” Washington, DC: U.S. Department of Education, 2011h. Available at http://www2.ed.gov/admins/lead/account/monitoring/reports11/index.html. Accessed February 13, 2013.

U.S. Department of Education. “Pennsylvania: Targeted Monitoring Review of School Improvement Grants, February 28–March 4, 2011.” Washington, DC: U.S. Department of Education, 2011i. Available at http://www2.ed.gov/admins/lead/account/monitoring/reports11/index.html. Accessed February 13, 2013.

U.S. Department of Education. “South Dakota: Targeted Monitoring Review of School Improvement Grants, May 16–20, 2011.” Washington, DC: U.S. Department of Education, 2011j. Available at http://www2.ed.gov/admins/lead/account/monitoring/reports11/index.html. Accessed February 13, 2013.

U.S. Department of Education. “U.S. Department of Education Releases Early Snapshot of School Improvement Grants Data [press release].” Washington, DC: U.S. Department of Education, November 19, 2012. Available at http://www.ed.gov/news/press-releases/us-department-education-releases-early-snapshot-school-improvement-grants-data. Accessed February 13, 2013.

U.S. Department of Education. Preview of Regression Discontinuity Design Standards. *What Works Clearinghouse,* 2015.

U.S. Department of Education, Office of the Inspector General. "School Improvement Grants: Selected States Generally Awarded Funds Only to Eligible Schools. Final Audit Report." Washington, DC: U.S. Department of Education, Office of the Inspector General, March 2012. Available at http://www2.ed.gov/about/offices/list/oig/auditreports/fy2012/a05l0002.pdf. Accessed February 13, 2013.

U.S. Government Accountability Office. "School Improvement Grants: Early Implementation Under Way, but Reforms Affected Short Time Frames." Report no. GAO-11-741. Washington, DC: U.S. Government Accountability Office, July 2011. Available at http://www.gao.gov/new.items/d11741.pdf. Accessed February 13, 2013.

Yatsko, S., R. Lake, E. C. Nelson, and M. Bowen. "Tinkering Toward Transformation: A Look at Federal School Improvement Grant Implementation." Seattle, WA: Center on Reinventing Public Education, 2012.

Zimmer, R., and R. Buddin. "Charter School Performance in Urban Districts." *Journal of Urban Economics*, vol. 60, no. 2, 2006, pp. 307–326.

Zimmer, R., B. Gill, K. Booker, S. Lavertu, and J. Witte. "Examining Charter Student Achievement Effects Across Seven States." *Economics of Education Review*, vol. 31, no. 2, 2012, pp. 213–224.

APPENDIX A

RDD IMPACT ANALYSIS

In this appendix, we detail the methods for and findings from our regression discontinuity design (RDD) impact analysis. We calculated the impact of implementing any of the four SIG-funded intervention models by estimating the local average treatment effect (LATE), sometimes called a "fuzzy" RDD. In Section A, we describe the benchmark approach to estimating main impacts. In Section B, we present the minimum detectable impacts for this analysis and additional impact findings not presented in Chapters IV and VI. In Section C, we present impacts for subgroups. In Section D, we present results from diagnostic analyses that we used to assess the validity of findings from the benchmark analysis. In Section E, we present results from sensitivity analyses to assess how findings would have changed if we had made different methodological choices. Finally, in Section F, we report findings from a variety of exploratory analyses.

A. Benchmark approach to estimating main impacts

Multiple alternative approaches are possible for several aspects of impact estimation (for example, selecting an RDD bandwidth). We base the main impact findings on a benchmark approach that consists of methods that we selected from among the available alternatives. In this section, we describe the sample used in the impact analysis, the impact regression equations, our approach to aggregating impacts into a single overall impact, and our approach to estimating standard errors.

1. Impact sample

We obtained the data for the RDD impact analysis from approximately 460 schools in approximately 50 school districts in 21 states. (One state and 20 schools were excluded from the RDD analysis because that state did not extend its IRB approval for the study while our analyses were being finalized.) We purposively selected the sample before outcome data were available, choosing states and school districts with a goal of maximizing the expected statistical precision of the RDD impacts. We based our estimates of statistical precision on the number of intervention and comparison schools in states and school districts and the fuzziness of the RDD. (We preferred districts and states in which a high proportion of schools in the intervention group received SIG funds to implement an intervention model and a low proportion of schools in the comparison group received SIG funds to implement an intervention model.) Using estimates of fuzziness and sample size based on information gathered through a review of states' SIG application materials and conversations with state administrative staff, we calculated the minimum detectable effect (MDE) corresponding to every opportunity to estimate an RDD impact in every state. (See Table A.1 for the full definitions of the SIG eligibility tiers and Table A.2 for details on the eight types of RDD opportunities created by those definitions, including the intervention and comparison groups and applicable RDD assignment variable for each opportunity.) We then ranked those opportunities and prioritized states and districts corresponding to the opportunities with lower MDE values. Because the sample was not randomly selected, one should use caution when interpreting the results. In particular, one should not assume that the findings presented in this report necessarily generalize to schools nationwide.

Table A.1. Eligibility requirements for implementing SIG-funded intervention models[a]

	Original Tier Definitions	Expanded Eligibility
Tier I	Any school receiving Title I funds for improvement, corrective action, or restructuring that falls into one of the following categories: • The lowest-achieving 5 percent of Title I schools in improvement, corrective action, or restructuring, or the lowest-achieving five Title I schools in improvement, corrective action, or restructuring in the state, whichever number of schools is greater • High schools that have had a graduation rate of less than 60 percent over a number of years	Title I-eligible elementary schools[b] that are no higher achieving than the highest achieving school that meets the original Tier I definition *and* fall into one of the following categories: • The bottom 20 percent of all schools in the state based on proficiency rates • Schools that have not made AYP for two consecutive years
Tier II	Any secondary school that is eligible for but does not receive Title I funds and that falls into one of the following categories: • The lowest-achieving 5 percent of secondary schools or the lowest-achieving five secondary schools in the state that are eligible for but do not receive Title I funds, whichever number of schools is greater • High schools with a graduation rate that is less than 60 percent over a number of years	Title I-eligible secondary schools[b] that are (1) no higher achieving than the highest-achieving school that meets the original Tier II definition or (2) high schools that have had a graduation rate of less than 60 percent over a number of years *and* fall into one of the following categories: • The bottom 20 percent of all schools in the state based on proficiency rates • Schools that have not made AYP for two consecutive years
Tier III	Schools receiving Title I funds for improvement, corrective action, or restructuring that are not in Tier I	Title I-eligible schools[b] that do not meet the requirements to be in Tier I or Tier II *and* that fall into the following categories: • The bottom 20 percent of all schools in the state based on proficiency rates • Schools that have not made AYP for two years

Source: U.S. Department of Education.

[a] The original tier definitions were published in the Federal Register on December 10, 2009. The expanded tier definitions were published in the Appropriations Act on December 16, 2009.

[b] Title I-eligible schools include all schools eligible to receive Title I funds, including both those that do and do not actually receive the funds.

AYP = adequate yearly progress.

Table A.2. Opportunities to conduct an RDD based on SIG eligibility tier definitions[a]

Opportunity	Intervention Group	Comparison Group	Assignment Variable
1	Original tier I elementary	Original tier III elementary	Achievement
2	Original tier I secondary	Original tier III secondary	Achievement
3	**Original tier I secondary**	**Original tier III secondary**	**Graduation rate**
4	Original tier II secondary	Original tier II secondary, but above the cutoff[b]	Achievement
5	**Original tier II secondary**	**Original tier II secondary, but above the cutoff[c]**	**Graduation rate**
6	Expanded tier I elementary	Expanded tier III elementary	Achievement
7	Expanded tier II secondary	Expanded tier III secondary	Achievement
8	**Expanded tier II secondary**	**Expanded tier III secondary**	**Graduation rate**

Source: State administrative records.

Note: We excluded from our impact analysis opportunities 3, 5, and 8 (indicated by red text) because our diagnostic analyses showed discontinuities in the density of the graduation rate assignment variable at the cutoff value.

[a] The original tiers were based on the definitions published in the Federal Register on December 10, 2009. The expanded tiers were based on the definitions published in the Appropriations Act on December 16, 2009.

[b] Schools that were eligible for but did not receive Title I funding and were above the 5 percent achievement cutoff.

[c] Schools that were eligible for but did not receive Title I funding and were above the 60 percent graduation rate cutoff.

RDD = regression discontinuity design.

For the benchmark analysis, we analyzed impacts using only the achievement assignment variable.[27] We used this approach because we found statistically significant discontinuities in the density of the graduation rate assignment variable at the cutoff (we describe this result in more detail in Section D).

The benchmark analysis included only the students who were present at the end of the school year, when standardized tests were administered. In other words, this impact estimate was place-based because it analyzed the students who were actually present in each place (school) at the time tests were administered, rather than students who were slated to attend each school. Therefore, the benchmark approach excluded schools that implemented the closure model by the end of the school year, as those schools were closed and did not contain any students at the end of the school year.

The study sample included a large number of students (roughly 15,000 to 300,000, depending on the outcome and year being examined). In Table A.3, we report student sample sizes in and out of the numerator and denominator bandwidths, above and below the cutoff, by year and by outcome. We also report sample sizes in terms of unique values of the assignment variable. Numerator and denominator bandwidths refer to the bandwidths selected for the numerator and denominator of the fuzzy RDD impact estimate, which were selected independently from each other (we present the equations for fuzzy impact estimation below). The numerator of the fuzzy RDD impact estimate is the magnitude of the discontinuity in the outcome regression at the cutoff; that is, the sharp RDD impact (Imbens and Wooldridge 2007). The denominator is the difference in SIG participation rates between schools below and above the cutoff (within the bandwidth); that is, the RDD impact on the likelihood of implementing a SIG-funded model.

[27] Per the SIG guidance provided by ED, states combined math and reading achievement into a single composite variable to determine SIG eligibility.

Table A.3. Assignment variable and bandwidth descriptions, place-based analysis

	Number of Students, Overall		Number of Unique Values of the Assignment Variable, Overall		Number of Students in Numerator Bandwidth		Number of Unique Values of the Assignment Variable in Numerator Bandwidth		Number of Students in Denominator Bandwidth		Number of Unique Values of the Assignment Variable in Denominator Bandwidth	
Outcome	**I**	**C**	**I**	**C**	**I**	**C**	**I**	**C**	**I**	**C**	**I**	**C**
Outcome year: 2010–2011												
Math test scores	94,890	199,540	210	640	64,260	73,190	160	220	67,050	78,610	170	240
Reading test scores	96,590	200,710	210	640	64,930	73,150	160	220	67,810	78,640	170	240
High school graduation	14,850	22,570	70	110	8,970	5,890	50	40	8,970	6,060	50	40
College enrollment	5,320	11,750	40	80	4,550	3,160	30	30	4,550	3,110	30	30
Outcome year: 2011–2012												
Math test scores	88,090	190,770	210	620	60,300	67,940	160	210	78,990	120,430	190	370
Reading test scores	89,190	189,180	210	610	61,180	68,210	160	210	77,500	112,100	190	350
High school graduation	12,960	20,620	70	120	9,570	8,340	50	60	8,500	6,840	50	50
College enrollment	3,910	10,670	40	80	3,470	3,000	30	30	3,290	2,640	30	30
Per-pupil spending[a]	170	140	170	140	120	90	110	100	100	70	90	80
Outcome year: 2012–2013												
Math test scores	88,740	189,990	200	600	68,700	87,540	170	260	74,750	111,080	180	330
Reading test scores	86,300	187,530	200	590	62,760	73,960	160	220	75,610	111,360	180	330
High school graduation	11,630	19,520	60	100	6,570	5,700	40	40	6,570	5,700	40	40
College enrollment	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Per-pupil spending[a]	160	140	160	140	120	90	110	90	90	70	80	70

Source: State and district administrative records; interviews with district administrators in spring 2012 and 2013.

Note: For each outcome, this table shows the number of students attending schools on either side (intervention or comparison) of the regression discontinuity design (RDD) cutoff, both overall and within the bandwidths. Bandwidths for the numerator and denominator of the fuzzy RDD impact estimate were selected independently from each other. The table also shows the number of unique values of the assignment variable on either side of the RDD cutoff, both overall and within the bandwidths. The number of unique values of the assignment variable is less than or equal to the number of schools because some schools could have the same value of the assignment variable. NA indicates cases for which we could not calculate impacts due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten.

[a] This row reports school sample sizes, rather than student sample sizes, because the outcome of per-pupil spending was a school-level outcome.
C = comparison; I = intervention.

We examined the characteristics of the intervention and comparison groups inside the bandwidth for two samples. Table A.4 reports characteristics of elementary schools that were assigned using the achievement variable, and Table A.5 reports characteristics of secondary schools assigned using the achievement variable.

Elementary schools in the intervention group had significantly lower reading achievement scores than schools in the comparison group (Table A.4). This finding was expected because schools were assigned to intervention and comparison groups using a cutoff on the achievement variable. Based on the Common Core of Data, intervention group schools also had (1) significantly more students eligible for free and reduced-price lunch, (2) significantly fewer students who were white, (3) significantly more students who were Hispanic, and (4) significantly more students who were Native American. Characteristics that did not differ significantly between the groups included percentage of students who were ELLs; per-pupil spending; total enrollment; percentage of schools that were receiving Title I funding schoolwide; and percentage of black, Asian, and multiracial students.

Table A.4. Characteristics of schools in RDD bandwidth for opportunities involving elementary schools

School Characteristic	Intervention Group	Comparison Group	*p*-Value
Characteristics from administrative data			
Average math achievement	-0.85	-0.79	0.20
Average reading achievement	-0.87	-0.76	0.02*
Percentage of students who are English language learners	21.4	16.4	0.11
Percentage of students eligible for free or reduced-price lunch[a]	89.3	88.0	0.36
Percentage of students who are white[a]	15.9	18.5	0.42
Characteristic from district survey			
Per-pupil spending in 2009–2010	$7,798	$8,102	0.72
CCD characteristics			
Total enrollment	446.9	444.4	0.92
Percentage of schools that receiving Title I funding schoolwide	94.7	95.6	0.75
Percentage of students eligible for free or reduced-price lunch[a]	88.7	82.5	0.00*
Percentage of students of the following race/ethnicity:			
White[a]	6.0	9.2	0.04*
Black	53.2	60.2	0.18
Hispanic	35.9	26.3	0.05*
Asian	1.8	2.0	0.80
Native American	1.4	0.5	0.05*
Multiracial	2.8	4.0	0.63
Number of Schools	**70–80**	**50–140**	

Source: Common Core of Data (CCD), state and district administrative records, and interviews with district administrators in spring 2012.

Note: The second and third columns show the mean characteristics of intervention and comparison group schools. Units for test scores are effect sizes (test scores were standardized to have a standard deviation of 1). We provide a range for the sample sizes because missing data varied across items in the table. In particular, the sample of schools for which we obtained per-pupil spending data from districts was smaller than the sample of schools within the RDD bandwidth because the school survey sample was based on the anticipated RDD bandwidth—which we estimated before collecting the administrative data—and the actual RDD bandwidth ended up being larger than the anticipated bandwidth. We aggregated administrative data to the school level, using the same set of students who were in the benchmark impact analysis for each school. The last column reports the *p*-value corresponding to the test of mean equivalence between the intervention and comparison groups. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten.

[a] The percentage of students eligible for free or reduced-price lunch and percentage of students who were white differed in the administrative data and the CCD because the set of grades differed. In particular, we did not collect administrative data for all grades from some schools, because those grades were not needed for the analysis.

*Significantly different at the 0.05 level, two-tailed test.

RDD = regression discontinuity design.

Secondary schools in the intervention group had significantly lower math and reading achievement scores than schools in the comparison group (Table A.5). Intervention group schools also had significantly larger total enrollments. In addition, a significantly higher proportion of intervention schools were receiving Title I funding schoolwide. Characteristics that did not differ significantly between the groups included percentage of students who were ELLs, per-pupil spending, percentage eligible for free and reduced-price lunch, and percentage of students in any race/ethnicity category (white, black, Hispanic, Asian, Native American, or multiracial).

Table A.5. Characteristics of schools in RDD bandwidth for opportunities involving secondary schools

School Characteristic	Intervention Group	Comparison Group	*p*-Value
Characteristics from administrative data			
Average math achievement	-0.72	-0.49	0.00*
Average reading achievement	-0.73	-0.52	0.00*
Percentage of students who are English language learners	12.8	10.1	0.14
Percentage of students eligible for free or reduced-price lunch[a]	80.5	81.0	0.79
Percentage of students who are white[a]	13.0	12.8	0.93
Characteristic from district survey			
Per-pupil spending in 2009–2010	$7,588	$8,061	0.47
CCD characteristics			
Total enrollment	1,002.5	742.8	0.00*
Percentage of schools that are receiving Title I funding schoolwide	99.1	92.0	0.01*
Percentage of students eligible for free or reduced-price lunch[a]	84.6	81.2	0.09
Percentage of students of the following race/ethnicity:			
White[a]	11.0	10.0	0.61
Black	54.5	63.0	0.08
Hispanic	27.1	21.2	0.15
Asian	2.4	2.1	0.53
Native American	4.0	2.9	0.59
Multiracial	1.1	1.0	0.97
Number of Schools	**100–120**	**80–130**	

Source: Common Core of Data (CCD), state and district administrative records, and interviews with district administrators in spring 2012.

Note: The second and third columns show the mean characteristics of intervention and comparison group schools. Units for test scores are effect sizes (test scores were standardized to have a standard deviation of 1). We provide a range for the sample sizes because missing data varied across items in the table. In particular, the sample of schools for which we obtained per-pupil spending data from districts was smaller than the sample of schools within the RDD bandwidth because the school survey sample was based on the anticipated RDD bandwidth—which we estimated before collecting the administrative data—and the actual RDD bandwidth ended up being larger than the anticipated bandwidth. We aggregated administrative data to the school level, using the same set of students who were in the benchmark impact analysis for each school. The last column reports the *p*-value corresponding to the test of mean equivalence between the intervention and comparison groups. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten.

[a] The percentage of students eligible for free or reduced-price lunch and percentage of students who were white differed in the administrative data and the CCD because the set of grades differed. In particular, we did not collect administrative data for all grades from some schools, because those grades were not needed for the analysis.

*Significantly different at the 0.05 level, two-tailed test.

RDD = regression discontinuity design.

2. Impact estimation by grade

We estimated impacts separately for each grade. The overall impact for the study consists of a weighted average of grade-specific impacts. We calculated impacts for all grades, 3 to 12, except in cases where the outcome was undefined (for example, many districts did not test students in grade 9, 11, or 12). For school-level outcomes (for example, SIG practice measures) we estimated a single impact.

3. Bandwidth selection

We estimated impacts using schools with values of the assignment variable that fell within a bandwidth around the cutoff values. To select the bandwidth, we first standardized the outcome and assignment variable from each grade using the statewide mean and standard deviation for each grade. We centered outcome variables at their means and the assignment variable at its cutoff value. We divided both by their standard deviations. We then pooled the standardized variables and applied the Imbens-Kalyanaraman (IK) procedure to the pooled data (Imbens and Kalyanaraman 2012). We estimated impacts locally in each grade within the single globally selected bandwidth (that is, the coefficient on the assignment variable varied across grade, but we used a single bandwidth studywide).

For all fuzzy RDD impact estimates reported, we selected separate bandwidths for the numerator and denominator of the impact estimate (we present the equations for fuzzy impact estimation below). For all reduced-form (that is, sharp) RDD impact estimates reported, we selected a single bandwidth for the impact estimate.

A challenge we faced in choosing an approach to bandwidth selection is that we estimated impacts separately for individual grades, but the impact of ultimate interest was a weighted average of the grade-specific impacts. Ideally, the bandwidth selection algorithm would minimize the mean square error for the aggregate impact, not for the individual grade impacts. However, we are not aware of a bandwidth selection algorithm that directly achieves this objective.

To address this challenge, we considered three approaches to bandwidth selection. Above, we describe the first approach—estimating a studywide bandwidth using pooled data. The second approach was to estimate a bandwidth for every grade. The third approach was a hybrid of the two methods, in which we calculated some components of the IK algorithm at the grade level and some studywide. Specifically, we calculated sample sizes and densities at the grade level; we estimated derivatives and conditional variances studywide using pooled data.[28]

We assessed the performance of these three bandwidth-selection approaches using Monte Carlo simulations. The simulations involved the following steps:

1. Using data from the year before SIG implementation, we estimated fourth-order polynomial regressions of the relationship between test score outcomes and assignment variable. We estimated these regressions separately by grade and, as data allowed, by state. We recorded the coefficients from these regressions and the regression R^2.
2. From among the sets of regression coefficients estimated in step 1, we selected the five that would yield the largest bias if we were to calculate an RDD impact for data generated using those sets of coefficients (using a cutoff value of 0). To contain the computational cost of the simulation, we focused on the five data-generating processes (DGPs) that yielded the largest bias. We focused on the ones that yielded the largest bias to be conservative with respect to bias in our choice of bandwidth selection algorithm. Table A.6 presents the coefficients for these five DGPs.
3. We generated 10,000 data sets. Each data set consisted of five grades. The DGPs for those five grades were those we selected in step 2.
4. For each of the 10,000 data sets, we estimated RDD impacts using the benchmark methods described in this appendix, but varying the bandwidth-selection approach. We saved the aggregate impact estimate from each of these 10,000 analyses.

Table A.6. Data-generating process coefficients

Data-Generating Process	b0	b1	b2	b3	b4	R^2
1	-0.474	0.057	0.006	0.025	-0.017	0.455
2	-0.552	-0.177	-0.327	0.549	-0.142	0.550
3	-0.787	0.332	0.361	-0.233	0.032	0.553
4	-0.768	0.066	0.021	0.177	-0.061	0.391
5	-0.795	0.459	-0.057	-0.169	0.030	0.358

Note: We used the coefficients in this table to generate an outcome Y from the equation $Y = ([R^2]^{0.5})*(b0 + b1*X + b2*X2 + b3*X3 + b4*X4) + ((1-R^2)^{0.5})*e$, where X and e are standard normal random variables.

Based on these simulation results, we decided to estimate a single, studywide IK bandwidth using pooled data. We made this decision after examining the mean and standard deviation of the impact estimates across the 10,000 replications (Table A.7). The examination had revealed that all three approaches yielded unbiased impact estimates, but the studywide approach yielded the most

[28] A member of the study's technical working group, Guido Imbens, suggested which components to estimate locally and globally.

precise impact estimate. (The mean of the impact estimates can be interpreted as bias because the data were generated under the null hypothesis of no impact.)

Table A.7. Comparing local, hybrid, and studywide IK procedures

Statistic	Bandwidth Estimation Approach		
	Local	Hybrid	Studywide
Mean impact	0.00	0.00	0.00
Impact standard deviation	0.70	0.62	0.59

Source: Monte Carlo simulations, 10,000 replications. Each replication consisted of 200 schools and five grades.

IK = Imbens-Kalyanaraman.

4. Impact equations

We calculated the impact of SIG-funded intervention model implementation using a fuzzy RDD by estimating the LATE.[29] The LATE equals the RDD impact on the outcome of interest divided by the RDD impact on the proportion of schools implementing a SIG-funded intervention model.

We estimated four equations using data where the unit of observation was the school, and schools were clustered within unique values of the assignment variable:

(1) $$Y_{ij}^{R} = \beta_0^{R} + \beta_1^{R} X_j^{R} + \beta_2^{R} Z_{ij}^{R} + u_j^{R} + \varepsilon_i^{R},$$

(2) $$Y_{ij}^{L} = \beta_0^{L} + \beta_1^{L} X_j^{L} + \beta_2^{L} Z_{ij}^{L} + u_j^{L} + \varepsilon_i^{L},$$

(3) $$P_{ij}^{R} = \gamma_0^{R} + \gamma_1^{R} X_j^{R} + \gamma_2^{R} Z_{ij}^{R} + \nu_j^{R} + e_i^{R},$$

(4) $$P_{ij}^{L} = \gamma_0^{L} + \gamma_1^{L} X_j^{L} + \gamma_2^{L} Z_{ij}^{L} + \nu_j^{L} + e_i^{L},$$

where the superscripts R and L denote the right and left sides of the RDD cutoff value for school-level achievement; Y is the average outcome (for example, scores on the state math assessment) for school i in cluster j, measured at the end of each school year; P is an indicator of whether a school implemented a SIG-funded intervention model; X is the assignment variable centered at the cutoff value; Z is a set of baseline covariates centered at the overall sample mean (that is, the mean calculated for the combined sample of units both above and below the cutoff) for those covariates (for example, demographic characteristics and prior test scores) and indicator variables for states, districts, and opportunity types (see Table A.2); u and v are cluster-level error terms; and ε and e are school-level error terms.[30] Schools to the left of (below) the cutoff were eligible for SIG

[29] To conduct robustness checks and diagnostic analyses (described below), we also estimated reduced-form, or "sharp," RDD impacts.

[30] Following Imbens and Lemieux (2008), we allowed the relationship between the outcome and the assignment variable, and the relationship between the outcome and the baseline covariates, to vary on either side of the cutoff.

funding. All equations were estimated using linear regression among observations within the bandwidth selected using the procedure described above.

The intercept in Equation (1) is the expected outcome for schools below the cutoff when all covariates equal zero; the intercept in Equation (2) is the expected outcome for schools above the cutoff when all covariates equal zero. Because the assignment variable is centered at the cutoff variable, and all other covariates are centered at their means, the intercept can be interpreted as the expected outcome at the cutoff value of the assignment variable and at the mean value of other covariates. The intercept in Equation (3) is the expected proportion of schools below the cutoff implementing a SIG-funded model when all covariates equal zero, and the intercept in Equation (4) is the expected proportion of schools above the cutoff implementing a SIG-funded intervention model when all covariates equal zero. Because the assignment variable is centered at the RDD cutoff value, the intercept is the predicted value of the dependent variable at the cutoff value. Therefore, the reduced-form RDD impact on Y is $\delta^Y = \beta_0^L - \beta_0^R$. The RDD impact on the likelihood of implementing a SIG-funded intervention model is $\delta^P = \gamma_0^L - \gamma_0^R$, and the LATE impact is $\delta^{LATE} = \dfrac{\delta^Y}{\delta^P}$.

5. Aggregating impacts across grades

We calculated the overall impact of SIG as a weighted average of impacts from all grades. The weight used in this calculation was based on sample size—specifically, the number of schools included in the IK bandwidth for each estimated impact. We interpret this overall impact as the average impact on schools at the assignment variable cutoff, which one can think about as the impact that would result from a marginal expansion in the availability of SIG funds by raising the cutoff used to determine eligibility. In other words, if the cutoff on the assignment variable is the policy lever used to expand or contract SIG, this overall impact is the average impact on the schools in our sample that would be affected by that lever.

6. Standard error estimation

To ensure that our standard errors fully reflected all sources of variability in our benchmark analysis—including variability introduced by the bandwidth selection algorithm—we calculated them using a residual bootstrap algorithm. The conceptual motivation for this algorithm is that RDD can be viewed as random assignment of residuals conditional on the assignment variable.

The residual bootstrap algorithm worked as follows:

1. Using all observations (not just those in the IK bandwidth), we regressed the outcome of interest on all variables used in the impact analysis plus the cutoff indicator and the square of the assignment variable and interactions between (a) the cutoff indicator and the assignment variable and (b) the cutoff indicator and square of the assignment variable. We also regressed the SIG participation variable on the same set of variables. We then saved the predicted outcome, predicted SIG participation probability, and residuals from those regressions. So that variance estimates accounted for the clustering of schools within unique values of the assignment variable (Lee and Card 2008), we created cluster- and individual-level predicted outcomes, participation probabilities, and residuals.

2. We randomly sampled pairs of outcome and SIG participation residuals (both at the cluster and individual level) to account for correlation between the outcome of interest and SIG participation. We constructed a bootstrap replicated data set by adding a randomly sampled residual to the predicted outcome for every observation (both at the cluster and individual levels). We did the same for predicted participation probabilities.
3. We calculated all grade-specific impacts (including bandwidth estimation) for the replicated data set.
4. We repeated steps 2 and 3 1,000 times. We calculated the covariance matrix of the grade-specific impacts using the impacts calculated in step 3 across bootstrap replications.

We used the covariance matrix resulting from this procedure to calculate the standard error for the overall aggregate impact. With this approach, we were able to account for (1) the bandwidth estimation algorithm, (2) covariance between impacts from different grades, (3) covariance between the numerator and denominator of the LATE impact estimate, and (4) clustering of schools within unique values of the assignment variable.

B. Additional impact estimates and minimum detectable effects

In this section, we report additional findings from the benchmark analysis that were not reported in Chapters IV and VI. We also report the MDEs for the benchmark impact analysis.

There was little evidence that SIG-funded models had an impact on the number of SIG-promoted practices used within any of the topic or subtopic areas (Table A.8). When examining practices separately for each of the four topic areas described in the SIG application, we found a significant negative impact of SIG-funded models for just one subtopic area in one year: tailoring strategies for secondary schools in spring 2012 (for which there was an impact of -2.82 practices).

Table A.8. Impacts of SIG-funded intervention models on the number of SIG-promoted practices used

Topic/Subtopic Area	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size
Outcome year: 2011–2012						
Total number of SIG-promoted practices used	21.76	21.47	0.28	0.41	0.89	320
Total number of ELL-focused SIG-promoted practices used	2.65	3.51	-0.87	-1.28	0.20	270
Comprehensive instructional reform strategies	7.09	7.50	-0.41	-0.57	0.35	320
Using data to identify and implement an instructional program	1.02	0.94	0.09	0.14	0.35	310
Promoting the continuous use of student data	1.93	1.95	-0.03	-0.04	0.74	310
Providing supports and professional development to staff to assist ELLs and students with disabilities	0.83	0.77	0.07	0.10	0.57	250
Using and integrating technology-based supports	0.79	0.94	-0.15	-0.23	0.09	310
Tailoring strategies for secondary schools	1.89	3.13	-1.23	-2.82*	0.00	100

Topic/Subtopic Area	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size
Teacher and principal effectiveness	10.03	9.77	0.26	0.38	0.84	320
Using rigorous, transparent, and equitable evaluation systems	1.56	2.27	-0.71	-0.89	0.17	310
Identifying and rewarding or removing teachers and principals	1.41	1.19	0.22	0.31	0.54	310
Providing high-quality, job-embedded professional development or supports	5.87	5.63	0.24	0.33	0.77	320
Implementing strategies to recruit, place, and retain staff	0.73	0.83	-0.10	-0.13	0.67	310
Learning time and community-oriented schools	3.54	3.57	-0.03	-0.04	0.95	320
Increasing learning time	0.55	0.48	0.07	0.10	0.55	300
Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs	3.01	3.02	-0.01	-0.02	0.98	320
Operational flexibility and support	0.94	0.94	0.01	0.01	0.97	310
Having operational flexibility	0.32	0.31	0.02	0.03	0.87	300
Receiving technical assistance and support	0.60	0.61	-0.01	-0.01	0.96	310
Outcome year: 2012–2013						
Total number of SIG-promoted practices used	21.28	19.03	2.25	3.26	0.21	330
Total number of ELL-focused SIG-promoted practices used	3.03	2.63	0.40	0.59	0.52	270
Comprehensive instructional reform strategies	7.32	6.85	0.47	0.79	0.10	330
Using data to identify and implement an instructional program	0.99	0.91	0.08	0.13	0.35	330
Promoting the continuous use of student data	1.93	1.91	0.01	0.02	0.75	330
Providing supports and professional development to staff to assist ELLs and students with disabilities	0.74	0.74	0.00	0.00	0.99	250
Using and integrating technology-based supports	0.97	0.86	0.11	0.18	0.15	320
Tailoring strategies for secondary schools	NA	NA	NA	NA	NA	NA
Teacher and principal effectiveness	9.99	8.89	1.10	1.60	0.43	330
Using rigorous, transparent, and equitable evaluation systems	1.89	1.69	0.20	0.33	0.57	320
Identifying and rewarding or removing teachers and principals	1.51	1.02	0.49	0.75	0.17	330
Providing high-quality, job-embedded professional development or supports	5.90	5.57	0.34	0.57	0.63	330
Implementing strategies to recruit, place, and retain staff	0.61	0.44	0.17	0.28	0.38	320
Learning time and community-oriented schools	3.20	2.93	0.26	0.44	0.56	330
Increasing learning time	0.60	0.43	0.17	0.27	0.08	310
Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs	2.47	2.49	-0.02	-0.03	0.97	330

Topic/Subtopic Area	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size
Operational flexibility and support	0.83	0.77	0.07	0.11	0.62	330
Having operational flexibility	0.24	0.21	0.03	0.05	0.71	310
Receiving technical assistance and support	0.61	0.54	0.07	0.11	0.49	330

Source: State and district administrative records; surveys of school administrators in spring 2012 and 2013.

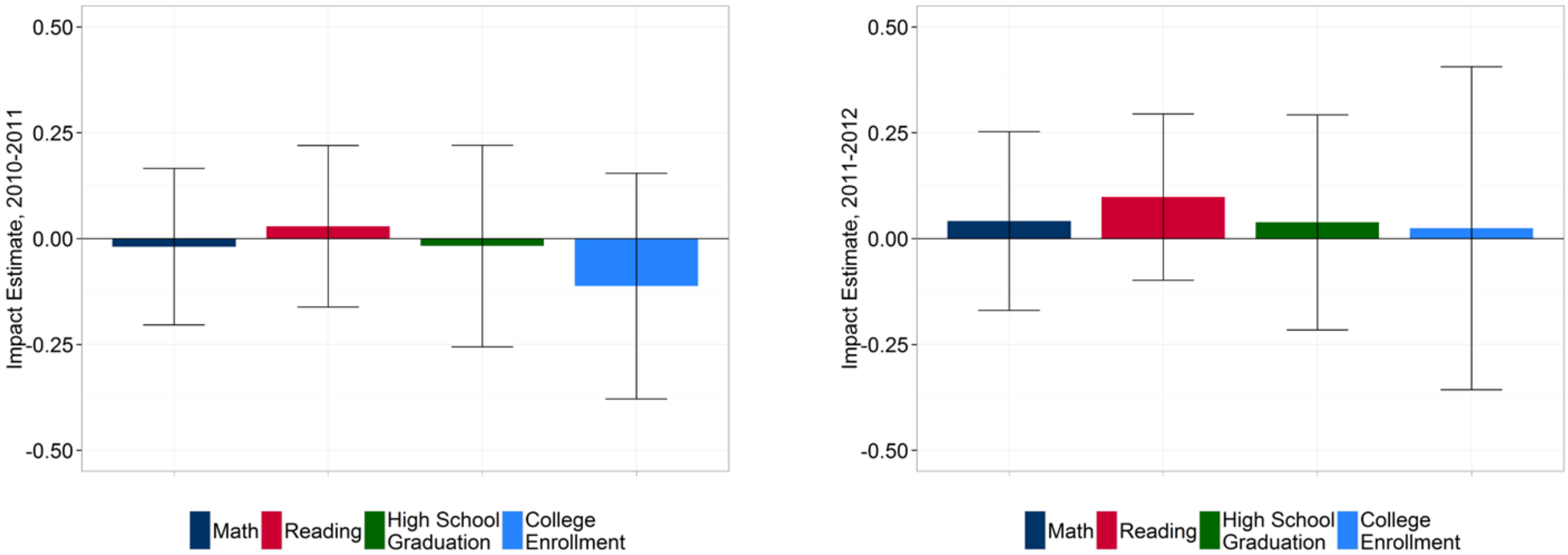
Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. Adjusted mean outcomes for the intervention and comparison groups are equal to the estimated intercept terms from the regressions of the outcome on the assignment variable and other covariates that were estimated separately on either side of the RDD cutoff value. The sample sizes reported in this table include both schools within the bandwidth and schools outside the bandwidth. The sample sizes differ across rows because missing data varied across items in the table. NA indicates cases for which we could not calculate impacts due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten.

*Significantly different from zero at the 0.05 level, two-tailed test.

ELL = English language learners.

We found no significant impacts of SIG on math test scores, reading test scores, high school graduation, or college enrollment for the 2010–2011 and 2011–2012 school years (see Figure A.1 and Table A.9).

Figure A.1. Impacts of SIG-funded models on student outcomes



Source: State and district administrative records.

Note: Units for test scores are effect sizes (test scores were standardized to have a standard deviation of 1). Units for high school graduation and college enrollment are percentage points/100. For example, an impact of 0.1 indicates an increase of 10 percentage points. Black bars show 95 percent confidence intervals. We calculated the results shown in these figures using the RDD methods described in this appendix.

For test score outcomes, we found that the smallest impacts that would be statistically significant in this study ranged from 0.19 to 0.22 standard deviations (Table A.10).[31] For high school graduation, the smallest impacts that would be statistically significant ranged from 0.15 to 0.26 standard deviations. For college enrollment, the smallest impacts that would be statistically significant were 0.27 and 0.39 standard deviations (in 2010–2011 and 2011–2012; impacts on college enrollment could not be calculated in 2012–2013 due to insufficient sample size). For per-pupil spending, the smallest impacts that would be statistically significant were $9,202 in 2011–2012 and $4,231 in 2012–2013. For the number of SIG-promoted practices used, the smallest impacts that would be statistically significant were 5.81 and 5.18 practices (in 2011–2012 and 2012–2013). For reference, we also included in this table our actual benchmark impacts for every outcome and year.

Table A.9. Impacts of SIG-funded intervention models on student outcomes and per-pupil spending, place-based analysis

Outcome	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
First outcome year: 2010–2011							
Math test scores	-0.67	-0.67	-0.01	-0.02	0.84	850	294,430
Reading test scores	-0.68	-0.71	0.03	0.03	0.76	840	297,300
High school graduation	0.72	0.73	-0.01	-0.02	0.89	160	37,420
College enrollment	0.44	0.51	-0.07	-0.11	0.41	90	17,070
Second outcome year: 2011–2012							
Math test scores	-0.63	-0.66	0.03	0.04	0.70	820	278,860
Reading test scores	-0.64	-0.71	0.07	0.10	0.33	810	278,370
High school graduation	0.78	0.75	0.03	0.04	0.77	160	33,570
College enrollment	0.46	0.45	0.01	0.02	0.90	90	14,580
Per-pupil spending	$8,025	$7,872	$153	$245	0.96	310	n.a.
Third outcome year: 2012–2013							
Math test scores	-0.64	-0.64	0.01	0.01	0.93	790	278,730
Reading test scores	-0.64	-0.69	0.05	0.08	0.46	790	273,830
High school graduation	0.76	0.79	-0.03	-0.05	0.53	150	31,150
College enrollment	NA	NA	NA	NA	NA	NA	NA
Per-pupil spending	$6,247	$6,187	$60	$100	0.96	300	n.a.

Source: State and district administrative records; interviews with district administrators and surveys of school administrators in spring 2012 and 2013.

Note: We standardized test score outcomes to have a standard deviation of 1, so we report test score impact estimates in effect-size units. Units for high school graduation and college enrollment are percentage points/100. For example, an impact of 0.1 indicates an increase of 10 percentage points. Units for per-pupil spending are dollars. Adjusted mean outcomes for the intervention and comparison groups are equal to the estimated intercept terms from the regressions of the outcome on the assignment variable and other covariates that were estimated separately on either side of the RDD cutoff value. The sample sizes reported in this table include schools/students both within the bandwidth and outside the bandwidth. NA indicates cases for which we could not calculate impacts due to insufficient sample sizes. To comply with

[31] We calculated these MDEs using our actual standard errors. The MDEs calculated using 50 percent power are the smallest impacts that could be statistically significant, given our standard errors.

NCES reporting requirements, we rounded sample sizes to the nearest ten. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

n.a. = not applicable; RDD = regression discontinuity design.

Table A.10. Minimum detectable effects

Outcome	Benchmark impact	MDE Using: 80 percent power	MDE Using: 50 percent power
Outcome year: 2010–2011			
Math test scores	-0.02	0.26	0.19
Reading test scores	0.03	0.27	0.19
High school graduation	-0.02	0.34	0.24
College enrollment	-0.11	0.38	0.27
Outcome year: 2011–2012			
Math test scores	0.04	0.30	0.22
Reading test scores	0.10	0.28	0.20
High school graduation	0.04	0.36	0.26
College enrollment	0.02	0.54	0.39
Per-pupil spending	$245	$12,883	$9,202
Number of SIG-promoted practices used	0.41	8.14	5.81
Outcome year: 2012–2013			
Math test scores	0.01	0.31	0.22
Reading test scores	0.08	0.29	0.20
High school graduation	-0.05	0.21	0.15
College enrollment	NA	NA	NA
Per-pupil spending	$100	$5,923	$4,231
Number of SIG-promoted practices used	3.26	7.25	5.18

Source: State and district administrative records; interviews with district administrators in spring 2012 and 2013.

Note: We calculated the MDE values in this table using the formula $MDE = \left[T^{-1}(1-0.05/2, df) + T^{-1}(\beta, df)\right] * \sqrt{Var\left(\hat{\delta}^{LATE}\right)}$, where T^{-1} is the inverse of the cumulative distribution function for student's t-distribution, β is the probability of a Type 2 error, df is degrees of freedom, and $Var\left(\hat{\delta}^{LATE}\right)$ is the estimated variance of the impact estimate. Except for per-pupil spending and the number of SIG-promoted practices used, we report all MDEs in effect-size units. NA indicates cases for which we could not calculate impacts due to insufficient sample sizes.

MDE = minimum detectable effect.

To interpret the magnitude of these MDEs, we drew on two of the benchmarks suggested by Hill (2008): (1) policy-relevant gaps in performance across different groups, and (2) effect sizes from past evaluations of other interventions targeting low-performing students and schools.[32] We believe these benchmarks are appropriate because SIG targets the lowest-performing schools with the goal of turning them around.

[32] Hill (2008) also proposes average achievement growth as a benchmark. We do not believe that benchmark is appropriate for this study because the schools and students in this study are far below average.

In terms of the first benchmark, our MDEs for test score outcomes are approximately one-fourth the size of the performance gaps that SIG schools are attempting to close. The gaps in performance between students in schools below the SIG eligibility cutoffs (that is, our intervention group) and the average performance of students in the same states are about 0.8 standard deviations (0.77 for math achievement and 0.78 for reading achievement, Table III.5).

With respect to the second benchmark, the MDEs for this study are within the range of impacts on academic achievement measures from past studies. In a review of meta-analyses, Hill (2008) finds mean effect sizes ranging from 0.22 to 0.27 standard deviations across grade levels. Among studies that meet What Works Clearinghouse (WWC) standards (with or without reservations), the 25th and 75th percentiles of impacts on math or reading achievement are 0.08 standard deviations and 0.41 standard deviations.[33]

C. Estimating subgroup impacts

We calculated impact estimates for the following subgroups: (1) ELLs and non-ELLs;[34] (2) grade level (grades 3 through 12); (3) elementary and secondary schools; (4) schools receiving Title I funds for improvement, corrective action, or restructuring, and secondary schools that were eligible for but did not receive Title I funds; and (5) schools in early RTT states, later RTT states, and non-RTT states.[35]

We estimated subgroup impacts such that all coefficients in Equations (1) to (4) were estimated separately for each subgroup. We used the same RDD bandwidth for subgroup impacts as we used for overall impacts. All other methods used for subgroup analyses were the same as the methods used for estimating main impacts.

All of the subgroup impacts on math and reading test scores, and high school graduation were statistically insignificant (Tables A.11, A.12, and A.13). It was not possible to estimate subgroup impacts on college enrollment due to small sample sizes and missing data. We also found no significant impacts involving the ELL and non-ELL subgroups (Table A.14).

---

[33] We calculated these percentiles using data available at http://ies.ed.gov/ncee/wwc/StudyFindings, restricting attention to impacts on reading or math achievement outcomes. We made our calculations for all impacts on reading and math achievement, not just those that were positive or statistically significant.

[34] ELL status is not fixed; students are expected to move out of this condition. Improvement in the achievement of ELL students should move them into the non-ELL category but may have no effect (or even a negative effect) on the achievement of the ELL group as a whole. We addressed this issue by defining students' ELL status in the benchmark analysis according to their ELL status prior to SIG (from the 2009–2010 school year) and leaving that status fixed afterward.

[35] Early RTT states received grants in the first two rounds of the competition. Later RTT states received grants in the third round. Non-RTT states did not receive grants. The RDD analysis sample included 7 of the 12 early RTT states, 4 of the 7 later RTT states, and 10 of the 32 non-RTT states, so the RTT versus non-RTT analysis is not fully representative of all the RTT and non-RTT states.

Table A.11. Impacts of SIG-funded intervention models on math test scores, by subgroup

Subgroup	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
Outcome year: 2010–2011							
Grade level							
Grade 3	-0.76	-0.77	0.01	0.01	0.96	490	26,060
Grade 4	-0.73	-0.80	0.07	0.08	0.54	490	26,440
Grade 5	-0.75	-0.66	-0.08	-0.11	0.44	500	26,640
Grade 6	-0.66	-0.64	-0.02	-0.02	0.90	550	44,860
Grade 7	-0.54	-0.63	0.09	0.12	0.57	460	48,390
Grade 8	-0.67	-0.63	-0.04	-0.06	0.81	430	44,190
Grade 9	-0.74	-0.75	0.01	0.01	0.98	170	24,820
Grade 10	-0.71	-0.64	-0.07	-0.19	0.64	210	20,940
Grade 11	-0.73	-0.72	-0.01	-0.02	0.95	180	27,890
Grade 12	-0.06	0.07	-0.12	-0.24	0.56	100	4,210
Elementary and secondary							
Elementary	-0.80	-0.80	-0.01	-0.01	0.93	490	113,420
Secondary	-0.57	-0.58	-0.01	-0.02	0.94	200	83,440
Difference				-0.01	0.98		
Title I status							
Title I receiving	-0.63	-0.57	0.06	0.16	0.49	110	55,960
Title I eligible	NA	NA	NA	NA	NA	NA	NA
Difference				NA	NA		
RTT status							
Early RTT states	-0.73	-0.83	-0.10	-0.19	0.35	370	78,080
Later RTT states	-0.87	-0.54	0.33	0.35	0.24	140	52,220
Non-RTT states	-0.68	-0.69	-0.01	-0.02	0.89	340	164,130
Difference between early RTT and non-RTT states				0.17	0.52		
Difference between later RTT and non-RTT states				-0.37	0.24		
Outcome year: 2011–2012							
Grade level							
Grade 3	-0.64	-0.76	0.12	0.16	0.34	480	25,200
Grade 4	-0.71	-0.69	-0.02	-0.03	0.84	480	24,860
Grade 5	-0.65	-0.78	0.13	0.19	0.30	490	24,930
Grade 6	-0.65	-0.67	0.02	0.03	0.90	530	42,470
Grade 7	-0.60	-0.64	0.03	0.05	0.83	450	45,250

Subgroup	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
Grade 8	-0.64	-0.58	-0.06	-0.08	0.73	400	42,640
Grade 9	-0.66	-0.66	-0.01	-0.01	0.98	150	21,950
Grade 10	-0.66	-0.69	0.03	0.05	0.90	200	21,100
Grade 11	-0.71	-0.73	0.01	0.03	0.94	160	25,780
Grade 12	-0.04	0.05	-0.09	-0.15	0.68	110	4,680
Elementary and secondary							
Elementary	-0.77	-0.65	0.13	0.14	0.31	470	108,010
Secondary	-0.53	-0.53	0.00	-0.01	0.97	200	79,920
Difference				-0.15	0.58		
Title I status							
Title I receiving	-0.51	-0.51	0.00	0.09	0.77	70	17,300
Title I eligible	NA	NA	NA	NA	NA	NA	NA
Difference				NA	NA		
RTT status							
Early RTT states	-0.69	-1.12	-0.44	-0.35	0.42	360	78,120
Later RTT states	-0.88	-0.43	0.45	0.71	0.14	140	49,060
Non-RTT states	-0.68	-0.67	0.01	0.01	0.96	330	151,680
Difference between early RTT and non-RTT states				0.36	0.48		
Difference between later RTT and non-RTT states				-0.70	0.15		
Outcome year: 2012–2013							
Grade level							
Grade 3	-0.56	-0.70	0.14	0.19	0.31	470	23,750
Grade 4	-0.68	-0.64	-0.03	-0.04	0.80	470	24,780
Grade 5	-0.68	-0.64	-0.05	-0.07	0.70	500	23,950
Grade 6	-0.57	-0.62	0.05	0.06	0.76	570	39,700
Grade 7	-0.69	-0.59	-0.11	-0.15	0.56	490	43,750
Grade 8	-0.58	-0.56	-0.01	-0.02	0.94	440	41,520
Grade 9	-0.81	-0.80	-0.01	-0.01	0.97	220	23,440
Grade 10	-0.70	-0.74	0.04	0.06	0.83	270	26,510
Grade 11	-0.68	-0.74	0.06	0.10	0.80	160	25,190
Grade 12	-0.37	-0.40	0.03	0.05	0.88	150	6,150
Elementary and secondary							
Elementary	-0.71	-0.60	0.11	0.11	0.46	460	104,800
Secondary	-0.65	-0.58	0.07	0.10	0.88	180	86,450
Difference				-0.02	0.99		

Subgroup	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
Title I status							
Title I receiving	-0.75	-0.81	-0.06	-0.12	0.70	110	55,340
Title I eligible	NA	NA	NA	NA	NA	NA	NA
Difference				NA	NA		
RTT status							
Early RTT states	-0.69	-0.83	-0.15	-0.31	0.21	350	75,550
Later RTT states	-0.67	-0.61	0.07	0.05	0.90	130	61,310
Non-RTT states	-0.61	-0.55	0.06	0.09	0.66	310	141,870
Difference between early RTT and non-RTT states				0.39	0.21		
Difference between Later RTT and non-RTT states				0.03	0.95		

Source: State and district administrative records.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. We standardized all outcomes to have a standard deviation of 1, so we report impact estimates in effect-size units. Adjusted mean outcomes for the intervention and comparison groups are equal to the estimated intercept terms from the regressions of the outcome on the assignment variable and other covariates that were estimated separately on either side of the RDD cutoff value. NA indicates cases for which we could not calculate impacts or *p*-values due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

Table A.12. Impacts of SIG-funded intervention models on reading test scores, by subgroup

Subgroup	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
Outcome year: 2010–2011							
Grade level							
Grade 3	-0.76	-0.81	0.05	0.07	0.65	490	25,750
Grade 4	-0.75	-0.85	0.10	0.13	0.39	490	26,060
Grade 5	-0.79	-0.85	0.06	0.07	0.62	500	26,260
Grade 6	-0.78	-0.77	-0.02	-0.02	0.91	540	44,430
Grade 7	-0.59	-0.73	0.14	0.18	0.40	460	47,900
Grade 8	-0.66	-0.65	-0.01	-0.02	0.95	420	43,490
Grade 9	-0.63	-0.54	-0.09	-0.13	0.68	160	25,570
Grade 10	-0.51	-0.44	-0.07	-0.14	0.73	200	25,160
Grade 11	-0.72	-0.65	-0.07	-0.13	0.68	170	28,160
Grade 12	0.15	0.05	0.10	0.12	0.76	80	4,520
Elementary and secondary							
Elementary	-0.86	-0.80	0.06	0.08	0.53	480	112,020
Secondary	-0.42	-0.46	-0.05	-0.09	0.67	200	89,030
Difference				-0.17	0.49		
Title I status							
Title I receiving	-0.39	-0.45	-0.07	-0.15	0.47	110	62,020
Title I eligible	NA	NA	NA	NA	NA	NA	NA
Difference				NA	NA		
RTT status							
Early RTT states	-0.84	-0.89	-0.05	-0.08	0.69	360	78,160
Later RTT states	-1.01	-0.78	0.22	0.20	0.55	140	53,780
Non-RTT states	-0.67	-0.65	0.03	0.02	0.89	340	165,360
Difference between early RTT and non-RTT states				0.10	0.74		
Difference between later RTT and non-RTT states				-0.18	0.63		
Outcome year: 2011–2012							
Grade level							
Grade 3	-0.65	-0.85	0.20	0.26	0.10	470	24,900
Grade 4	-0.71	-0.75	0.03	0.04	0.78	470	24,470
Grade 5	-0.61	-0.82	0.21	0.30	0.06	490	24,590
Grade 6	-0.68	-0.72	0.04	0.05	0.81	530	42,010
Grade 7	-0.73	-0.75	0.03	0.04	0.86	440	44,770

Subgroup	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
Grade 8	-0.66	-0.64	-0.02	-0.03	0.89	400	41,770
Grade 9	-0.66	-0.71	0.05	0.07	0.83	130	22,290
Grade 10	-0.54	-0.59	0.04	0.07	0.85	200	22,390
Grade 11	-0.58	-0.58	-0.01	-0.01	0.97	160	26,050
Grade 12	-0.08	-0.16	0.07	0.12	0.77	100	5,120
Elementary and secondary							
Elementary	-0.79	-0.66	0.14	0.17	0.21	470	106,610
Secondary	-0.52	-0.50	0.02	0.03	0.90	190	82,380
Difference				-0.14	0.59		
Title I status							
Title I receiving	-0.35	-0.49	-0.14	-0.30	0.29	110	37,480
Title I eligible	NA	NA	NA	NA	NA	NA	NA
Difference				NA	NA		
RTT status							
Early RTT states	-0.80	-0.93	-0.13	-0.15	0.51	350	76,110
Later RTT states	-0.80	-0.57	0.23	0.25	0.65	140	49,700
Non-RTT states	-0.66	-0.62	0.04	0.05	0.74	330	152,560
Difference between early RTT and non-RTT states				0.20	0.33		
Difference between later RTT and non-RTT states				-0.19	0.75		
Outcome year: 2012–2013							
Grade level							
Grade 3	-0.64	-0.77	0.13	0.17	0.34	470	23,510
Grade 4	-0.65	-0.72	0.07	0.10	0.57	470	24,400
Grade 5	-0.67	-0.66	0.00	0.00	0.98	490	23,600
Grade 6	-0.65	-0.69	0.04	0.05	0.80	570	39,290
Grade 7	-0.67	-0.72	0.05	0.06	0.79	480	43,340
Grade 8	-0.64	-0.68	0.03	0.05	0.83	430	40,490
Grade 9	-0.63	-0.66	0.03	0.04	0.91	140	18,140
Grade 10	-0.69	-0.67	-0.01	-0.02	0.95	260	29,070
Grade 11	-0.66	-0.68	0.02	0.03	0.95	150	25,330
Grade 12	-0.31	-0.57	0.25	0.44	0.28	140	6,680
Elementary and secondary							
Elementary	-0.75	-0.59	0.16	0.18	0.20	460	103,300
Secondary	-0.55	-0.50	0.06	0.07	0.88	180	84,760
Difference				-0.11	0.89		

Subgroup	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
Title I status							
Title I receiving	-0.72	-0.71	0.00	0.03	0.94	100	42,550
Title I eligible	NA	NA	NA	NA	NA	NA	NA
Difference				NA	NA		
RTT status							
Early RTT states	-0.79	-0.79	0.00	-0.05	0.82	350	75,650
Later RTT states	-1.05	-0.65	0.41	0.72	0.08	130	55,490
Non-RTT states	-0.64	-0.54	0.11	0.15	0.41	310	142,690
Difference between early RTT and non-RTT states				0.20	0.47		
Difference between later RTT and non-RTT states				-0.57	0.21		

Source: State and district administrative records.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. We standardized all outcomes to have a standard deviation of 1, so we report impact estimates in effect-size units. Adjusted mean outcomes for the intervention and comparison groups are equal to the estimated intercept terms from the regressions of the outcome on the assignment variable and other covariates that were estimated separately on either side of the RDD cutoff value. NA indicates cases for which we could not calculate impacts or p-values due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

Table A.13. Impacts of SIG-funded intervention models on high school graduation, by subgroup

Subgroup	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
Outcome year: 2010–2011							
Title I status							
Title I receiving	0.75	0.81	0.06	0.09	0.56	80	20,930
Title I eligible	NA	NA	NA	NA	NA	NA	NA
Difference				NA	NA		
RTT status							
Early RTT states	NA	NA	NA	NA	NA	NA	NA
Later RTT states	NA	NA	NA	NA	NA	NA	NA
Non-RTT states	0.72	0.73	0.01	0.02	0.91	100	30,340
Difference between early RTT and non-RTT states				NA	NA		
Difference between later RTT and non-RTT states				NA	NA		
Outcome year: 2011–2012							
Title I status							
Title I receiving	0.72	0.79	0.07	0.14	0.30	90	19,420
Title I eligible	NA	NA	NA	NA	NA	NA	NA
Difference				NA	NA		
RTT status							
Early RTT states	NA	NA	NA	NA	NA	NA	NA
Later RTT states	NA	NA	NA	NA	NA	NA	NA
Non-RTT states	0.78	0.80	0.02	0.02	0.87	90	26,850
Difference between early RTT and non-RTT states				NA	NA		
Difference between later RTT and non-RTT states				NA	NA		
Outcome year: 2012–2013							
Title I status							
Title I receiving	0.75	0.79	0.03	0.05	0.61	70	17,730
Title I eligible	NA	NA	NA	NA	NA	NA	NA
Difference				NA	NA		

Subgroup	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
RTT status							
Early RTT states	NA	NA	NA	NA	NA	NA	NA
Later RTT states	NA	NA	NA	NA	NA	NA	NA
Non-RTT states	0.76	0.76	0.00	0.00	0.96	80	23,370
Difference between early RTT and non-RTT states				NA	NA		
Difference between later RTT and non-RTT states				NA	NA		

Source: State and district administrative records.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. We standardized all outcomes to have a standard deviation of 1, so we report impact estimates in effect-size units. Adjusted mean outcomes for the intervention and comparison groups are equal to the estimated intercept terms from the regressions of the outcome on the assignment variable and other covariates that were estimated separately on either side of the RDD cutoff value. NA indicates cases for which we could not calculate impacts or p-values due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

Table A.14. Impacts of SIG-funded intervention models, by ELL subgroup

Outcome/Subgroup	Intervention Mean	Comparison Mean	Impact Estimate	*p*-Value
Outcome year: 2010–2011				
Math test scores for ELL students	-0.90	-0.87	0.02	0.90
Math test scores for non-ELL students	-0.63	-0.61	0.02	0.89
Difference			0.00	0.98
Reading test scores for ELL students	-1.18	-1.10	0.09	0.58
Reading test scores for non-ELL students	-0.63	-0.59	0.04	0.74
Difference			-0.05	0.79
High school graduation for ELL students	0.46	0.58	0.15	0.54
High school graduation for non-ELL students	0.76	0.73	-0.04	0.81
Difference			-0.19	0.42
College enrollment for ELL students	0.35	0.16	-0.23	0.54
College enrollment for non-ELL students	0.52	0.46	-0.10	0.70
Difference			0.14	0.74
Outcome year: 2011–2012				
Math test scores for ELL students	-0.95	-0.90	0.07	0.68
Math test scores for non-ELL students	-0.64	-0.57	0.11	0.38
Difference			0.04	0.83
Reading test scores for ELL students	-1.16	-1.12	0.06	0.75
Reading test scores for non-ELL students	-0.64	-0.54	0.14	0.22
Difference			0.09	0.68
High school graduation for ELL students	0.71	0.71	0.00	0.99
High school graduation for non-ELL students	0.81	0.77	-0.05	0.76
Difference			-0.05	0.83
College enrollment for ELL students	0.43	0.43	-0.01	1.00
College enrollment for non-ELL students	0.43	0.47	0.09	0.72
Difference			0.10	0.96
Outcome year: 2012–2013				
Math test scores for ELL students	-0.97	-0.96	-0.01	0.96
Math test scores for non-ELL students	-0.59	-0.58	0.02	0.85
Difference			0.03	0.88
Reading test scores for ELL students	-1.19	-1.15	0.04	0.81
Reading test scores for non-ELL students	-0.60	-0.53	0.09	0.51
Difference			0.04	0.84
High school graduation for ELL students	0.66	0.44	-0.29	0.16
High school graduation for non-ELL students	0.79	0.79	-0.01	0.96
Difference			0.28	0.18
College enrollment for ELL students	NA	NA	NA	NA
College enrollment for non-ELL students	0.36	0.44	0.10	0.91
Difference			NA	NA

Source: State and district administrative records.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. We standardized test score outcomes to have a standard deviation of 1, so we report test score impact estimates in effect-size units. Units for high school graduation and college enrollment are percentage points/100. For example, an impact of 0.1 indicates an increase of 10 percentage points. Adjusted mean outcomes for the intervention and comparison groups are equal to the estimated intercept terms from the regressions of the outcome on the assignment variable and other covariates that were estimated separately on either side of the RDD cutoff value. NA indicates cases for which we could not calculate *p*-values due to insufficient sample sizes. No impacts were statistically significant at the 0.05 percent level, using a two-tailed test.

ELL = English language learner.

D. Diagnostic analyses

For each benchmark impact estimate, we conducted diagnostic analyses focused on the following issues: (1) integrity of the assignment variable, (2) attrition from the study sample, (3) continuity of the relationship between the assignment variable and the outcome, (4) graphical analysis to assess the appropriateness of the selected bandwidth, and (5) finite sample bias.

1. Integrity of the assignment variable

A key condition for an RDD to produce consistent impact estimates is that there was no systematic manipulation of the assignment variable. Manipulation means altering the true values of the assignment variable to influence intervention assignments—for example, falsifying the graduation rate for a high school to affect the school's SIG eligibility.

We assessed the integrity of the assignment variables in this study using three approaches: (1) an institutional approach in which we examined the process for calculating and reporting values of the assignment variable, (2) a statistical test for such discontinuities (the McCrary [2008] test), and (3) a graphical approach in which we looked for discontinuities in the density of the assignment variable on either side of the cutoff.

Institutional assessment. We carefully reviewed the process each state used to categorize schools into SIG eligibility tiers and award grants. This review consisted of three components:

1. **An examination of submitted SIG applications**. Each state submitted a detailed application for SIG in which they were required to list SIG eligible schools by tier and to explain the process they used to categorize schools into tiers. This information served as the starting point for understanding the process of categorizing schools into eligibility tiers and awarding grants.
2. **Detailed conversations with state staff**. We contacted state staff to further refine our understanding of how schools were categorized into eligibility tiers and how grants were awarded. From these conversations, we learned which continuous variables were used to divide schools into Tiers I, II, or III.
3. **A school-level analysis**. We obtained the school-level data (including the continuous achievement and graduation rate variables that we planned to use as RDD assignment variables) that states used to award grants. This data enabled us to examine whether schools were divided into tiers and awarded grants in the manner described by state staff.

Our review indicates that it is unlikely that assignment variables in this study were manipulated, because (1) the variables (school-level achievement and graduation rates) existed in administrative records prior to the SIG program (and, therefore, it would have been difficult to revise those data); (2) the staff assessing which schools were eligible for SIG were at the state rather than district level, which means they were less likely to have personal preferences regarding which schools were deemed eligible for SIG; and (3) ED, not the states, set the cutoffs on these variables (the 5th percentile of achievement and a 60 percent graduation rate).

Statistical test. We did not find any discontinuities involving the achievement assignment variable, but we found three statistically significant results involving the graduation rate (Table A.15). We used Dimmery (2013) to carry out the McCrary (2008) test. Due to these discontinuities, we decided to exclude from our benchmark analysis all impacts estimated using the graduation rate (we conducted a sensitivity analysis that includes impacts using the graduation rate, however).

Graphical assessment. We found no graphical evidence of density discontinuities. To make this assessment, we created and examined histograms (represented by the dots in each figure) and densities (represented by the curve) of the achievement assignment variable using Dimmery (2013) (Figures A.2 to A.24). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines). In each figure, the confidence intervals at the cutoff overlap each other, indicating no graphical evidence of density discontinuities. (Due to the large number of figures in this appendix, we include Figures A.2 to A.93 at the end of the appendix.)

2. Sample attrition

Sample attrition could contribute bias to impact estimates if it resulted in a lack of equivalence between intervention and comparison units at the cutoff value of the assignment variable. We focus on the attrition rate at the cutoff value of the assignment variable because RDD impacts apply to the cutoff value of the assignment variable. (Conceptually, RDD impacts on outcomes, and RDD estimates of attrition, apply to schools at the cutoff. In practice, it is not possible to estimate impacts or attrition using only schools at the cutoff because there are too few schools at the cutoff. Instead, we estimate impacts and attrition at the cutoff using regression lines based on data within the bandwidth both above and below the cutoff.)

Due to this potential source of bias, studies can meet WWC evidence standards without reservations only if they have low levels of overall and differential (between the intervention and comparison groups) attrition (U.S. Department of Education 2010b, 2015). To assess the potential for attrition bias in our benchmark analysis, we calculated RDD impacts on school-level attrition and reported the mean predicted school-level attrition rates for the intervention and comparison groups at the cutoff values of the assignment variables. We used the same analytic approach to calculating impacts on attrition that we used for our main impact analysis (including the same bandwidth selection algorithm). Because our benchmark analysis was a place-based analysis (that is, it was an impact focused on the students present in study schools at follow-up that combined impacts on student achievement with impacts on mobility), we did not report student-level attrition rates.

For three out of four student outcomes, we found very little attrition, either overall or differentially between the intervention and comparison groups (Table A.16). For math and reading test score outcomes, there was almost no attrition. For high school graduation, there was almost no attrition in the first two years, but some (about 5 percent overall, 2 percent differential between the intervention and comparison groups) in the third year. The attrition rate was high for college enrollment, particularly in the third year. In addition, the overall attrition rate for per-pupil spending, another outcome we examined, was not high, but the differential rate was high in the second year (about 9 percentage points).

Table A.15. McCrary test results, place-based analysis

	McCrary Test *p*-Value, by Outcome and Year											
	Math test scores			Reading test scores			High school graduation			College enrollment		
Grade and assignment variable	2010–11	2011–12	2012–13	2010–11	2011–12	2012–13	2010–11	2011–12	2012–13	2010–11	2011–12	2012–13
Grade 3, achievement	0.92	0.69	0.75	0.92	0.73	0.71	NA	NA	NA	NA	NA	NA
Grade 4, achievement	0.88	0.69	0.70	0.87	0.73	0.70	NA	NA	NA	NA	NA	NA
Grade 5, achievement	0.90	0.31	0.65	0.93	0.33	0.60	NA	NA	NA	NA	NA	NA
Grade 6, achievement	0.89	0.23	0.76	0.94	0.23	0.78	NA	NA	NA	NA	NA	NA
Grade 7, achievement	0.69	0.68	0.89	0.71	0.71	0.96	NA	NA	NA	NA	NA	NA
Grade 8, achievement	0.98	0.55	0.69	0.83	0.57	0.71	NA	NA	NA	NA	NA	NA
Grade 9, both, 1[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 9, both, 2[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 9, achievement	0.71	0.39	0.88	0.76	0.31	0.73	NA	NA	NA	NA	NA	NA
Grade 9, grad rate	NA	0.76	0.89	0.55	0.70	0.68	NA	NA	NA	NA	NA	NA
Grade 10, both, 1[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 10, both, 2[a]	NA	NA	NA	NA	NA	0.59	NA	NA	NA	NA	NA	NA
Grade 10, achievement	0.85	0.59	0.86	0.89	0.78	0.81	NA	NA	NA	NA	NA	NA
Grade 10, grad rate	0.28	0.16	0.09	0.13	0.20	0.20	NA	NA	NA	NA	NA	NA
Grade 11, both, 1[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 11, both, 2[a]	NA	0.64	NA	NA	NA	0.59	NA	NA	NA	NA	NA	NA
Grade 11, achievement	0.91	0.57	0.35	1.00	0.89	0.37	NA	NA	NA	NA	NA	NA
Grade 11, grad rate	0.07	0.06	0.03*	0.10	0.04*	0.08	NA	NA	NA	NA	NA	NA
Grade 12, both, 1[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 12, both, 2[a]	NA	NA	NA	NA	NA	0.59	NA	NA	NA	NA	NA	NA
Grade 12, achievement	0.83	0.74	0.53	0.55	0.25	0.62	0.90	0.99	0.63	0.59	0.62	NA
Grade 12, grad rate	0.17	0.22	0.05[b]	0.27	NA	0.07	0.28	0.15	0.03*	0.23	0.15	0.35

Source: State and district administrative records.

Note: This table reports results of the McCrary test (2008) for discontinuities in the density of the assignment variable at the cutoff. We report McCrary *p*-values by outcome, year, grade, and assignment variable. NA indicates cases for which we could not calculate impacts due to insufficient sample sizes, and so we did not conduct the McCrary test.

[a] Both, 1 indicates a mini-study with two assignment variables, with this row reporting on the achievement assignment variable. Both, 2 indicates a mini-study with two assignment variables, with this row reporting on the graduation rate assignment variable.

[b] Rounds to—but is larger than—0.05.

*Statistically significant at the 0.05 level, two-tailed test.

Table A.16. Impacts of SIG-funded intervention models on attrition

Outcome	Intervention Mean	Comparison Mean	Impact Estimate	*p*-Value
Outcome year: 2010–2011				
Math test scores	0.00	0.01	0.00	1.00
Reading test scores	0.01	0.01	0.00	1.00
High school graduation	0.01	0.01	0.00	0.98
College enrollment	0.07	0.02	0.05	0.87
Outcome year: 2011–2012				
Math test scores	0.00	0.01	0.00	1.00
Reading test scores	0.01	0.01	0.00	1.00
High school graduation	0.01	0.01	0.00	0.99
College enrollment	0.09	0.53	-0.45	0.10
Per-pupil spending[a]	-0.01	0.08	-0.09	0.50
Outcome year: 2012–2013				
Math test scores	0.01	0.01	0.00	0.84
Reading test scores	0.01	0.00	0.00	0.87
High school graduation	0.04	0.05	-0.02	0.94
College enrollment	0.13	0.79	-0.66	0.08
Per-pupil spending[a]	0.07	0.03	0.04	0.77

Source: State and district administrative records; interviews with district administrators in spring 2012 and 2013.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. Adjusted mean outcomes for the intervention and comparison groups were equal to the estimated intercept terms from the regressions of the outcome on the assignment variable (student achievement) and other covariates that were estimated separately on either side of the RDD cutoff value. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

[a] The outcome of per-pupil spending came from the district interview. School-level attrition could occur because a school closed or because the district did not provide data for that school. For the outcome of per-pupil spending, most of the attrition was due to closures.

3. Continuity of the relationship between the assignment variable and the outcome

In the absence of the intervention, there should be a smooth relationship between the outcome and the assignment variable at the cutoff for an RDD to be valid. If this condition is satisfied, then a discontinuity in the outcome at the cutoff can be attributed to the intervention.

It is never possible to directly assess the smoothness condition under the counterfactual of no intervention, but we were able to make three types of indirect assessments. First, we calculated RDD impacts on baseline covariates related to the outcome variable. The presence of such impacts would strongly suggest a violation of the smoothness condition because the intervention cannot affect baseline covariates and baseline measures of the outcomes. Second, we estimated impacts at multiple values of the assignment variable other than the actual cutoff. Third, we used graphical analysis to look for unexplained discontinuities in the relationship between the outcome and the assignment variable at values of the assignment variable other than the cutoff value. Unexplained violations of the smoothness condition away from the cutoff suggest a higher probability that the smoothness condition is violated at the cutoff.

When we estimated RDD impacts on baseline covariates, we found no statistically significant differences with respect to any covariate, which is consistent with the smoothness condition being satisfied (Table A.17). When we estimated RDD impacts on baseline math and reading test scores for each analytic sample we examined (recall that we estimated impacts on math test scores, reading test scores, high school graduation, and college enrollment in 2010–2011, 2011–2012, and 2012–2013), we also found no statistically significant impacts (Tables A.18 and A.19).

Findings from our analysis of discontinuities in the relationship between the outcome and the assignment variable at values other than the cutoff were consistent with the smoothness condition being satisfied. For each grade that contributed an impact to the overall aggregate impact, we calculated outcome discontinuities at 10 or more values of the assignment variable other than the actual cutoff. (Table A.20 reports the number of significant discontinuities by outcome and by year.) These values were all within the IK bandwidth. Half of the values were above and half were below the actual cutoff. The values were evenly spaced across the distribution of the assignment variable inside the bandwidth. The proportion of discontinuities that were statistically significant was smaller than 5 percent (the percent that would be expected to be significant due to chance) for every student outcome (that is, math and reading test scores, high school graduation, and college enrollment) and year, which is consistent with the smoothness condition being satisfied.

Table A.17. Comparison of baseline characteristics between the intervention and comparison groups for the benchmark impacts

Baseline Characteristic	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size
Characteristics from administrative data						
Average math achievement	-0.78	-0.72	-0.06	-0.08	0.57	890
Average reading achievement	-0.72	-0.77	0.05	0.07	0.55	890
Percentage of students who are						
White	0.16	0.12	0.04	0.05	0.47	900
Eligible for free or reduced-price lunch	0.77	0.80	-0.03	-0.04	0.39	900
English language learners	0.15	0.18	-0.03	-0.03	0.53	900
Characteristic from district survey						
Per-pupil spending in 2009–2010	$7,374	$7,374	$0	$0	1.00	330
CCD characteristics						
Total enrollment	933.16	1,001.61	-68.45	-91.04	0.56	930
Percentage of schools that are receiving Title I funding schoolwide	0.97	0.96	0.01	0.02	0.85	860
Percentage of students who are						
White	0.10	0.07	0.03	0.04	0.47	920
Black	0.48	0.51	-0.03	-0.05	0.63	920
Hispanic	0.04	0.02	0.01	0.02	0.19	920
Asian	0.04	0.02	0.01	0.02	0.19	920
Native American	0.01	0.01	0.00	0.00	0.93	920
Multiracial	0.04	0.03	0.01	0.01	0.80	140
Eligible for free or reduced-price lunch	0.81	0.85	-0.04	-0.06	0.29	920

Source: State and district administrative records; surveys of school administrators in spring 2012; Common Core of Data (CCD).

Note: We calculated intervention-comparison group differences using the benchmark approach to calculating regression discontinuity design (RDD) impact estimates described in this appendix, except that no additional covariates were included in the model. We standardized test score outcomes to have a standard deviation of 1, so we report test score impact estimates in effect-size units. Units for per-pupil spending are dollars. The sample sizes reported in this table include both schools within the bandwidth and schools outside the bandwidth. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

Table A.18. Comparison of baseline math achievement for the intervention and comparison groups used to estimate impacts on student outcomes, place-based analysis

Baseline Math Achievement for Samples Used to Estimate Impacts on the Following Outcomes:[a]	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
First outcome year: 2010–2011							
Math test scores	-0.66	-0.64	-0.02	-0.06	0.48	760	249,460
Reading test scores	-0.65	-0.64	-0.01	-0.04	0.61	750	252,180
High school graduation	-1.13	-0.83	-0.30	-0.63	0.10	120	32,740
College enrollment	NA	NA	NA	NA	NA	NA	NA
Second outcome year: 2011–2012							
Math test scores	-0.62	-0.62	-0.01	-0.01	0.94	750	216,500
Reading test scores	-0.62	-0.61	-0.01	-0.02	0.89	740	213,690
High school graduation	-0.75	-0.71	-0.05	-0.06	0.73	90	23,830
College enrollment	NA	NA	NA	NA	NA	NA	NA
Per-pupil spending	NA	NA	NA	NA	NA	NA	NA
Third outcome year: 2012–2013							
Math test scores	-0.59	-0.63	0.04	0.10	0.57	680	201,300
Reading test scores	-0.61	-0.63	0.02	0.04	0.82	680	197,170
High school graduation	-0.81	-0.69	-0.12	-0.16	0.61	110	25,910
College enrollment	NA	NA	NA	NA	NA	NA	NA
Per-pupil spending	NA	NA	NA	NA	NA	NA	NA

Source: State and district administrative records; surveys of school administrators in spring 2012; Common Core of Data (CCD).

Note: We calculated intervention-comparison group differences using the benchmark approach to calculating regression discontinuity design (RDD) impact estimates described in this appendix, except that no additional covariates were included in the model. In particular, for each outcome and year, we used the same bandwidth to calculate intervention-comparison group differences as we used to estimate RDD impacts for the relevant outcome and year. We standardized test score outcomes to have a standard deviation of 1, so we report impact estimates in effect-size units. NA indicates cases for which we could not calculate impacts or *p*-values due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

[a] The outcome in this table is baseline math achievement. The first column indicates which sample was used to examine baseline math achievement. For example, the second row examines baseline math achievement for the sample of schools that were used in the place-based analysis of reading test scores in 2010–2011.

Table A.19. Comparison of baseline reading achievement for the intervention and comparison groups used to estimate impacts on student outcomes, place-based analysis

Baseline Reading Achievement for Samples Used to Estimate Impacts on the Following Outcomes:[a]	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
First outcome year: 2010–2011							
Math test scores	-0.69	-0.66	-0.03	-0.06	0.56	760	249,000
Reading test scores	-0.68	-0.68	-0.01	-0.03	0.77	750	252,110
High school graduation	-0.95	-0.94	-0.02	-0.04	0.94	120	32,500
College enrollment	NA	NA	NA	NA	NA	NA	NA
Second outcome year: 2011–2012							
Math test scores	-0.66	-0.64	-0.02	-0.02	0.90	740	216,480
Reading test scores	-0.64	-0.64	0.01	0.02	0.90	740	213,690
High school graduation	-0.49	-0.71	0.22	0.31	0.14	90	23,810
College enrollment	NA	NA	NA	NA	NA	NA	NA
Per-pupil spending	NA	NA	NA	NA	NA	NA	NA
Third outcome year: 2012–2013							
Math test scores	-0.66	-0.66	0.00	0.03	0.85	680	200,810
Reading test scores	-0.67	-0.68	0.01	0.04	0.83	680	196,630
High school graduation	-0.71	-0.45	-0.26	-0.34	0.33	110	25,910
College enrollment	NA	NA	NA	NA	NA	NA	NA
Per-pupil spending	NA	NA	NA	NA	NA	NA	NA

Source: State and district administrative records; surveys of school administrators in spring 2012; Common Core of Data (CCD).

Note: We calculated intervention-comparison group differences using the benchmark approach to calculating regression discontinuity design (RDD) impact estimates described in this appendix, except that no additional covariates were included in the model. In particular, for each outcome and year, we used the same bandwidth to calculate intervention-comparison group differences as we used to estimate RDD impacts for the relevant outcome and year. We standardized all outcomes to have a standard deviation of 1, so we report impact estimates in effect-size units. NA indicates cases where we could not calculate impacts or *p*-values due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

[a] The outcome in this table is baseline reading achievement. The first column indicates which sample was used to examine baseline reading achievement. For example, the first row examines baseline reading achievement for the sample of schools that were used in the place-based analysis of math test scores in 2010–2011.

Table A.20. Assessing the continuity of the relationship between the outcome and assignment variable for student outcomes and per-pupil spending, place-based analysis

Outcome	Number of Assignment Variable Values Other than the Cutoff for Which We Tested for Discontinuities in the Outcome	Percentage of Assignment Variable Values with Statistically Significant Discontinuities in the Outcome
Outcome year: 2010–2011		
Math test scores	100	1
Reading test scores	100	1
High school graduation	10	0
College enrollment	10	0
Outcome year: 2011–2012		
Math test scores	100	3
Reading test scores	100	4
High school graduation	10	0
College enrollment	10	0
Per-pupil spending	10	10
Outcome year: 2012–2013		
Math test scores	100	2
Reading test scores	100	1
High school graduation	10	0
College enrollment	NA	NA
Per-pupil spending	10	10

Source: State and district administrative records; interviews with district administrators in spring 2012 and 2013.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. For each outcome and grade, we tested for discontinuities in the outcome-assignment variable relationship at several values of the assignment variable other than the actual cutoff (that is, at values of the assignment variable where no discontinuities are expected). Specifically, we tested for discontinuities at four values above and four values below the actual cutoff. These values were all inside the IK bandwidth used for the numerator of the fuzzy impact estimates. The values were evenly spaced across the distribution of the assignment variable inside the bandwidth. NA indicates cases for which we could not conduct tests due to insufficient sample sizes.

IK = Imbens and Kalyanaraman.

Graphically, we found no clear evidence of discontinuities in the relationship between the outcome and the assignment variable in the bandwidth at points other than the cutoff. Figures A.25 to A.47 show the relationship between the assignment variable and the outcome for every grade that contributed an impact to the overall aggregate impact. In these figures, we present the raw data, a best-fit curve, and the 95 percent confidence interval for the best-fit curve. The cutoff value in each plot is zero (indicated by a red vertical line). The grey shaded region around the cutoff value is the bandwidth. As a gauge to assess the magnitude of any apparent discontinuities, we included two sets of horizontal black lines. The distance between the two solid black lines equals two times the standard error of the grade-specific reduced-form impact estimate. The distance between the two dashed lines equals two times the standard error of the grade-specific fuzzy impact estimate. We used two times the standard error to comply with the WWC standards for RDD studies. The best-fit curve was estimated using local linear regression as implemented by the loess function in R (R Core Team 2015). The green shaded region around the best-fit curve shows the 95 percent confidence interval for the curve. We concluded that none of these figures provided clear evidence of discontinuities in the relationship between the outcome and the assignment variable in the bandwidth at points other than the cutoff because we saw no discontinuities larger than the distance between either set of two black lines. To assess discontinuities, we looked for sharp changes in the slope of the best-fit curve and for obvious discontinuities in the vertical location of points in the scatter plot.

From these analyses, we concluded that the relationship between the assignment variable and the outcome was likely smooth in the absence of the intervention.

4. Graphical analysis

We can also use the graphical analysis presented in Figures A.25 to A.47 to assess the appropriateness of the selected bandwidth. The bandwidth is appropriate if the relationship between the outcome and assignment variable inside the bandwidth appears to be approximately linear.

Based on a review of these figures, we found little evidence to suggest that the selected bandwidth was inappropriate. In most cases, the relationship appeared to be reasonably linear (perfect linearity is not required—the bandwidth selection algorithm chooses a bandwidth to optimize the tradeoff between functional form misspecification bias and statistical precision, not to pick the bandwidth that guarantees perfect linearity). Although some of the best-fit curves appeared nonlinear within the bandwidth (for example, in Figure A.46, the best fit curve does not appear to be linear within the bandwidth to the left of the cutoff), this finding appeared to be due to overfitting sparse data.

5. Finite sample bias

The LATE estimator is susceptible to finite sample bias if the cutoff on the assignment variable is a weak instrument for SIG receipt—that is, if δ^P is too small or if the standard error of δ^P is too large. Using the criterion for weak instruments from Stock and Yogo (2005), we deemed δ^P to be a weak instrument in grades in which its *t*-statistic was less than 4.

We found no evidence of weak instruments. In all cases, the *t*-statistics were above 4 (Table A.21).

E. Sensitivity analyses

We assessed the sensitivity of the benchmark impact estimates to (1) the choice of bandwidth, (2) accounting for student mobility, (3) including impacts estimated using the graduation rate assignment variable, (4) the choice of aggregation weight, (5) functional form, (6) covariate adjustment, and (7) taking an approach that differs from the benchmark in multiple ways simultaneously (described below).

1. Choice of bandwidth

To assess the sensitivity of impact findings to the choice of bandwidth, we examined how impacts varied across the distribution of bandwidths estimated through bootstrapping. Specifically, for each bootstrap replication, we saved the selected bandwidth. We then calculated an impact by applying each bandwidth to the actual (not bootstrapped) data. Because bandwidth estimates varied considerably across bootstrap replications, this approach provided a comprehensive view of the relationship between impacts and bandwidth choice.

We concluded that our main finding—that the impact of SIG was statistically insignificant — was not sensitive to bandwidth selection. To arrive at this conclusion, we calculated fuzzy RDD impact estimates corresponding to a number of bandwidths (Table A.22). Regardless of bandwidth, no estimates were statistically significant, providing evidence that our main finding was not sensitive to bandwidth choice.

Table A.21. Assessing the strength of the cutoff on the assignment variable as an instrument for SIG receipt

	Student's *t*-Statistic for δ^P, by Outcome and Year											
	Math test scores			Reading test scores			High school graduation			College enrollment		
Impacts by grade and assignment variable	**2010–11**	**2011–12**	**2012–13**	**2010–11**	**2011–12**	**2012–13**	**2010–11**	**2011–12**	**2012–13**	**2010–11**	**2011–12**	**2012–13**
Grade 3, achievement	47.87	40.60	39.25	48.40	41.17	39.43	NA	NA	NA	NA	NA	NA
Grade 4, achievement	47.35	41.57	41.84	50.28	40.18	43.78	NA	NA	NA	NA	NA	NA
Grade 5, achievement	49.38	35.22	37.94	48.02	35.96	38.83	NA	NA	NA	NA	NA	NA
Grade 6, achievement	36.42	33.71	37.59	35.36	34.50	37.06	NA	NA	NA	NA	NA	NA
Grade 7, achievement	35.98	26.02	31.09	34.51	26.93	30.67	NA	NA	NA	NA	NA	NA
Grade 8, achievement	30.41	23.91	26.90	29.67	24.93	28.27	NA	NA	NA	NA	NA	NA
Grade 9, achievement	25.45	22.40	25.96	22.89	20.20	21.94	NA	NA	NA	NA	NA	NA
Grade 10, achievement	10.77	16.34	22.07	15.66	16.70	27.95	NA	NA	NA	NA	NA	NA
Grade 11, achievement	22.24	16.22	14.98	21.23	17.57	17.53	NA	NA	NA	NA	NA	NA
Grade 12, achievement	15.57	16.31	19.30	18.32	16.54	18.02	19.47	22.46	22.92	10.17	5.89	NA

Source: State and district administrative records.

Note: The *t*-statistics shown in this table correspond to the impact on SIG receipt from the benchmark fuzzy regression discontinuity design (RDD) impacts. We report impacts by outcome, year, grade, and assignment variable. NA indicates cases for which we could not calculate impacts because those outcomes did not exist for grades 3-11.

Table A.22. Sensitivity of findings to alternative bandwidths, for impacts on student outcomes and per-pupil spending, place-based analysis

Outcome	25th Percentile of Bootstrap Bandwidths		Benchmark Bandwidth		75th Percentile of Bootstrap Bandwidths	
	Bandwidth	Fuzzy impact	Bandwidth	Fuzzy impact	Bandwidth	Fuzzy impact
First outcome year: 2010–2011						
Math test scores	0.79	-0.03	0.86	-0.02	1.10	0.01
Reading test scores	0.79	0.03	0.87	0.03	1.14	0.04
High school graduation	0.73	0.00	0.62	-0.02	1.04	0.07
College enrollment	0.79	-0.11	0.77	-0.11	1.06	-0.08
Second outcome year: 2011–2012						
Math test scores	0.79	0.05	0.84	0.04	1.10	0.05
Reading test scores	0.78	0.12	0.86	0.10	1.08	0.10
High school graduation	0.69	-0.05	1.01	0.04	0.96	0.00
College enrollment	0.75	0.01	0.82	0.02	1.00	0.14
Per-pupil spending	0.31	$742	0.50	$245	0.45	$234
Third outcome year: 2012–2013						
Math test scores	0.82	0.06	1.08	0.01	1.14	0.02
Reading test scores	0.80	0.13	0.96	0.08	1.10	0.06
High school graduation	0.68	-0.05	0.70	-0.05	0.96	0.00
College enrollment	NA	NA	NA	NA	NA	NA
Per-pupil spending	0.35	$1,916	0.56	$100	0.47	$38

Source: State and district administrative records; interviews with district administrators and surveys of school administrators in spring 2012 and 2013.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. For each bootstrap replication, we saved the selected bandwidths for the numerator and denominator of the fuzzy impact estimate. We then calculated the 25th and 75th percentiles of the numerator bandwidths. This table shows the fuzzy impacts corresponding to those bandwidths. The bandwidths used for the denominator of the fuzzy impact are the ones that corresponded to the numerator bandwidths. Units for the bandwidth are standard deviations (assignment variables were standardized to have a standard deviation of 1). NA indicates cases for which we could not calculate impacts due to insufficient sample sizes. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

2. Accounting for student mobility

Our benchmark analysis included students present at study schools at the end of each follow-up year. If the intervention affected which schools students attended, this impact estimate conflated impacts on student achievement with impacts on mobility. In a sensitivity analysis, we controlled for mobility by focusing on the students that were slated to attend each school (based on the school they attended in the baseline year and typical school feeder patterns in the district) rather than the school they actually attended. Unlike the benchmark analysis, this sensitivity test included schools that implemented the closure model, because we analyzed outcomes for students who were slated to attend the closure schools had they not closed. In Table A.23, we report student sample sizes for this sensitivity analysis overall and within the numerator and denominator bandwidths, above and below the cutoff, by year and by outcome. We also report sample sizes in terms of unique values of the assignment variable.

Our findings did not change when we accounted for student mobility; all impacts were insignificant (Table A.24).

Table A.23. Sample sizes by assignment variable and bandwidth descriptions, accounting for student mobility

Outcome	Number of Students, Overall		Number of Unique Values of the Assignment Variable, Overall		Number of Students in Numerator Bandwidth		Number of Unique Values of the Assignment Variable in Numerator Bandwidth		Number of Students in Denominator Bandwidth		Number of Unique Values of the Assignment Variable in Denominator Bandwidth	
	I	C	I	C	I	C	I	C	I	C	I	C
Outcome year: 2010–2011												
Math test scores	91,230	187,010	230	670	70,790	100,110	190	340	67,410	83,180	180	290
Reading test scores	93,140	188,710	230	660	84,840	144,960	210	490	68,750	85,340	180	290
High school graduation	15,490	23,650	80	120	10,380	8,530	50	50	11,350	10,020	60	60
College enrollment	5,540	11,770	50	80	4,680	3,380	30	30	4,830	3,760	40	30
Outcome year: 2011–2012												
Math test scores	80,920	172,000	220	660	53,640	63,430	170	240	59,640	74,480	180	280
Reading test scores	81,890	170,630	220	660	54,150	61,240	170	230	57,030	67,200	170	250
High school graduation	14,920	24,490	80	120	9,380	9,930	50	60	11,190	11,910	60	70
College enrollment	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Outcome year: 2012–2013												
Math test scores	76,930	146,330	210	650	57,520	70,230	180	300	56,880	66,950	170	280
Reading test scores	74,770	144,430	210	640	51,750	57,800	170	250	55,650	66,720	180	290
High school graduation	12,770	22,120	80	120	9,940	8,910	60	60	8,320	7,710	50	50
College enrollment	3,330	7,950	40	70	2,860	1,780	30	20	2,700	1,690	30	20

Source: State and district administrative records; interviews with district administrators in spring 2012 and 2013.

Note: For each outcome, this table shows the number of students attending schools on either side (intervention or comparison) of the regression discontinuity design (RDD) cutoff, both overall and within the bandwidths. Bandwidths for the numerator and denominator of the fuzzy RDD impact estimate were selected independently from each other. The table also shows the number of unique values of the assignment variable on either side of the RDD cutoff, both overall and within the bandwidths. The number of unique values of the assignment variable is less than or equal to the number of schools, because some schools could have the same value of the assignment variable. NA indicates cases for which we could not calculate impacts due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten.

C = comparison; I = intervention.

Table A.24. Impacts of SIG-funded intervention models on student outcomes, accounting for student mobility

Outcome	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
First outcome year: 2010–2011							
Math test scores	-0.64	-0.65	0.01	0.02	0.79	890	278,240
Reading test scores	-0.67	-0.69	0.03	0.03	0.67	880	281,850
High school graduation	0.73	0.75	-0.02	-0.02	0.86	170	39,140
College enrollment	0.48	0.48	0.00	0.00	0.99	90	17,310
Second outcome year: 2011–2012							
Math test scores	-0.61	-0.64	0.03	0.00	0.99	880	252,920
Reading test scores	-0.63	-0.69	0.06	0.06	0.43	870	252,520
High school graduation	0.78	0.76	0.02	0.04	0.63	170	39,400
College enrollment	NA	NA	NA	NA	NA	NA	NA
Third outcome year: 2012–2013							
Math test scores	-0.56	-0.58	0.02	0.06	0.49	860	223,260
Reading test scores	-0.60	-0.63	0.03	0.06	0.50	850	219,200
High school graduation	0.76	0.72	0.04	0.06	0.58	170	34,890
College enrollment	0.39	0.41	-0.02	-0.04	0.77	80	11,280

Source: State and district administrative records; interviews with district administrators and surveys of school administrators in spring 2012 and 2013.

Note: We standardized test score outcomes to have a standard deviation of 1, so we report test score impact estimates in effect-size units. Units for high school graduation and college enrollment are percentage points/100. For example, an impact of 0.1 indicates an increase of 10 percentage points. Adjusted mean outcomes for the intervention and comparison groups are equal to the estimated intercept terms from the regressions of the outcome on the assignment variable and other covariates that were estimated separately on either side of the regression discontinuity design (RDD) cutoff value. The sample sizes reported in this table include both schools/students within the bandwidth and schools/students outside the bandwidth. NA indicates cases for which we could not calculate impacts due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

a. Diagnostic analyses for impact estimates accounting for student mobility

We conducted the same diagnostic analyses that we conducted for our benchmark, place-based impact estimates. We summarize the findings from these diagnostic analyses below:

- **Integrity of the assignment variable.** Using the McCrary test, we found no discontinuities involving the achievement assignment variable, but we found two statistically significant discontinuities involving the graduation rate assignment variable (Table A.25). Due to these discontinuities, we decided to exclude from the analysis all impacts estimated using the graduation rate assignment variable. We found no graphical evidence of discontinuities involving the density of the achievement assignment variable (Figures A.48 to A.69). In each figure, the confidence intervals at the cutoff overlap each other, indicating no graphical evidence of density discontinuities.
- **Sample attrition.** Overall and differential attrition were very low for math and reading test score outcomes. Attrition was moderate for the high school graduation outcome, and high for the college enrollment outcome. (Table A.26).

Table A.25. McCrary test results, accounting for student mobility

	McCrary Test *p*-Value, by Outcome and Year											
	Math test scores			Reading test scores			High school graduation			College enrollment		
Grade and assignment variable	**2010–11**	**2011–12**	**2012–13**	**2010–11**	**2011–12**	**2012–13**	**2010–11**	**2011–12**	**2012–13**	**2010–11**	**2011–12**	**2012–13**
Grade 3, achievement	0.82	0.76	NA	0.57	0.77	NA	NA	NA	NA	NA	NA	NA
Grade 4, achievement	0.75	0.94	0.53	0.50	0.95	0.61	NA	NA	NA	NA	NA	NA
Grade 5, achievement	0.72	0.98	0.83	0.56	0.97	0.92	NA	NA	NA	NA	NA	NA
Grade 6, achievement	0.88	0.91	0.91	0.57	0.89	1.00	NA	NA	NA	NA	NA	NA
Grade 7, achievement	0.51	0.60	0.42	0.32	0.59	0.37	NA	NA	NA	NA	NA	NA
Grade 8, achievement	0.57	0.59	0.42	0.34	0.56	0.41	NA	NA	NA	NA	NA	NA
Grade 9, both, 1[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 9, both, 2[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 9, achievement	0.17	0.39	0.61	0.10	0.39	0.26	NA	NA	NA	NA	NA	NA
Grade 9, grad rate	0.49	NA	NA	0.69	0.97	NA	NA	NA	NA	NA	NA	NA
Grade 10, both, 1[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 10, both, 2[a]	NA	NA	0.46	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 10, achievement	0.46	0.98	0.83	0.43	0.80	0.74	NA	NA	NA	NA	NA	NA
Grade 10, grad rate	0.33	0.33	0.14	0.09	0.62	0.05*	NA	NA	NA	NA	NA	NA
Grade 11, both, 1[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 11, both, 2[a]	NA	0.82	0.39	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 11, achievement	0.71	0.49	0.67	0.50	0.66	0.80	NA	NA	NA	NA	NA	NA
Grade 11, grad rate	0.07	0.07	0.11	0.07	0.09	0.04*	NA	NA	NA	NA	NA	NA
Grade 12, both, 1[a]	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 12, both, 2[a]	NA	NA	0.39	NA	NA	NA	NA	NA	NA	NA	NA	NA
Grade 12, achievement	0.86	0.88	0.63	0.36	0.53	0.56	1.00	0.82	0.77	0.64	NA	0.10
Grade 12, grad rate	0.19	0.12	0.11	0.12	NA	0.17	0.28	0.25	0.07	0.22	0.19	0.40

Source: State and district administrative records.

Note: This table reports results of the McCrary test (2008) for discontinuities in the density of the assignment variable at the cutoff. We report McCrary *p*-values by outcome, year, grade, and assignment variable. NA indicates cases for which we could not calculate impacts due to insufficient sample sizes, and so we did not conduct the McCrary test.

[a] Both, 1 indicates a mini-study with two assignment variables, and with this row is reporting on the achievement assignment variable. Both, 2 indicates a mini-study with two assignment variables, and with this row is reporting on the graduation rate assignment variable.

*Statistically significant at the 0.05 level, two-tailed test.

Table A.26. Impacts of SIG-funded intervention models on student-level attrition, accounting for student mobility

Outcome	Intervention Mean	Comparison Mean	Impact Estimate	*p*-Value
Outcome year: 2010–2011				
Math test scores	0.01	0.01	0.00	0.68
Reading test scores	0.01	0.00	0.00	1.00
High school graduation	0.23	0.33	-0.10	0.53
College enrollment	0.48	0.51	-0.04	0.87
Outcome year: 2011–2012				
Math test scores	0.01	0.01	0.00	1.00
Reading test scores	0.01	0.00	0.00	1.00
High school graduation	0.15	0.17	-0.02	0.89
College enrollment	0.35	0.67	-0.32	0.16
Outcome year: 2012–2013				
Math test scores	0.01	0.00	0.00	0.60
Reading test scores	0.01	0.00	0.00	0.90
High school graduation	0.17	0.29	-0.13	0.39
College enrollment	0.35	0.86	-0.50*	0.03

Source: State and district administrative records; interviews with district administrators in spring 2012 and 2013.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. Adjusted mean outcomes for the intervention and comparison groups were equal to the estimated intercept terms from the regressions of the outcome on the assignment variable and other covariates that were estimated separately on either side of the RDD cutoff value.

*Statistically significant at the 0.05 level, two-tailed test.

- **Continuity of the relationship between the assignment variable and the outcome.** There were no statistically significant baseline differences with respect to any covariate (Table A.17) or baseline outcome measure (Tables A.27 and A.28), suggesting a smooth relationship between the outcome and the assignment variable at the cutoff. Table A.17 applies to both the place-based and mobility-robust analyses because it focuses on the sample of students and schools for which baseline covariates were available, as opposed to the sample for which either the place-based or mobility-robust outcomes were available.
  - In addition, we found no clear evidence of discontinuities in the relationship between the outcome and the assignment variable in the bandwidth at points other than the cutoff for math and reading test scores or college enrollment (Figures A.70 to A.91; Table A.29). For the high school graduation rate outcome, however, there was some evidence of a discontinuity in 2011–2012: we found one statistically significant discontinuity out of 10 values tested.
- **Bandwidth selection.** The relationship between the outcome and assignment variable inside the bandwidth appeared to be approximately linear (Figures A.70 to A.91).
  - In addition, we concluded that our main finding accounting for student mobility—that the impact of SIG was statistically insignificant—was not sensitive to bandwidth selection (Table A.30).
- **Finite sample bias.** We found no evidence of weak instruments using the criterion for weak instruments from Stock and Yogo (2005). In all cases, the *t*-statistics were above 4 (Table A.31).

Table A.27. Comparison of baseline math achievement for the intervention and comparison groups used to estimate impacts on student outcomes, mobility-robust analysis

Baseline Math Achievement for Samples Used to Estimate Impacts on the Following Outcomes:[a]	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
First outcome year: 2010–2011							
Math test scores	-0.64	-0.67	0.03	0.02	0.69	810	225,890
Reading test scores	-0.64	-0.67	0.03	0.04	0.46	800	228,970
High school graduation	-1.04	-1.08	0.04	0.06	0.84	130	35,960
College enrollment	-1.48	-1.25	-0.23	-0.30	0.60	70	14,750
Second outcome year: 2011–2012							
Math test scores	-0.58	-0.60	0.01	0.01	0.88	790	185,390
Reading test scores	-0.58	-0.60	0.01	0.01	0.89	790	184,360
High school graduation	-0.98	-0.85	-0.13	-0.16	0.53	100	25,060
College enrollment	NA	NA	NA	NA	NA	NA	NA
Per-pupil spending	NA	NA	NA	NA	NA	NA	NA
Third outcome year: 2012–2013							
Math test scores	-0.46	-0.55	0.08	0.14	0.22	690	184,710
Reading test scores	-0.45	-0.53	0.08	0.13	0.24	680	181,190
High school graduation	-0.68	-0.73	0.05	0.07	0.76	130	27,630
College enrollment	NA	NA	NA	NA	NA	NA	NA
Per-pupil spending	NA	NA	NA	NA	NA	NA	NA

Source: State and district administrative records; surveys of school administrators in spring 2012; Common Core of Data (CCD).

Note: We calculated intervention-comparison group differences using the mobility-robust approach to calculating regression discontinuity design (RDD) impact estimates described in this appendix, except that no additional covariates were included in the model. In particular, for each outcome and year, we used the same bandwidth to calculate intervention-comparison group differences as we used to estimate RDD impacts for the relevant outcome and year. We standardized all outcomes to have a standard deviation of 1, so we report impact estimates in effect size units. NA indicates cases for which we could not calculate impacts or *p*-values due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

[a] The outcome in this table is baseline math achievement. The first column indicates which sample was used to examine baseline math achievement. For example, the second row examines baseline math achievement for the sample of schools that were used in the mobility-robust analysis of reading test scores in 2010–2011.

Table A.28. Comparison of baseline reading achievement for the intervention and comparison groups used to estimate impacts on student outcomes, mobility-robust analysis

Baseline Reading Achievement for Samples Used to Estimate Impacts on the Following Outcomes:[a]	Intervention Mean	Comparison Mean	Sharp Impact Estimate	Fuzzy Impact Estimate	*p*-Value for Fuzzy Impact Estimate	School Sample Size	Student Sample Size
First Outcome year: 2010–2011							
Math test scores	-0.66	-0.71	0.05	0.07	0.28	810	225,620
Reading test scores	-0.63	-0.69	0.06	0.08	0.14	800	228,880
High school graduation	-0.82	-0.86	0.04	0.06	0.86	130	35,790
College enrollment	-1.15	-0.59	-0.56	-0.74	0.23	70	14,670
Second outcome year: 2011–2012							
Math test scores	-0.62	-0.60	-0.02	-0.03	0.71	790	185,250
Reading test scores	-0.61	-0.62	0.01	0.00	0.99	790	184,240
High school graduation	-0.73	-0.65	-0.07	-0.09	0.73	100	25,060
College enrollment	NA	NA	NA	NA	NA	NA	NA
Per-pupil spending	NA	NA	NA	NA	NA	NA	NA
Third outcome year: 2012–2013							
Math test scores	-0.51	-0.52	0.02	0.06	0.61	690	184,730
Reading test scores	-0.51	-0.51	0.00	0.03	0.78	680	181,200
High school graduation	-0.62	-0.68	0.05	0.08	0.79	130	27,460
College enrollment	NA	NA	NA	NA	NA	NA	NA
Per-pupil spending	NA	NA	NA	NA	NA	NA	NA

Source: State and district administrative records; surveys of school administrators in spring 2012; Common Core of Data (CCD).

Note: We calculated intervention-comparison group differences using the mobility-robust approach to calculating regression discontinuity design (RDD) impact estimates described in this appendix, except that no additional covariates were included in the model. In particular, for each outcome and year, we used the same bandwidth to calculate intervention-comparison group differences as we used to estimate RDD impacts for the relevant outcome and year. We standardized all outcomes to have a standard deviation of 1, so we report impact estimates in effect-size units. NA indicates cases for which we could not calculate impacts or *p*-values due to insufficient sample sizes. To comply with NCES reporting requirements, we rounded sample sizes to the nearest ten. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

[a] The outcome in this table is baseline reading achievement. The first column indicates which sample was used to examine baseline reading achievement. For example, the first row examines baseline reading achievement for the sample of schools that were used in the mobility-robust analysis of math test scores in 2010–2011.

Table A.29. Assessing the continuity of the relationship between the outcome and assignment variable for student outcomes, accounting for student mobility

Outcome	Number of Assignment Variable Values Other than the Cutoff for Which We Tested for Discontinuities in the Outcome	Percentage of Assignment Variable Values with Statistically Significant Discontinuities in the Outcome
Outcome year: 2010–2011		
Math test scores	100	0
Reading test scores	100	3
High school graduation	10	0
College enrollment	10	0
Outcome year: 2011–2012		
Math test scores	100	2
Reading test scores	100	3
High school graduation	10	10
College enrollment	NA	NA
Outcome year: 2012–2013		
Math test scores	90	3
Reading test scores	90	3
High school graduation	10	0
College enrollment	10	0

Source: State and district administrative records; interviews with district administrators in spring 2012 and 2013.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. For each outcome and grade, we tested for discontinuities in the outcome-assignment variable relationship at several values of the assignment variable other than the actual cutoff (that is, at values of the assignment variable where no discontinuities are expected). Specifically, we tested for discontinuities at four values above and four values below the actual cutoff. These values were all inside the IK bandwidth used for the numerator of the fuzzy impact estimates. The values were evenly spaced across the distribution of the assignment variable inside the bandwidth. NA indicates cases for which we could not conduct tests due to insufficient sample sizes.

IK = Imbens and Kalyanaraman.

Table A.30. Sensitivity of findings to alternative bandwidths, for impacts on student outcomes, accounting for student mobility

Outcome	25th Percentile of Bootstrap Bandwidths		Benchmark Bandwidth		75th Percentile of Bootstrap Bandwidths	
	Bandwidth	Fuzzy impact	Bandwidth	Fuzzy impact	Bandwidth	Fuzzy impact
First outcome year: 2010–2011						
Math test scores	0.79	0.00	1.21	0.02	1.10	0.02
Reading test scores	0.81	0.05	1.82	0.03	1.12	0.06
High school graduation	0.73	-0.06	0.85	-0.02	1.01	0.00
College enrollment	0.78	-0.01	0.76	0.00	1.07	-0.02
Second outcome year: 2011–2012						
Math test scores	0.79	0.00	0.91	0.00	1.07	0.01
Reading test scores	0.75	0.07	0.90	0.06	1.00	0.08
High school graduation	0.70	0.00	0.93	0.04	0.99	0.05
College enrollment	NA	NA	NA	NA	NA	NA
Third outcome year: 2012–2013						
Math test scores	0.79	0.09	1.12	0.06	1.10	0.07
Reading test scores	0.77	0.11	0.96	0.06	1.08	0.07
High school graduation	0.67	0.01	1.02	0.06	0.92	0.05
College enrollment	0.77	-0.04	0.84	-0.04	1.04	0.01

Source: State and district administrative records; interviews with district administrators and surveys of school administrators in spring 2012 and 2013.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. For each bootstrap replication, we saved the selected bandwidths for the numerator and denominator of the fuzzy impact estimate. We then calculated the 25th and 75th percentiles of the numerator bandwidths. This table shows the fuzzy impacts corresponding to those bandwidths. The bandwidths used for the denominator of the fuzzy impact are the ones that corresponded to the numerator bandwidths. Units for the bandwidth are standard deviations (assignment variables were standardized to have a standard deviation of 1). NA indicates cases for which we could not calculate impacts due to insufficient sample sizes. No impacts were statistically significant at the 0.05 level, using a two-tailed test.

Table A.31. Assessing the strength of the cutoff on the assignment variable as an instrument for SIG receipt, accounting for student mobility

	Student's *t*-Statistic for δ^P, by Outcome and Year											
	Math test scores			Reading test scores			High school graduation			College enrollment		
Impacts by grade and assignment variable	**2010–11**	**2011–12**	**2012–13**	**2010–11**	**2011–12**	**2012–13**	**2010–11**	**2011–12**	**2012–13**	**2010–11**	**2011–12**	**2012–13**
Grade 3, achievement	39.70	54.71	NA	41.24	54.89	NA	NA	NA	NA	NA	NA	NA
Grade 4, achievement	60.99	61.22	60.80	58.80	58.53	62.51	NA	NA	NA	NA	NA	NA
Grade 5, achievement	61.41	57.38	62.10	59.75	52.88	65.42	NA	NA	NA	NA	NA	NA
Grade 6, achievement	41.88	39.52	46.35	40.13	38.75	46.47	NA	NA	NA	NA	NA	NA
Grade 7, achievement	41.85	38.56	40.42	41.27	37.21	43.14	NA	NA	NA	NA	NA	NA
Grade 8, achievement	33.51	34.96	39.45	33.94	34.64	40.14	NA	NA	NA	NA	NA	NA
Grade 9, achievement	15.37	16.39	24.12	15.13	16.14	18.31	NA	NA	NA	NA	NA	NA
Grade 10, achievement	15.79	5.46	12.01	17.11	4.13	16.72	NA	NA	NA	NA	NA	NA
Grade 11, achievement	26.67	22.46	11.70	25.03	24.53	10.98	NA	NA	NA	NA	NA	NA
Grade 12, achievement	25.83	32.34	30.71	28.58	27.20	27.56	21.24	20.27	22.28	10.54	NA	9.66

Source: State and district administrative records.

Note: The *t*-statistics shown in this table correspond to the impact on SIG receipt from the benchmark fuzzy regression discontinuity design (RDD) impacts. We report impacts by outcome, year, grade, and assignment variable. NA indicates cases for which we could not calculate impacts because those outcomes did not exist for grades 3-11.

3. Including impacts estimated using the graduation rate assignment variable

As mentioned previously, because we found significant discontinuities in the density of the graduation rate assignment variable, we decided to exclude from our benchmark analysis all impacts estimated using the graduation rate. However, we conducted a sensitivity analysis that includes impacts using the graduation rate. No impacts from this sensitivity analysis were significant (Table A.32).

4. Choice of aggregation weight

We calculated the overall impact of SIG as a weighted average of the grade-specific impacts. In the benchmark analysis, we used a sample size weight, but we conducted two sensitivity tests to assess whether findings changed when we used different weights.

In the first sensitivity analysis, we weighted by the inverse of the covariance matrix of the grade-specific impacts, meaning that the more precisely estimated impacts received a greater weight. Our findings did not change when we used the inverse variance weight. None of the impacts were significant (Table A.32).

In the second sensitivity analysis, we used a weight equal to the student sample at baseline for the school that was closest to the cutoff value of the assignment variable for each grade-specific impact. We focused on the school closest to the cutoff because, as mentioned in Chapter II, the RDD impact estimate applies only to schools at the cutoff, which means that the impact corresponds to what might be expected if a policy change slightly expanded or contracted the SIG eligibility cutoff. Our findings did not change when we used this weighting approach; impacts were still not significant (Table A.32).

5. Functional form

In our benchmark approach, we estimated the relationship between the outcome and assignment variable using a linear functional form within a bandwidth. An alternative approach is to use all the data and estimate the relationship between the outcome and assignment variable using a higher order polynomial regression. Gelman and Imbens (2014) explicitly cautioned *against* using this approach, but we conducted this sensitivity analysis because past authors have used this approach in RDD analysis (for example, Dee 2012).

This alternative approach involved estimating a cubic polynomial regression involving all the data (not just data in a bandwidth) and produced findings very similar to those generated by the benchmark model (Table A.32). None of the impacts estimated using the polynomial functional form were statistically significant.

6. Covariate adjustment

Consistent (that is, asymptotically unbiased) RDD impacts can be estimated without including additional covariates beyond the assignment variable. In our benchmark analysis, we included covariates that we believed might be related to the outcome, in an effort to increase the statistical precision of impact estimates. In this sensitivity analysis, we examined how impacts changed if we excluded those additional covariates.

Our findings did not change when we excluded covariates. We found no significant impacts (Table A.32).

Table A.32. Impacts of SIG-funded intervention models using alternative analysis methods

Outcome	Benchmark	Include Impacts Using the Graduation Rate Assignment Variable	Accounting for Student Mobility	Aggregated Using Inverse-Variance Weight	Aggregated Using Weights Equal to Baseline Student Sample for School Closest to Cutoff	Alternative Functional Form (Polynomial)	Excluding Covariates	Alternative Analysis with IK Bandwidth	Alternative Analysis with CV Bandwidth
Outcome year: 2010–2011									
Math test scores	-0.02	0.02	0.02	-0.01	-0.11	-0.03	0.01	0.00	0.01
Reading test scores	0.03	0.06	0.03	0.06	0.02	0.12	0.06	0.05	0.04
High school graduation	-0.02	-0.04	-0.02	-0.02	-0.02	0.07	-0.02	0.15	0.15
College enrollment	-0.11	-0.04	0.00	-0.11	-0.11	NA	-0.11	0.04	0.04
Outcome year: 2011–2012									
Math test scores	0.04	0.06	0.00	0.04	0.01	0.11	0.03	-0.07	-0.05
Reading test scores	0.10	0.09	0.06	0.11	0.05	0.41	0.07	0.01	-0.07
High school graduation	0.04	0.01	0.04	0.04	0.04	-0.07	-0.06	0.06	0.06
College enrollment	0.02	0.08	NA	0.02	0.02	NA	-0.23	-0.05	-0.05
Per-pupil spending	$245	$1,173	n.a.[a]	$245	$245	NA	$-553	$-635	$644
Outcome year: 2012–2013									
Math test scores	0.01	0.01	0.06	0.01	0.02	-0.16	0.12	-0.04	-0.02
Reading test scores	0.08	0.08	0.06	0.08	0.23	0.34	0.14	0.02	0.03
High school graduation	-0.05	-0.07	0.06	-0.05	-0.05	NA	-0.05	0.14	0.12
College enrollment	NA	0.04	-0.04	NA	NA	NA	-0.11	-0.02	0.00
Per-pupil spending	$100	$1,253	n.a.[a]	$100	$100	NA	$-562	$-1,425	$20

Source: State and district administrative records; interviews with district administrators in spring 2012 and 2013.

Note: We calculated the results shown in this table using the regression discontinuity design (RDD) methods described in this appendix. NA indicates cases for which we could not calculate impacts due to insufficient sample sizes. No estimates were statistically significant at the 0.05 level, using a two-tailed test.

[a] Per-pupil spending is a school-level outcome (as opposed to a student-level outcome), so the analysis that accounted for student mobility is not applicable to this outcome.

n.a. = not applicable; IK = Imbens and Kalyanaraman; CV = cross-validation.

7. A substantially altered approach

The sensitivity analyses described above all varied one aspect of our benchmark approach while holding other aspects constant. For this sensitivity analysis, we altered several aspects simultaneously, resulting in a substantially different approach to estimation. The altered methods were as follows:

1. We estimated a single pooled impact rather than aggregating grade-specific impacts.
2. We estimated impacts using the achievement and graduation rate assignment variables.
3. We used the binding score method to account for multiple assignment variables (Reardon and Robinson 2010).
4. Rather than estimating standard errors using the residual bootstrap algorithm described previously, we estimated bias-corrected standard errors as in Calonico et al. (2014).
5. We used either the IK algorithm or cross-validation to select the bandwidth.

Our findings did not change when we used this substantially altered approach. All of the impacts remained insignificant (Table A.32).[36]

F. Exploratory analyses

We conducted two exploratory analyses: (1) examining how impacts varied across grades, and (2) examining whether the descriptive differences in SIG practices between schools that implemented a SIG-funded intervention model and schools that did not were large enough to be statistically significant if the descriptive analysis had the same statistical power as the RDD analysis.

1. Variation in impacts across grades

We found considerable variation in the precision of impacts across grades (as represented by variation in the size of the confidence intervals), but the most precisely estimated impacts were also the ones that were closest to zero (Figure A.92). We likewise found no statistically significant impacts on any outcome in any year for any of the grades.

2. Interpreting statistical power of RDD impacts on practices relative to descriptive differences in practices

In Chapter IV, we explored why there were no significant RDD impacts of SIG-funded models on the use of SIG-promoted practices, even though the descriptive analysis showed significant differences in use of practices between schools that did and did not implement a SIG-funded model. In this section, we report the details of our analysis, investigating the extent to which the lower statistical power of the RDD analysis might explain the difference in statistical significance between the RDD and descriptive analyses.

---

[36] The last two columns of the table report these results. In each of these two columns, we applied the alternate methods (1) through (4) and used the bandwidth indicated in the column heading.

We calculated illustrative *t*-statistics that demonstrate how the statistical significance of descriptive findings would change if the descriptive analysis had the same power as the RDD impact analysis. The numerator in these *t*-statistics is the descriptive difference in practices between schools implementing a SIG-funded intervention model and schools not implementing such a model. The denominator is the standard error from an RDD impact analysis involving the same practices as outcomes.

We found that none of the descriptive differences would be statistically significant if the descriptive analysis had the same statistical power as the RDD impact analysis (Table A.33). None of these *t*-statistics is greater than 1.96, which is the threshold for statistical significance for a normally distributed test statistic when controlling type 1 errors at a rate of 5 percent.

From these findings, we concluded that the lack of statistical power in the RDD analysis is sufficient to explain why the RDD impacts were not statistically significant even though the descriptive analysis findings were significant. In other words, although our analyses show that schools implementing SIG-funded models used more SIG-promoted practices than other schools, given the lack of statistical power in the RDD analysis, we cannot conclude that SIG *caused* those observed differences.

Table A.33. Illustrative *t*-statistics

Outcome	Descriptive Difference	RDD Standard Error	Illustrative *t*-Statistic
Outcome year: 2011–2012			
Total number of SIG-promoted practices used	2.61	2.91	0.90
Total number of ELL-focused SIG-promoted practices used	0.05	1.00	0.05
Comprehensive instructional reform strategies	0.33	0.60	0.55
Using data to identify and implement an instructional program	0.04	0.15	0.27
Promoting the continuous use of student data	0.05	0.11	0.44
Providing supports and professional development to staff to assist ELLs and students with disabilities	0.04	0.18	0.22
Using and integrating technology-based supports	0.07	0.13	0.53
Tailoring strategies for secondary schools	0.06	0.85	0.07
Teacher and principal effectiveness	1.43	1.84	0.78
Using rigorous, transparent, and equitable evaluation systems	0.20	0.64	0.31
Identifying and rewarding or removing teachers and principals	0.32	0.51	0.62
Providing high quality, job-embedded professional development or supports	0.82	1.10	0.74
Implementing strategies to recruit, place, and retain staff	0.03	0.32	0.09
Learning time and community-oriented schools	0.72	0.68	1.06
Increasing learning time	0.10	0.17	0.60
Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs	0.62	0.61	1.02
Operational flexibility and support	0.17	0.29	0.59
Having operational flexibility	0.03	0.15	0.20

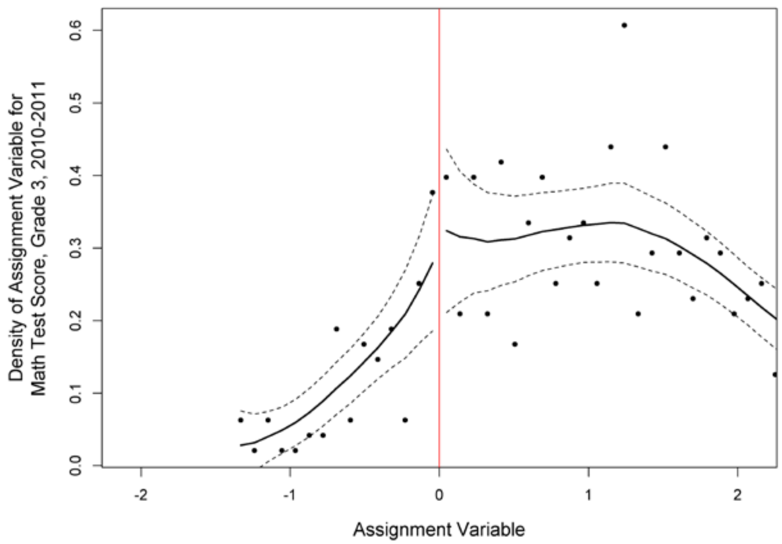
Outcome	Descriptive Difference	RDD Standard Error	Illustrative *t*-Statistic
Receiving technical assistance and support	0.14	0.18	0.77
Outcome year: 2012–2013			
Total number of SIG-promoted practices used	2.50	2.59	0.97
Total number of ELL-focused SIG-promoted practices used	0.08	0.93	0.09
Comprehensive instructional reform strategies	0.22	0.48	0.46
Using data to identify and implement an instructional program	0.01	0.14	0.07
Promoting the continuous use of student data	0.03	0.08	0.40
Providing supports and professional development to staff to assist ELLs and students with disabilities	0.02	0.16	0.12
Using and integrating technology-based supports	0.07	0.12	0.57
Tailoring strategies for secondary schools	0.06	NA	NA
Teacher and principal effectiveness	1.60	2.02	0.79
Using rigorous, transparent, and equitable evaluation systems	0.17	0.58	0.29
Identifying and rewarding or removing teachers and principals	0.33	0.55	0.60
Providing high quality, job-embedded professional development or supports	0.86	1.16	0.74
Implementing strategies to recruit, place, and retain staff	0.24	0.32	0.75
Learning time and community-oriented schools	0.47	0.76	0.62
Increasing learning time	0.13	0.15	0.85
Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs	0.35	0.74	0.48
Operational flexibility and support	0.17	0.22	0.78
Having operational flexibility	0.04	0.14	0.29
Receiving technical assistance and support	0.13	0.16	0.83

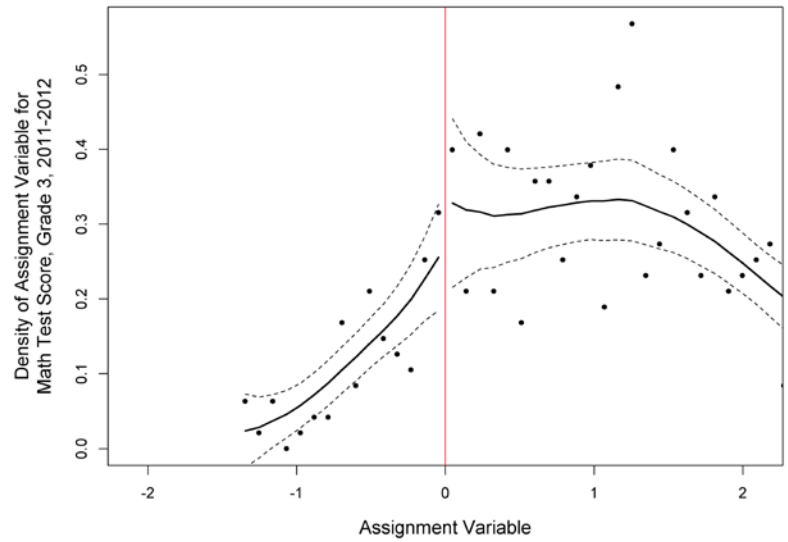
Source: State and district administrative records; surveys of school administrators in spring 2012 and 2013.

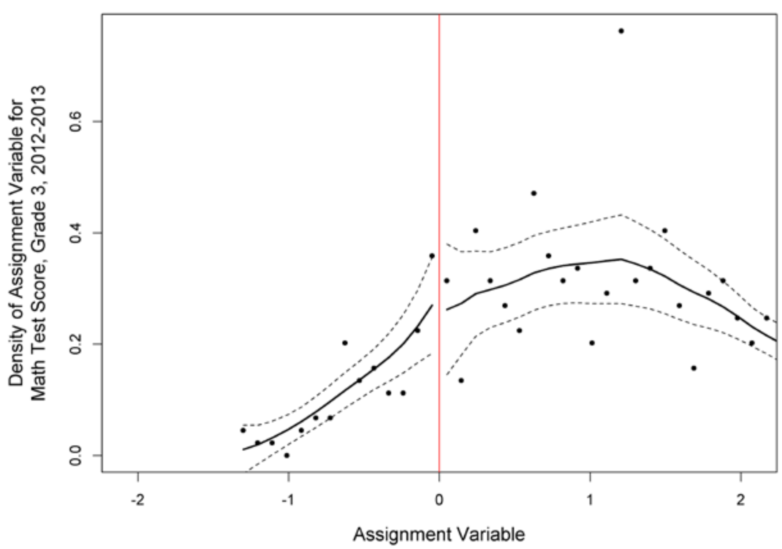
Note: The descriptive difference is the difference in mean practices between schools implementing a SIG-funded intervention model and schools not implementing one. The RDD standard error is the standard error from an RDD impact analysis using the specified outcome variable. The illustrative *t*-statistic is the ratio of the descriptive difference to the RDD standard error. NA indicates cases for which we could not calculate standard errors or t-statistics due to insufficient sample sizes.

RDD = regression discontinuity design.

Figure A.2. Density of the assignment variable for math test score in grade 3, place-based analysis



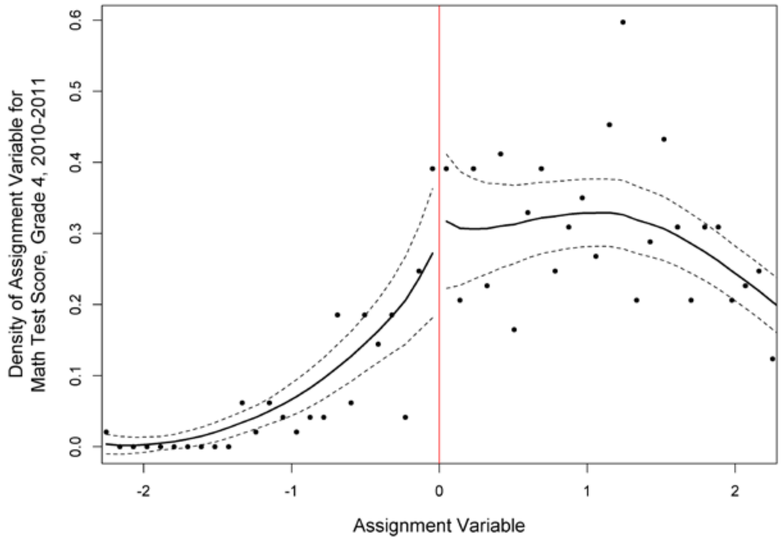


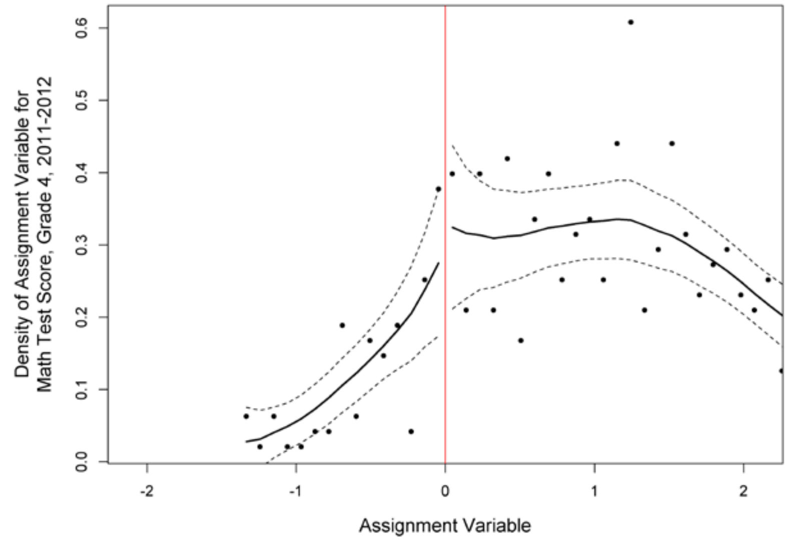


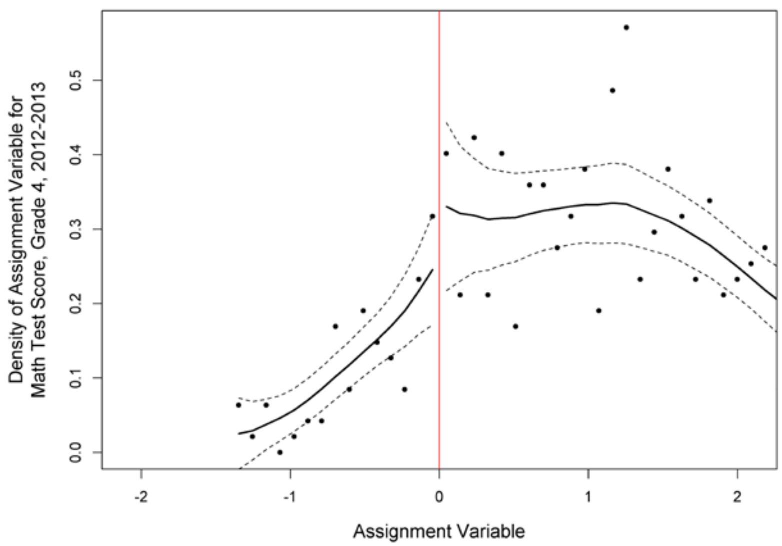
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.3. Density of the assignment variable for math test score in grade 4, place-based analysis



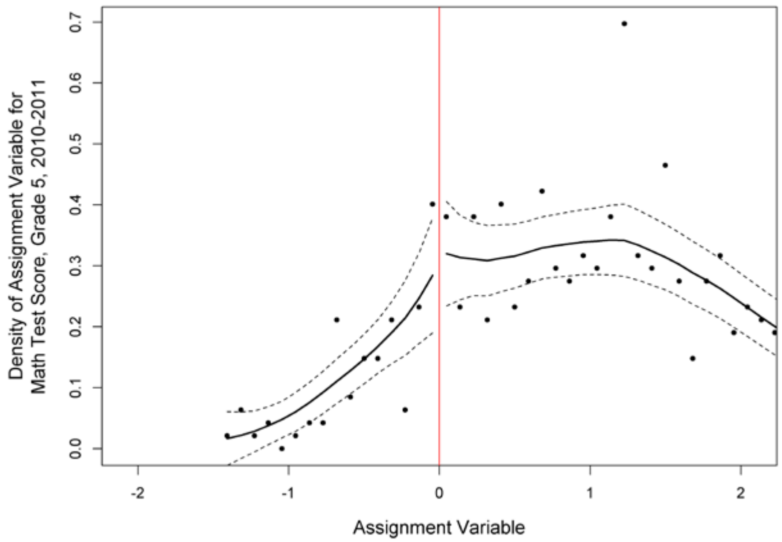


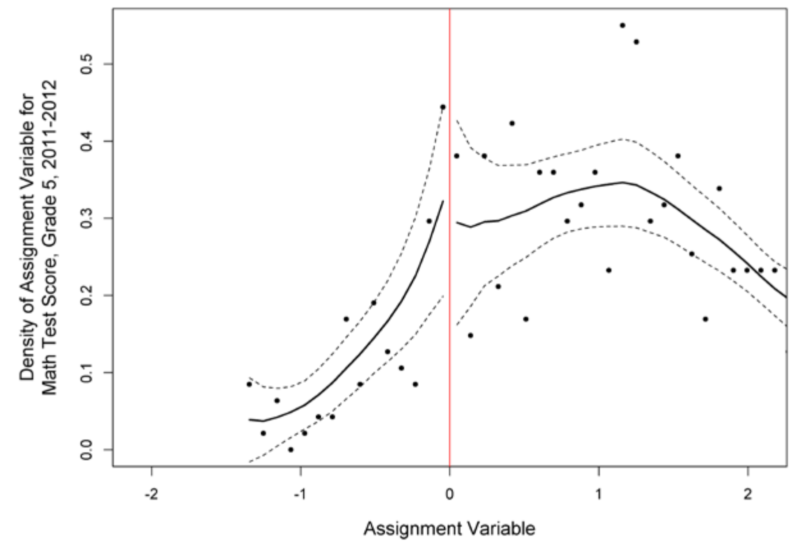


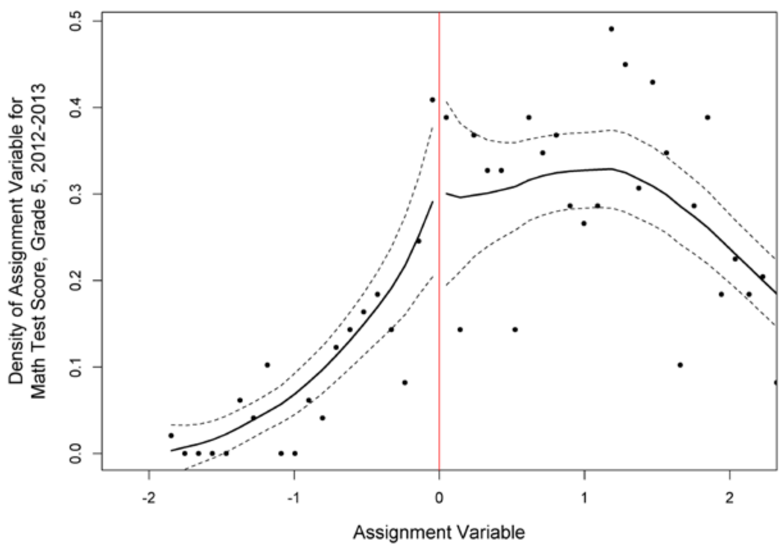
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.4. Density of the assignment variable for math test score in grade 5, place-based analysis



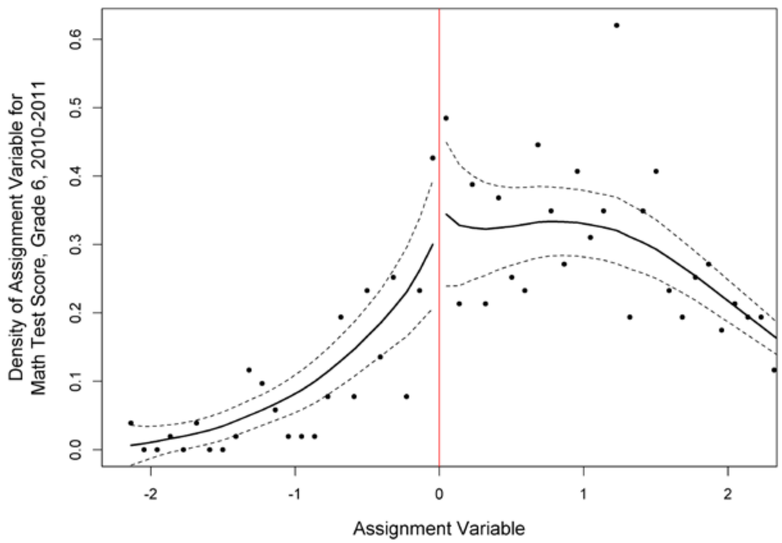


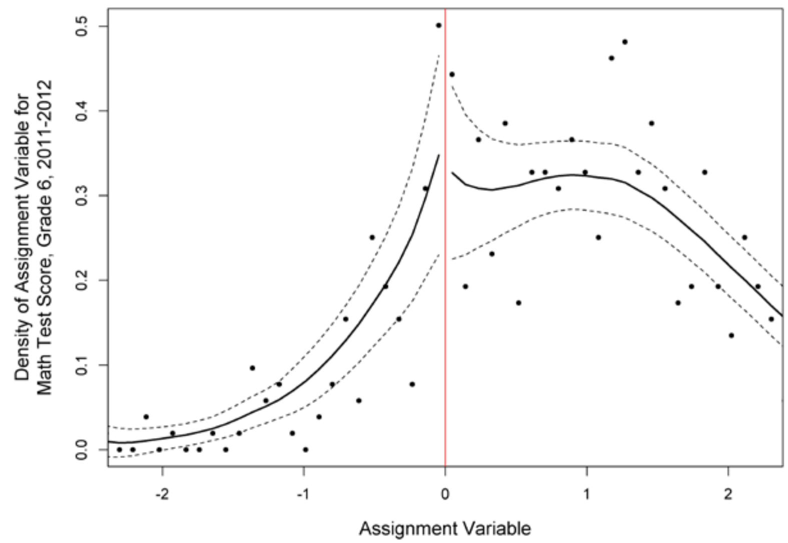


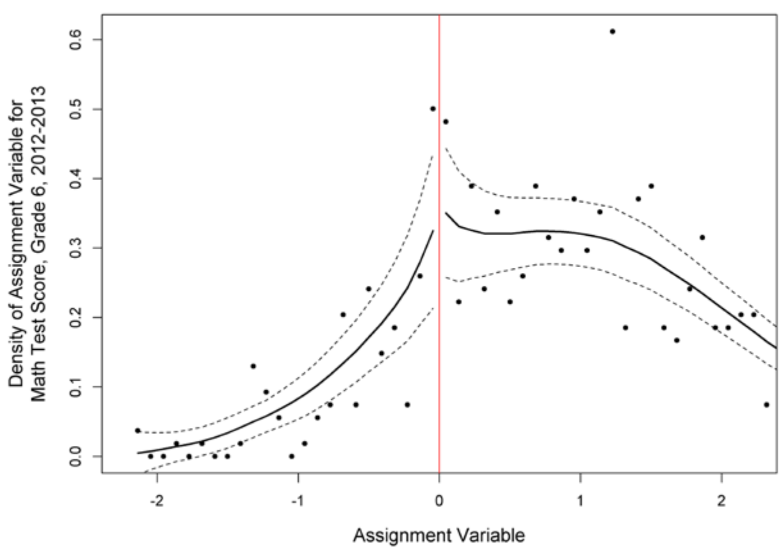
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.5. Density of the assignment variable for math test score in grade 6, place-based analysis



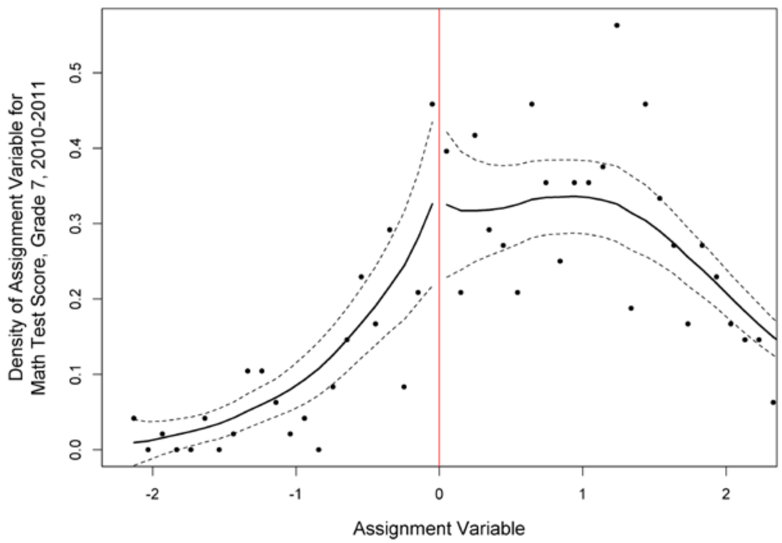


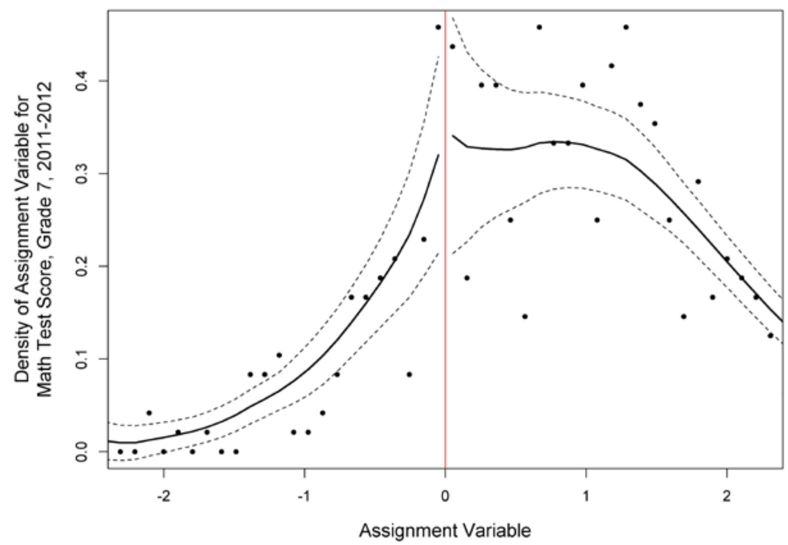


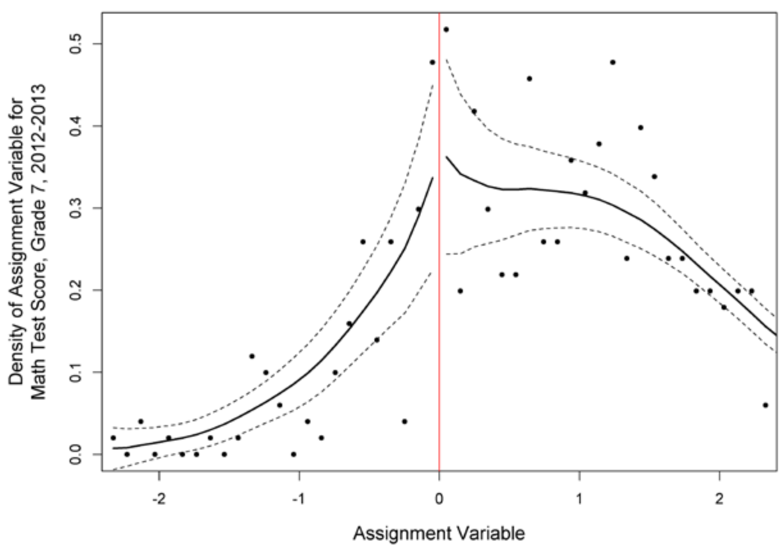
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.6. Density of the assignment variable for math test score in grade 7, place-based analysis



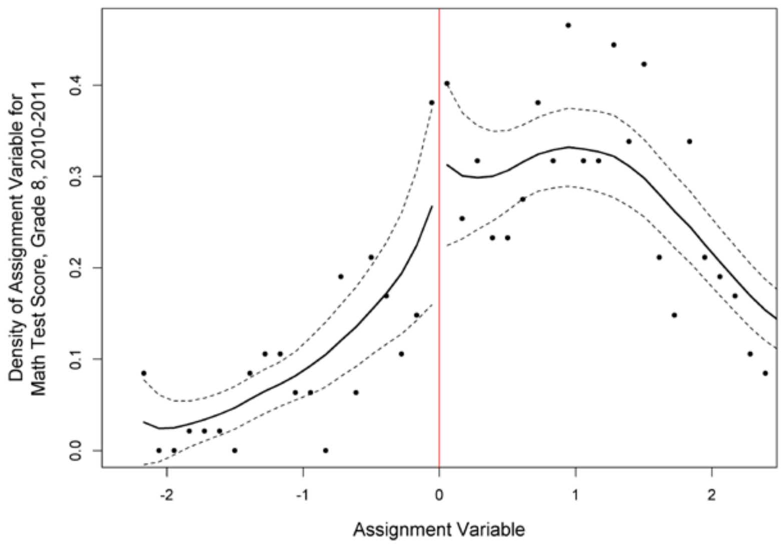


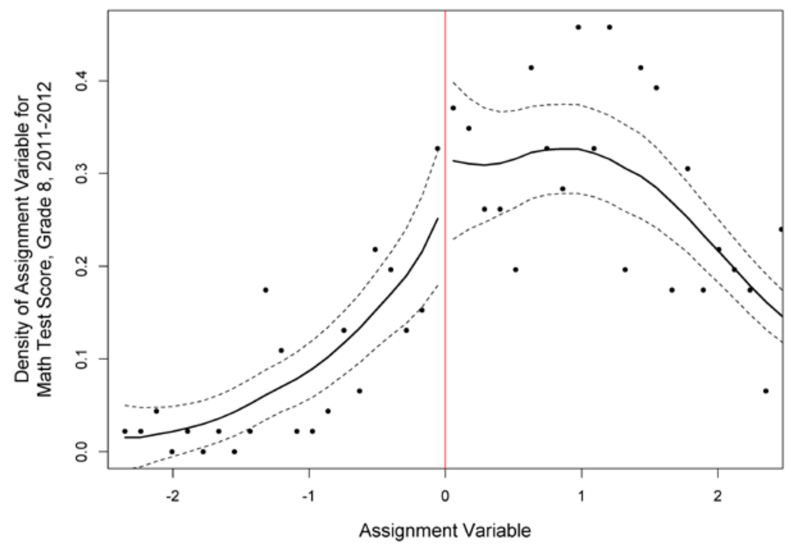


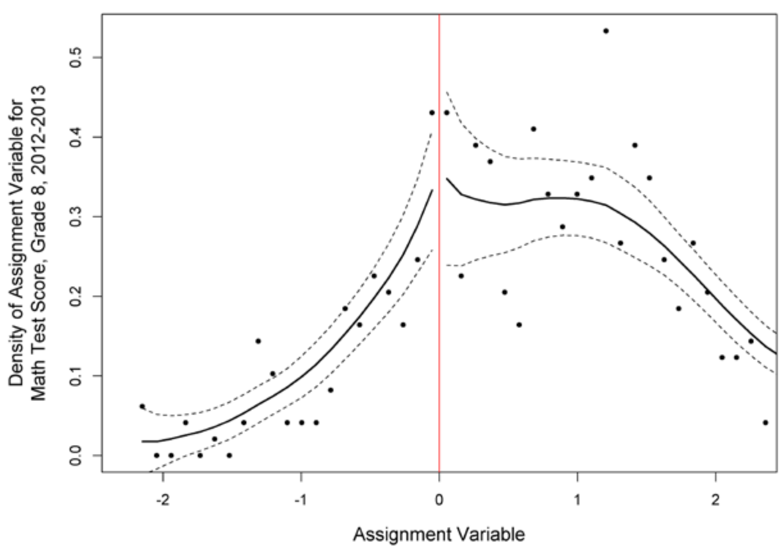
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines)

Figure A.7. Density of the assignment variable for math test score in grade 8, place-based analysis



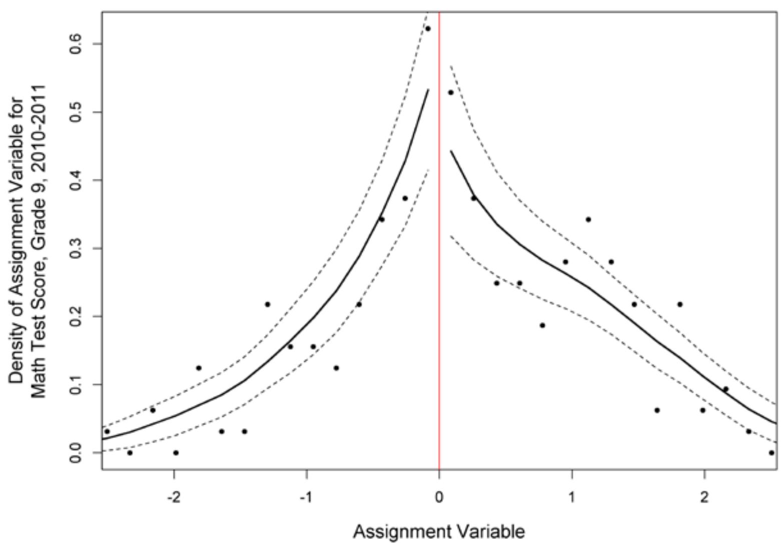


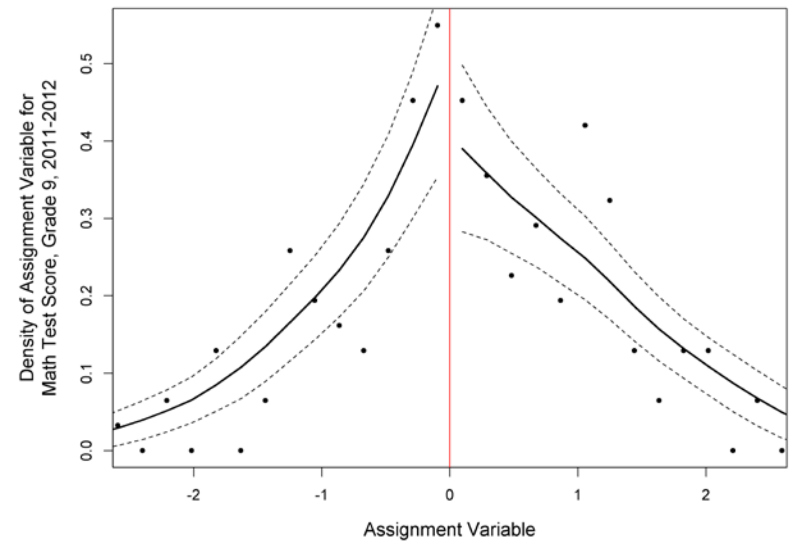


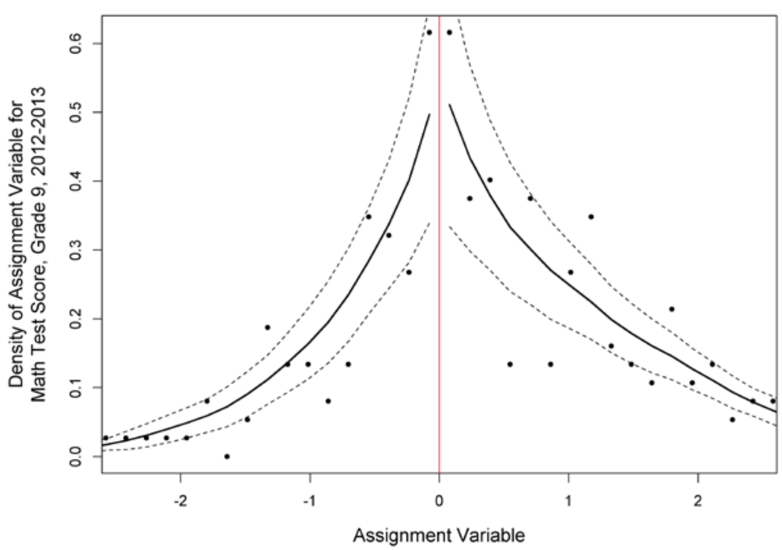
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.8. Density of the assignment variable for math test score in grade 9, place-based analysis



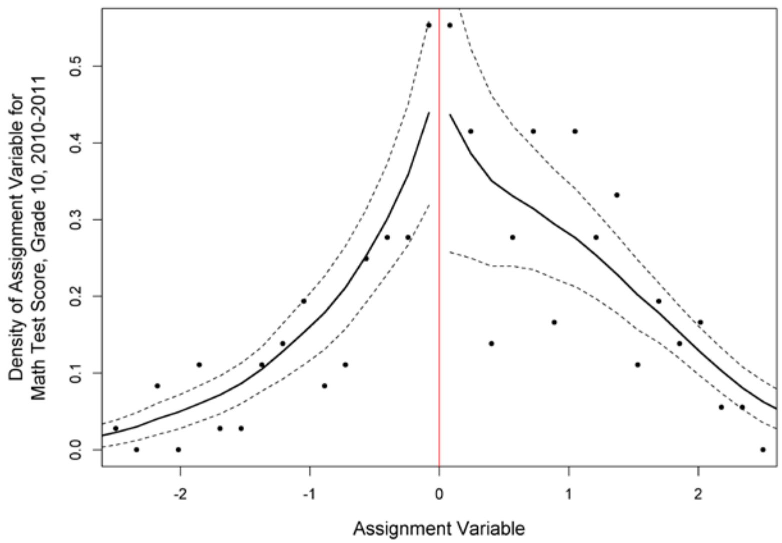


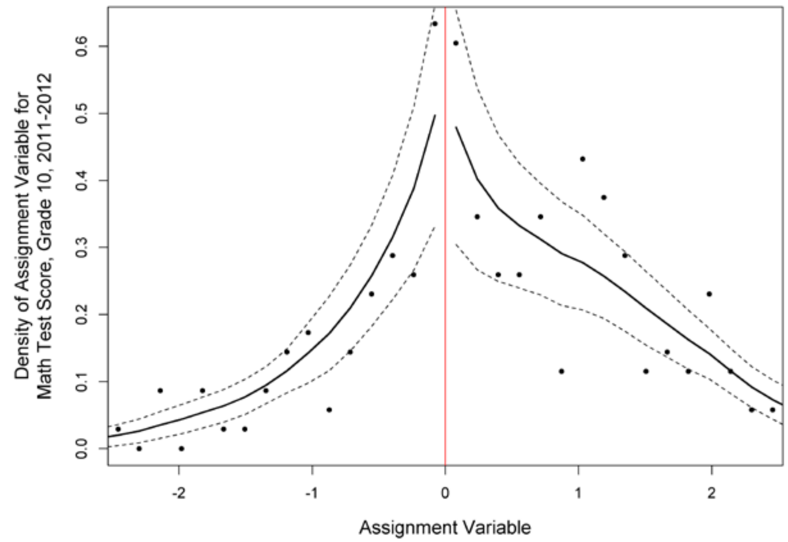


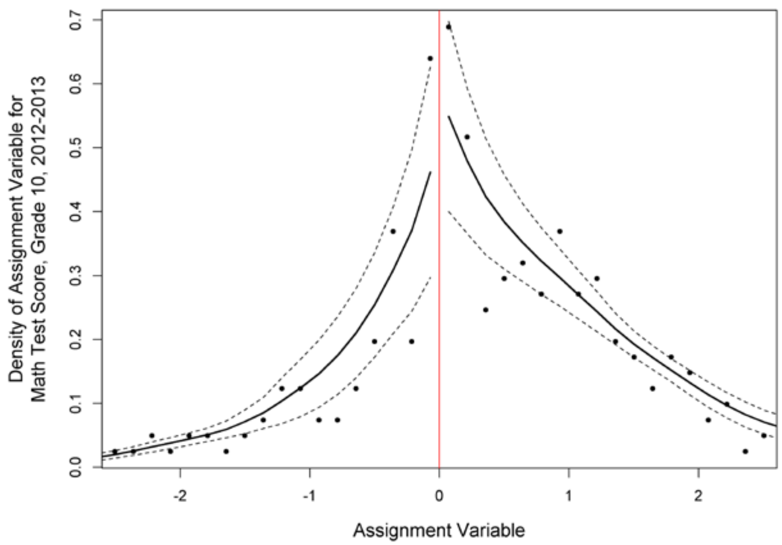
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.9. Density of the assignment variable for math test score in grade 10, place-based analysis



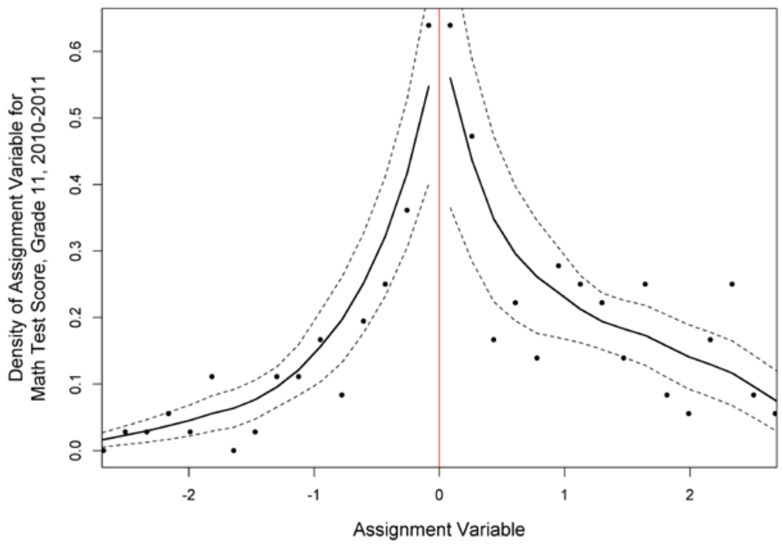


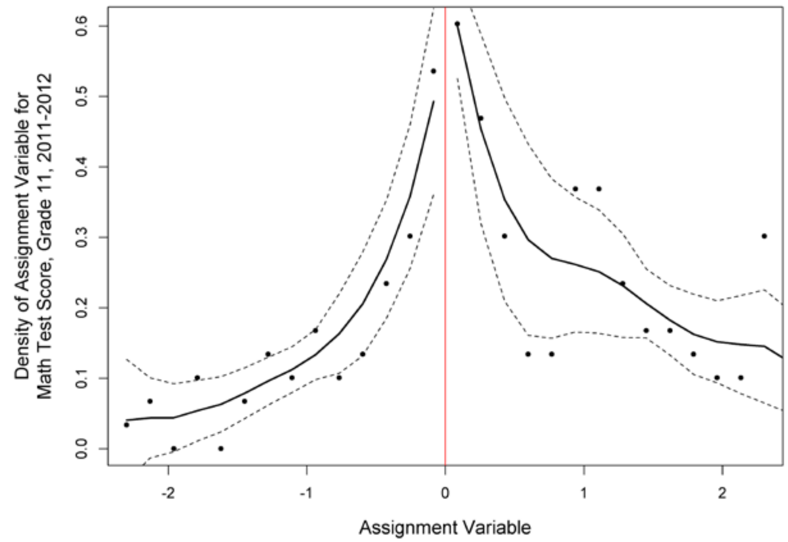


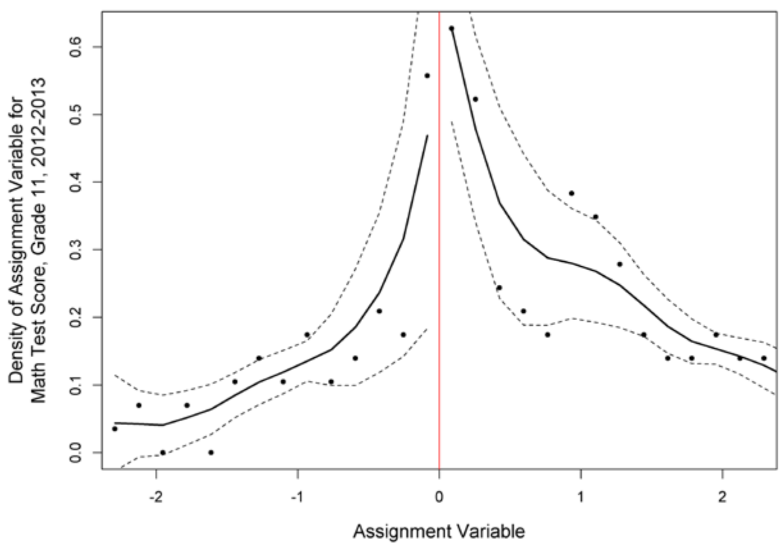
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.10. Density of the assignment variable for math test score in grade 11, place-based analysis



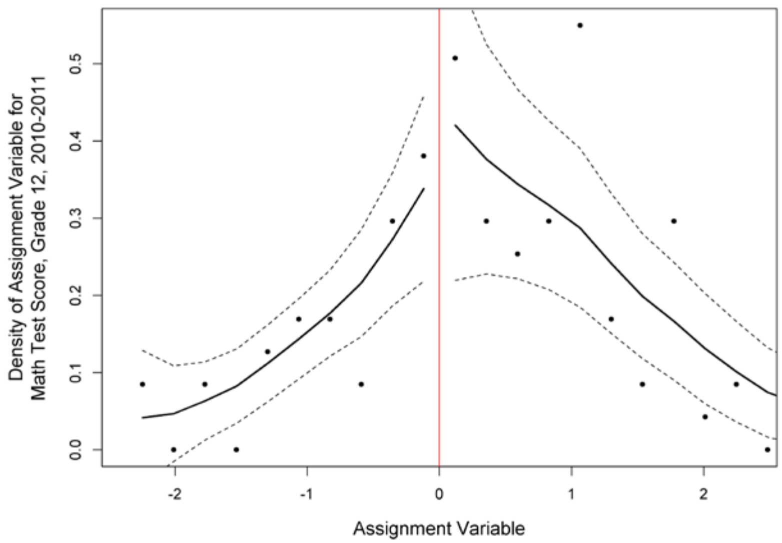


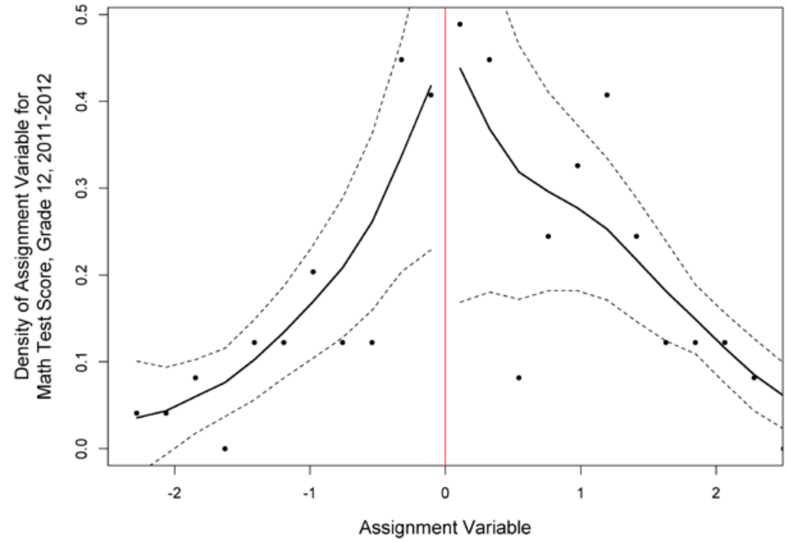


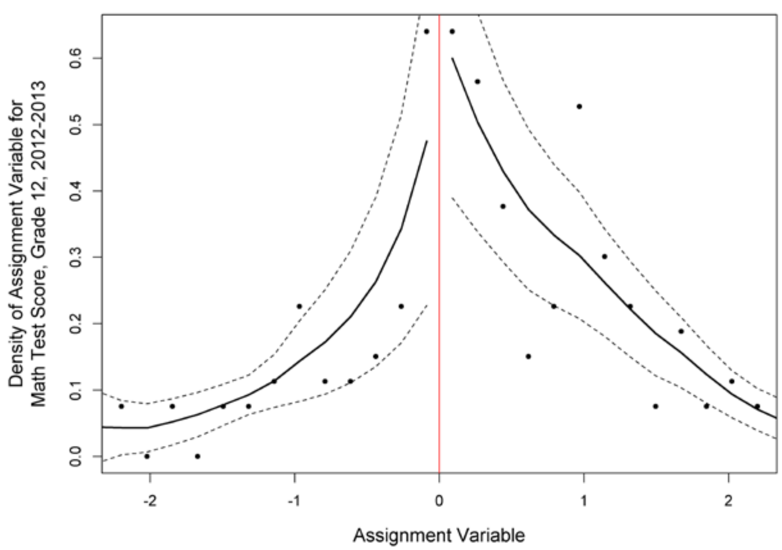
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.11. Density of the assignment variable for math test score in grade 12, place-based analysis



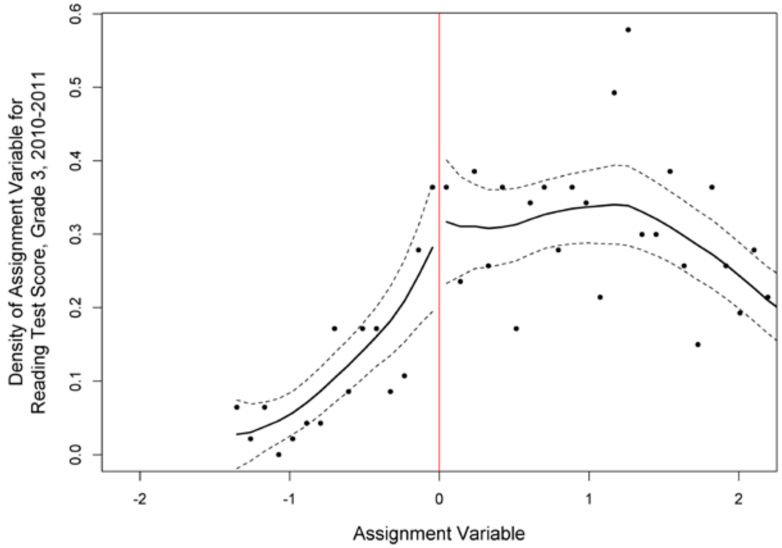


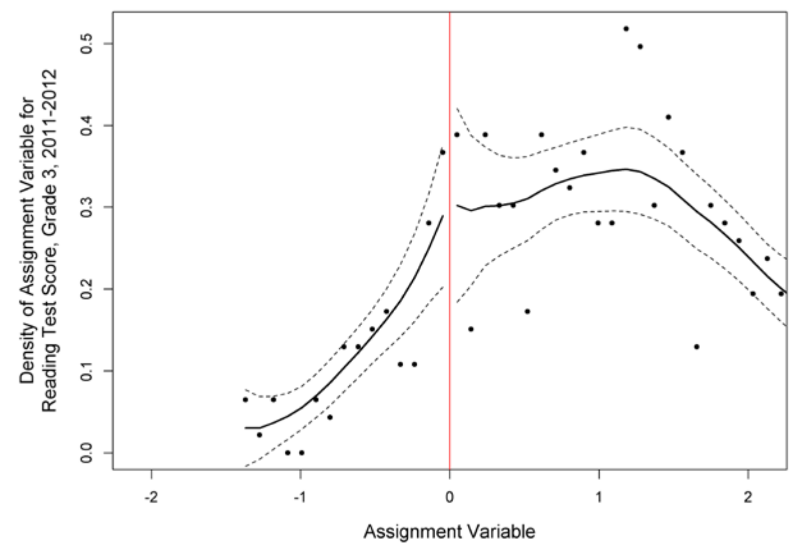


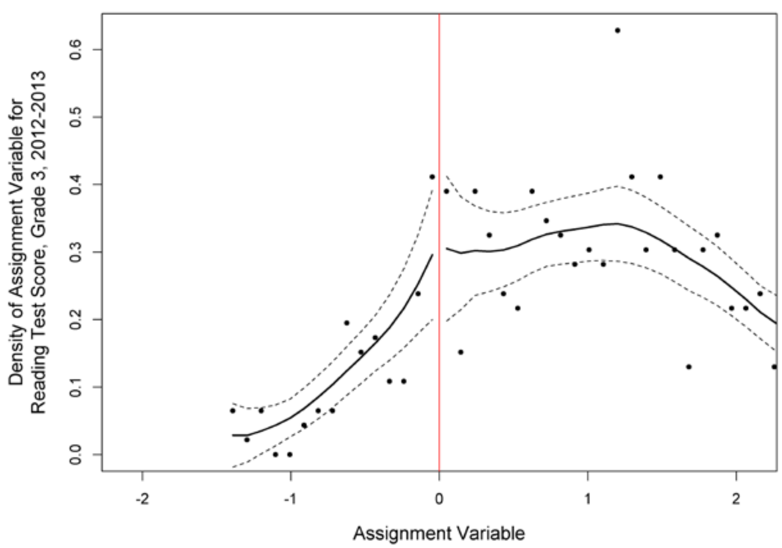
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.12. Density of the assignment variable for reading test score in grade 3, place-based analysis



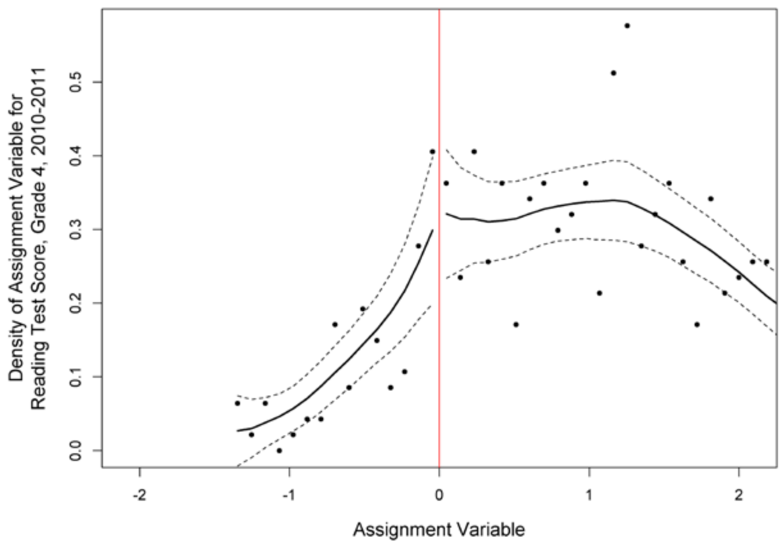


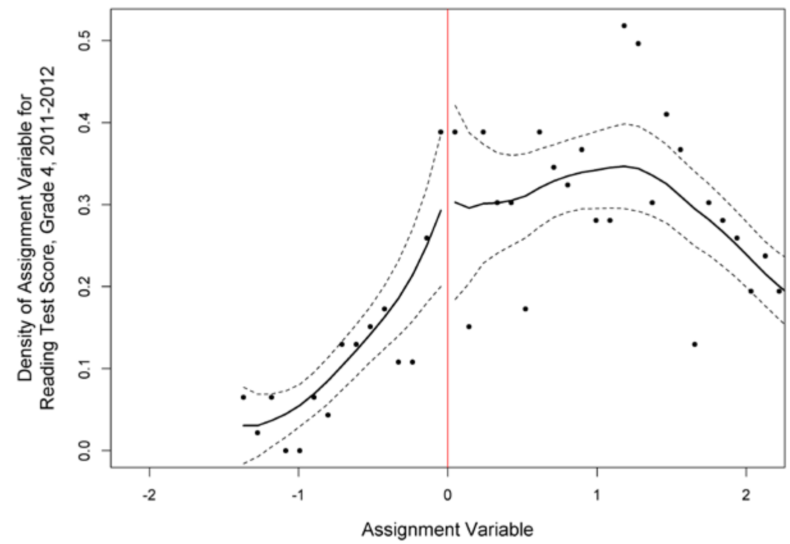


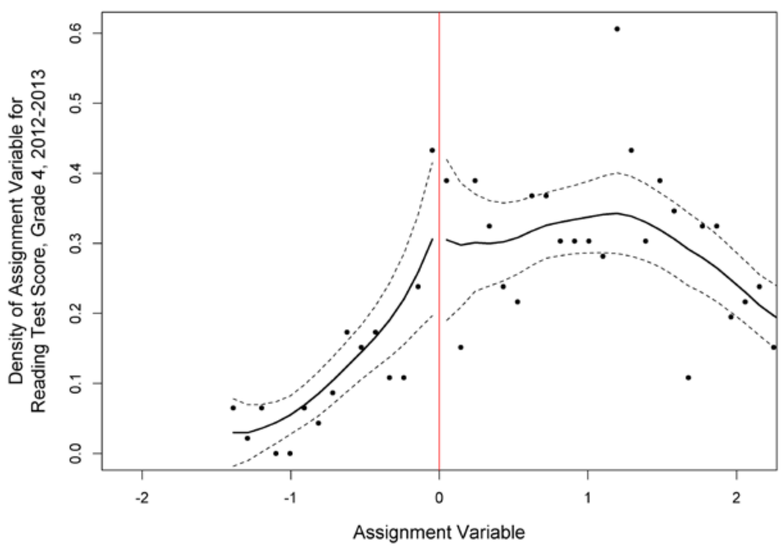
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.13. Density of the assignment variable for reading test score in grade 4, place-based analysis



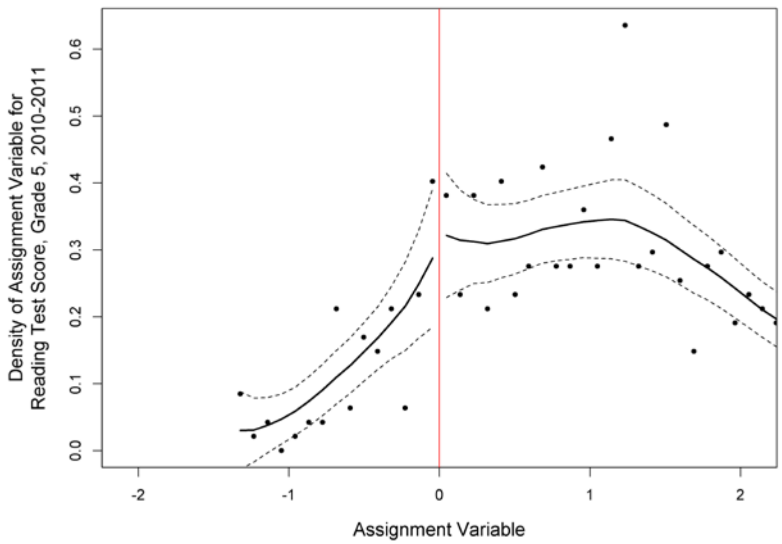


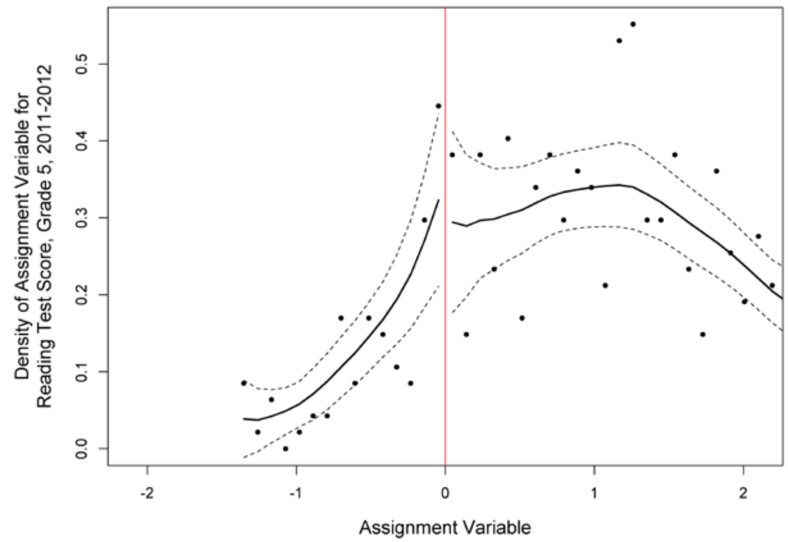


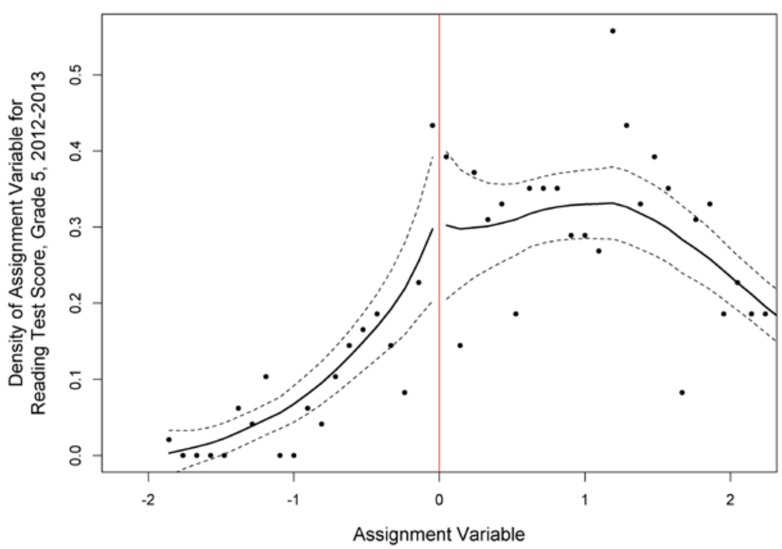
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.14. Density of the assignment variable for reading test score in grade 5, place-based analysis



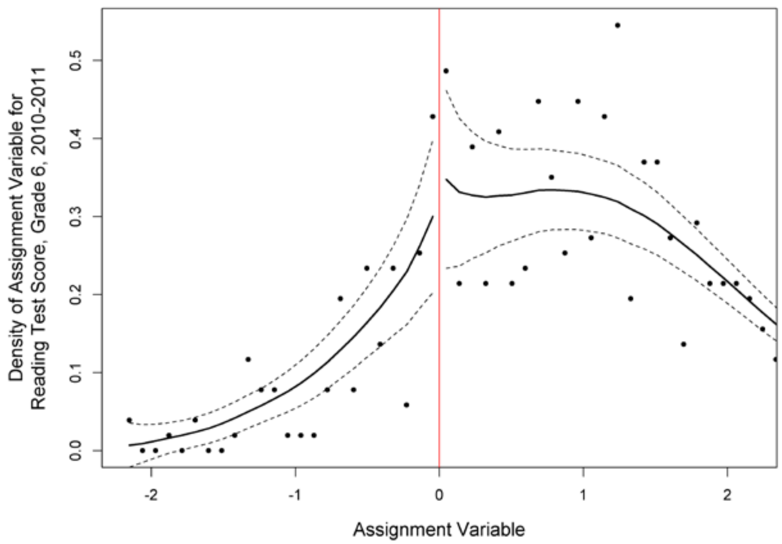


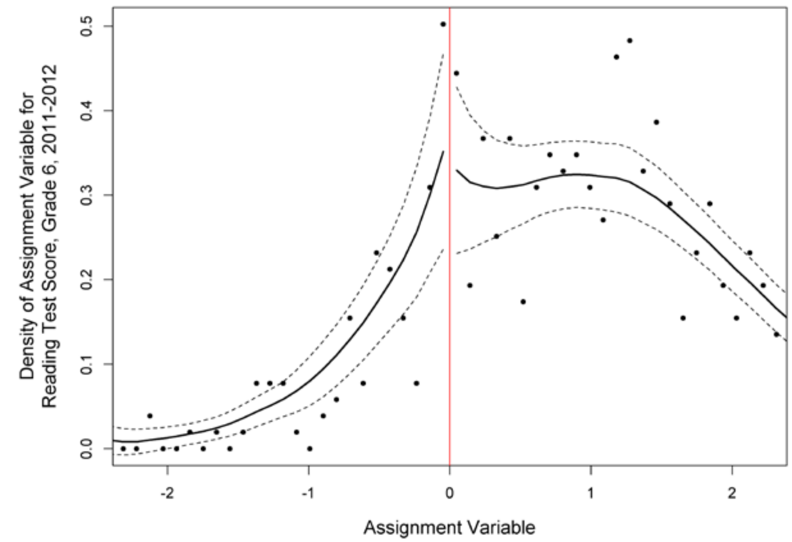


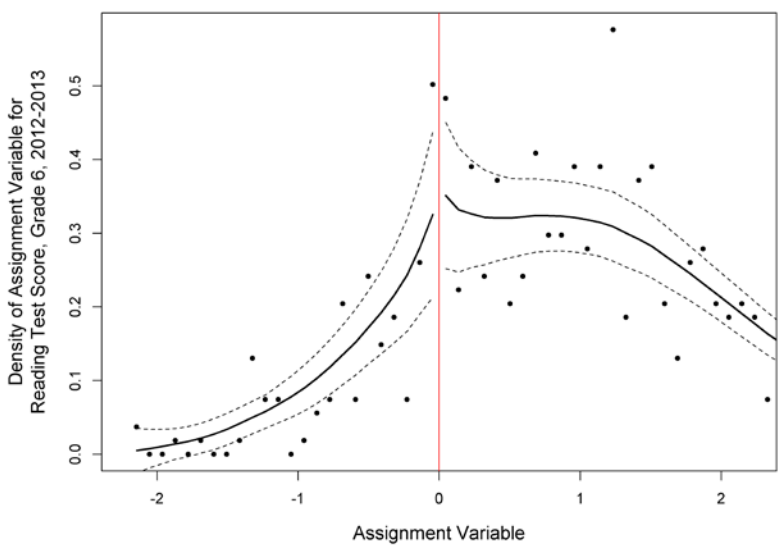
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.15. Density of the assignment variable for reading test score in grade 6, place-based analysis



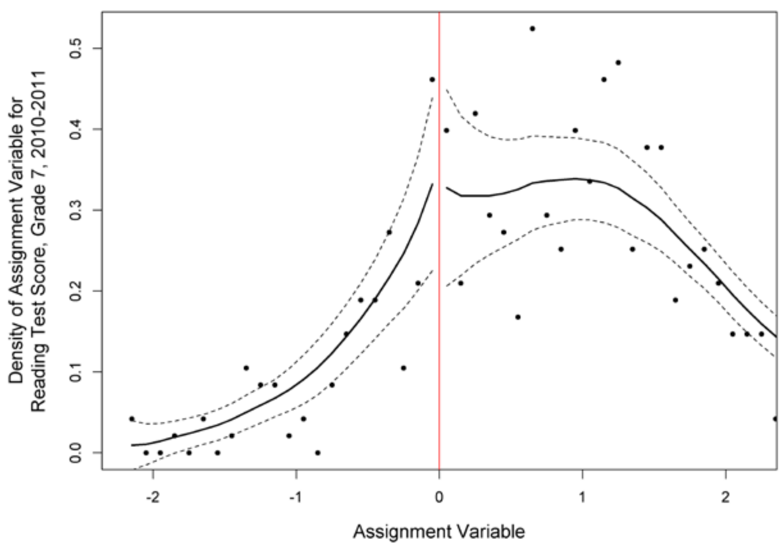


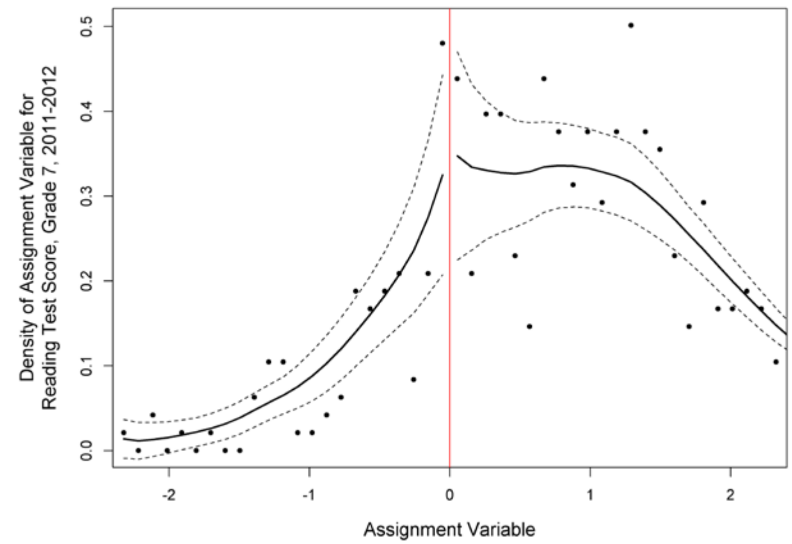


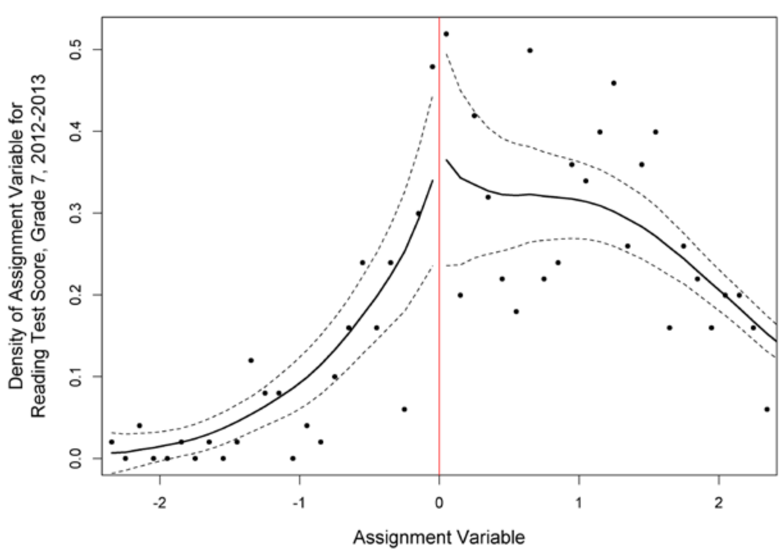
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.16. Density of the assignment variable for reading test score in grade 7, place-based analysis



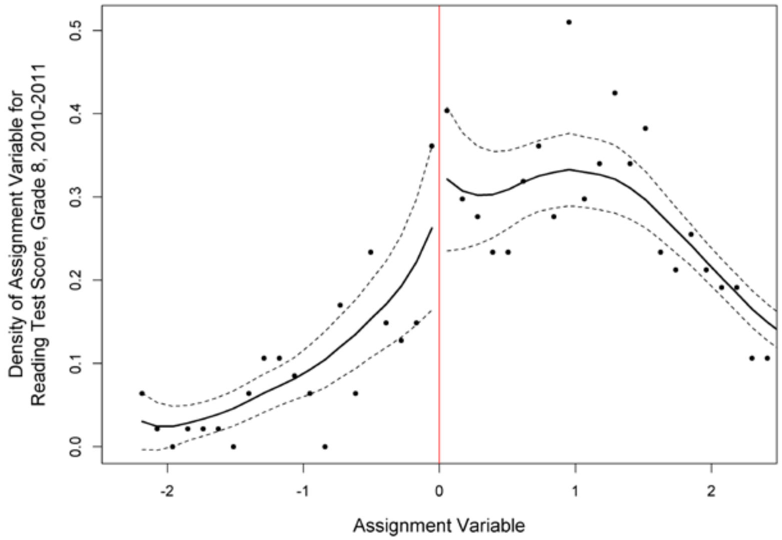


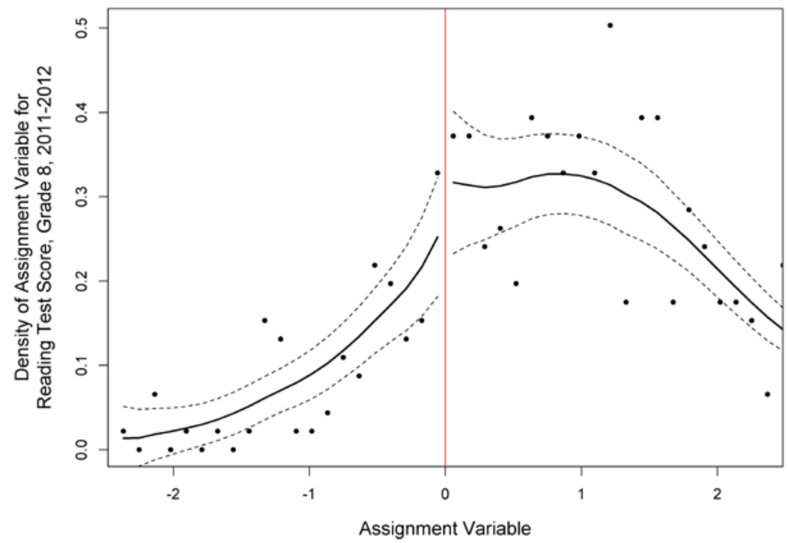


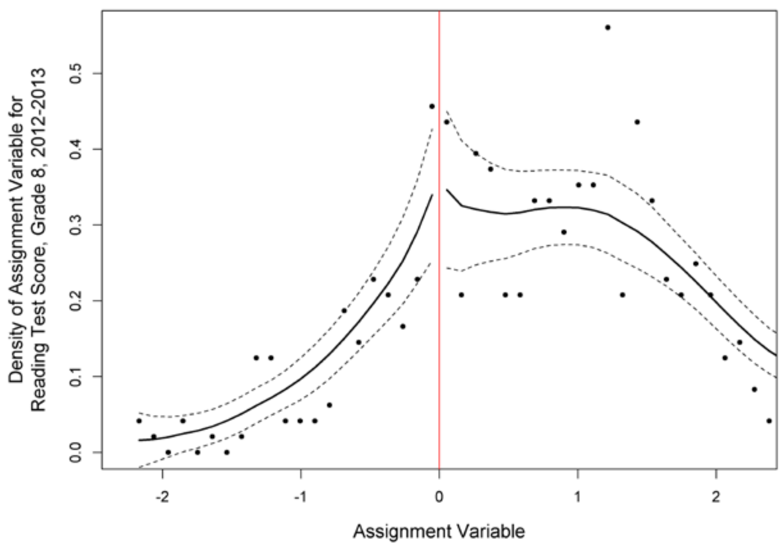
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.17. Density of the assignment variable for reading test score in grade 8, place-based analysis



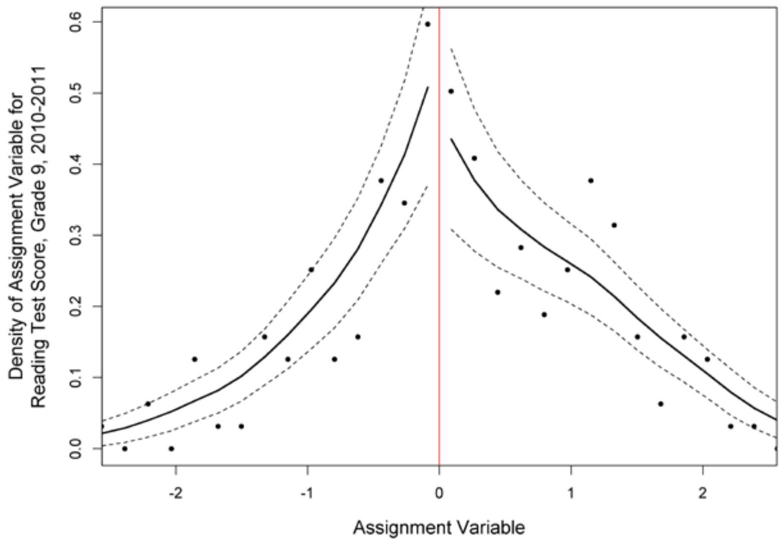


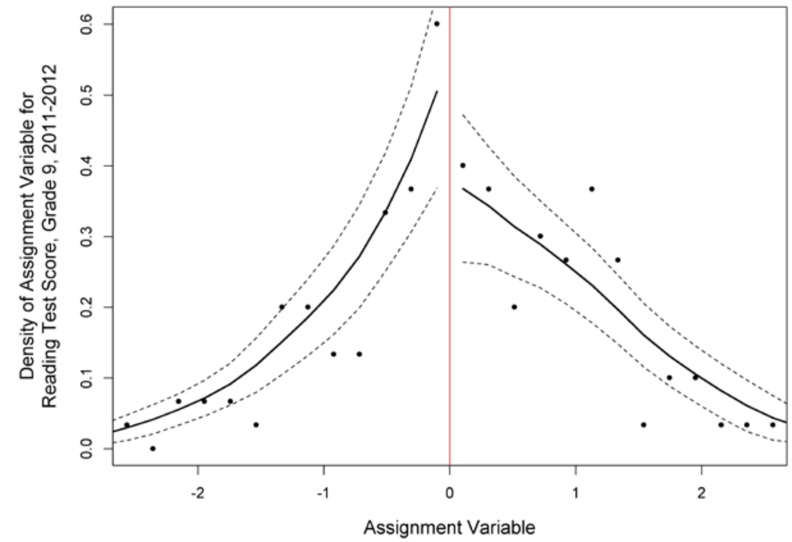


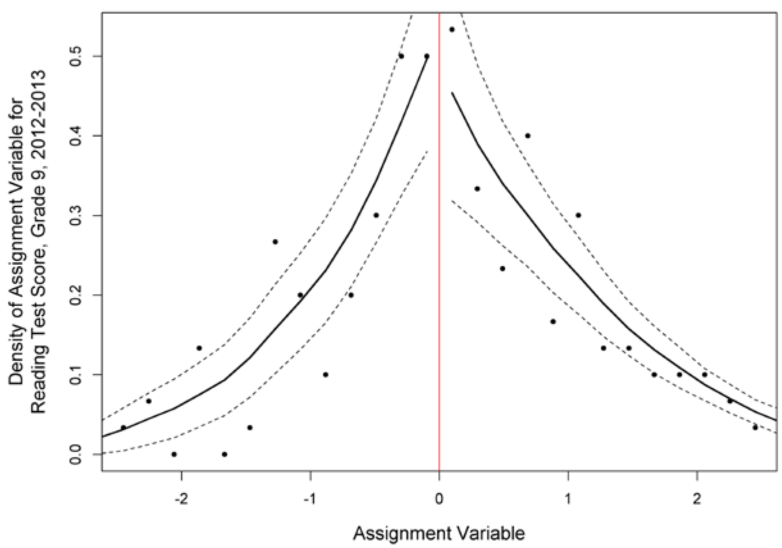
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.18. Density of the assignment variable for reading test score in grade 9, place-based analysis



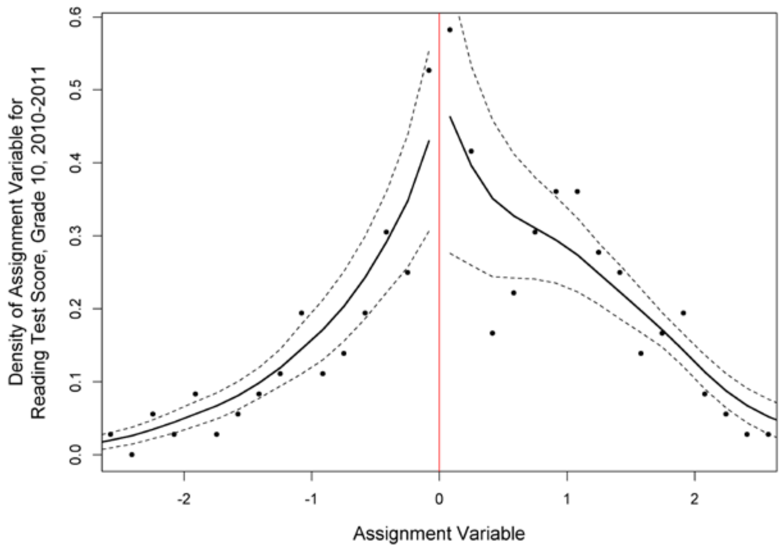


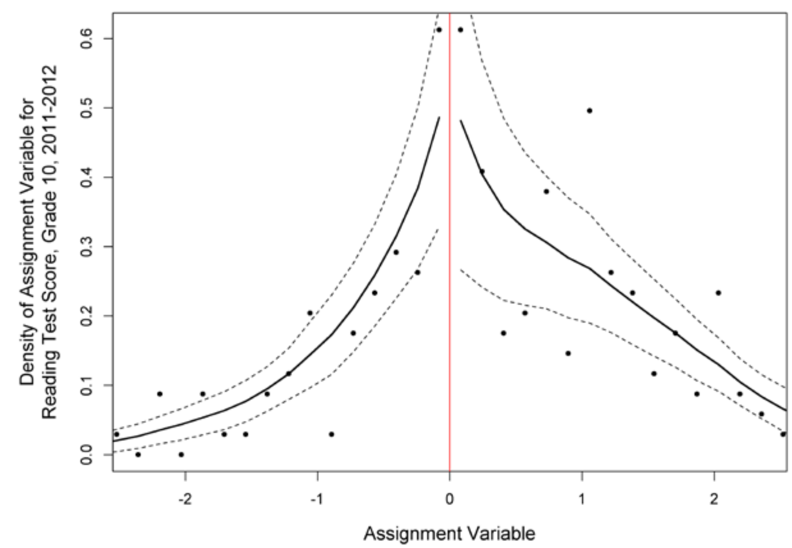


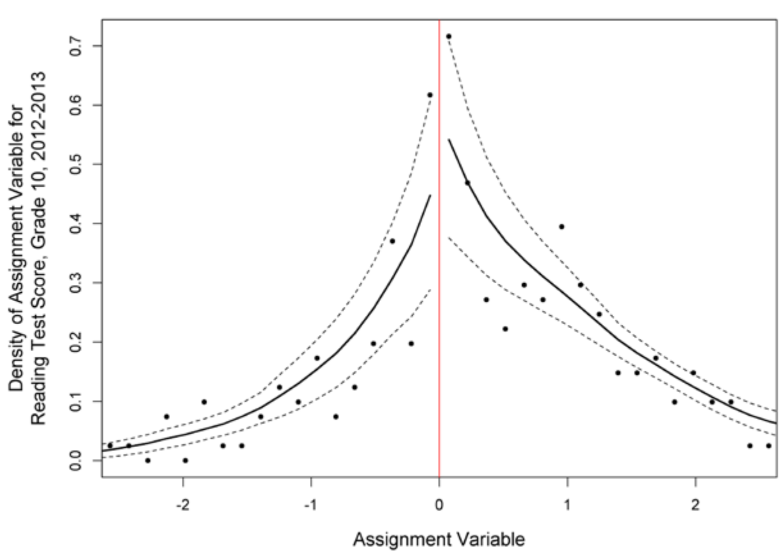
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.19. Density of the assignment variable for reading test score in grade 10, place-based analysis



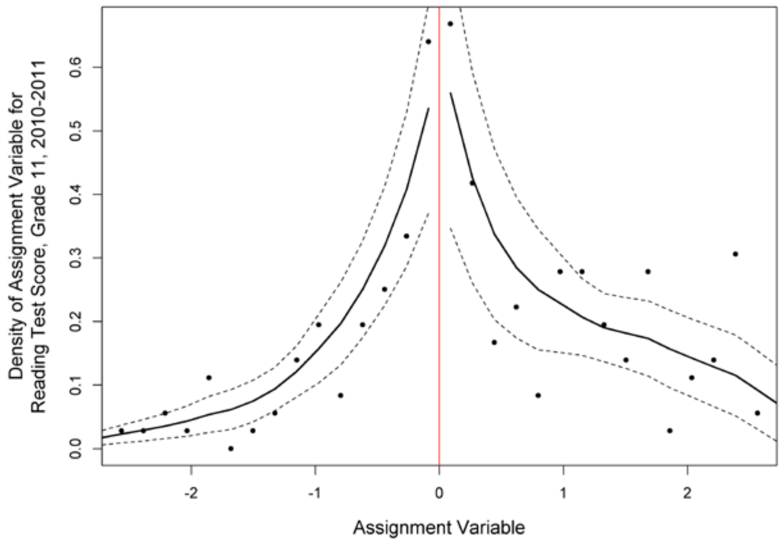


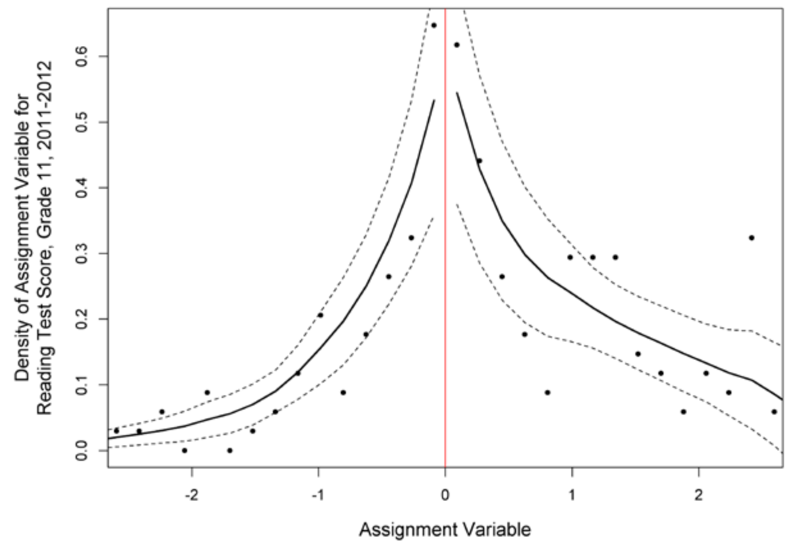


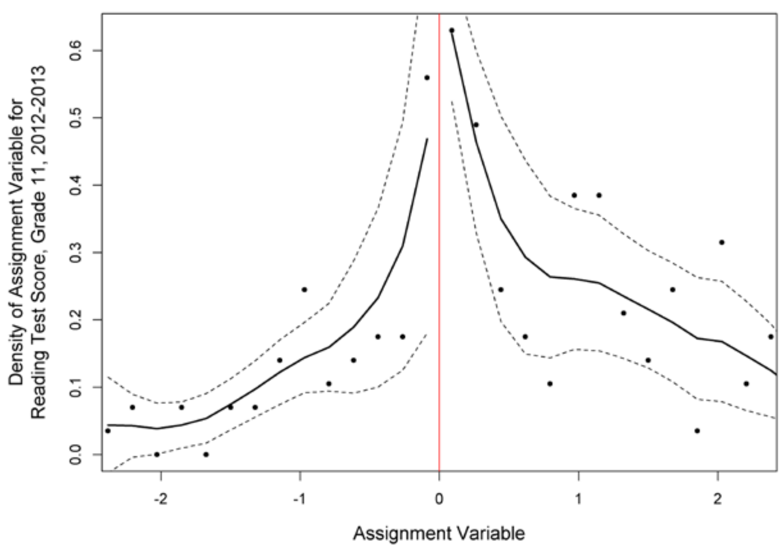
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.20. Density of the assignment variable for reading test score in grade 11, place-based analysis



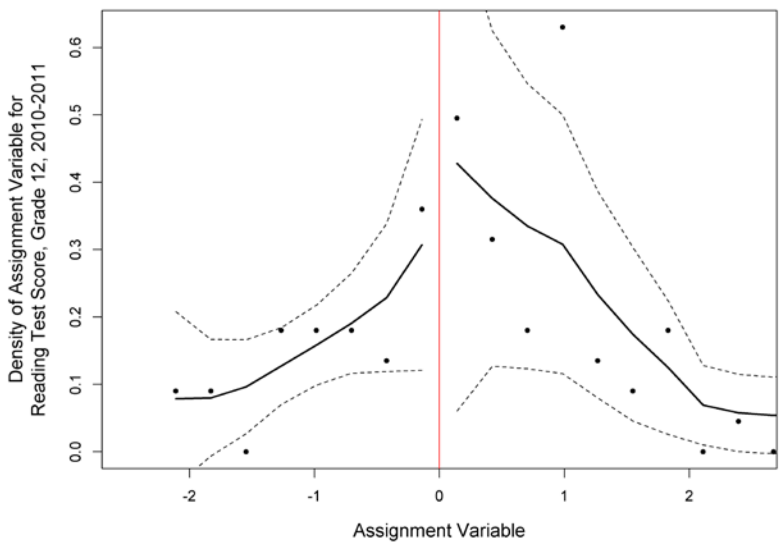


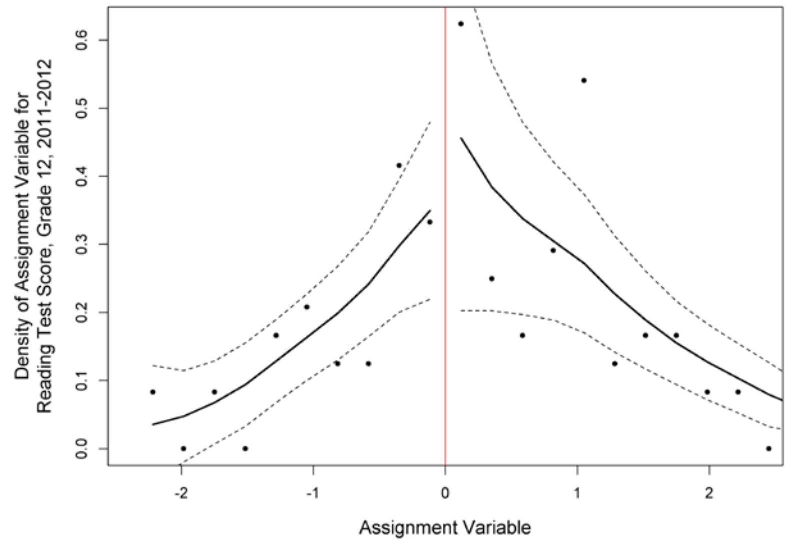


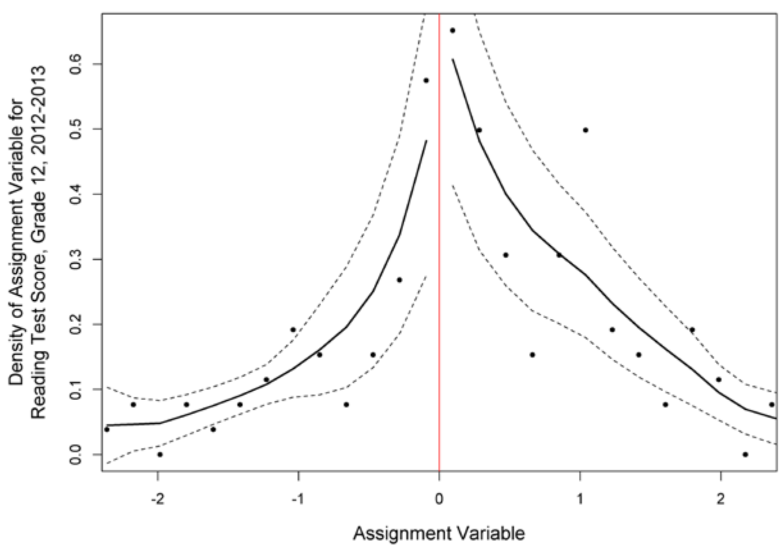
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.21. Density of the assignment variable for reading test score in grade 12, place-based analysis



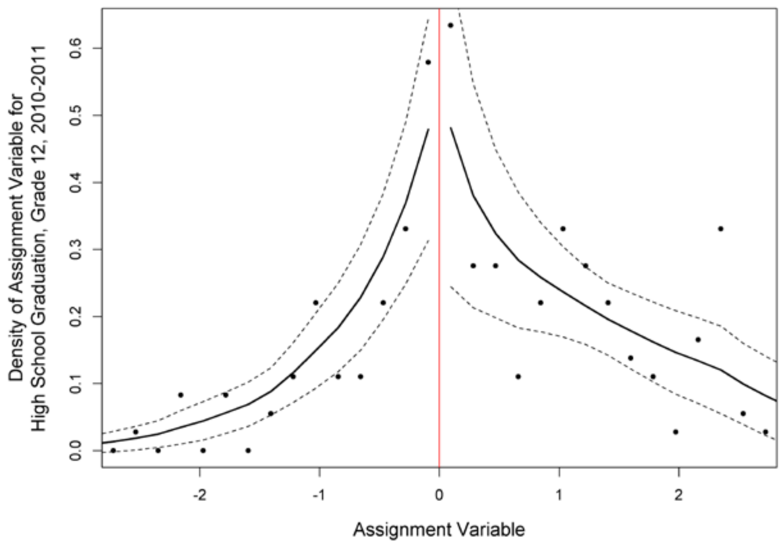


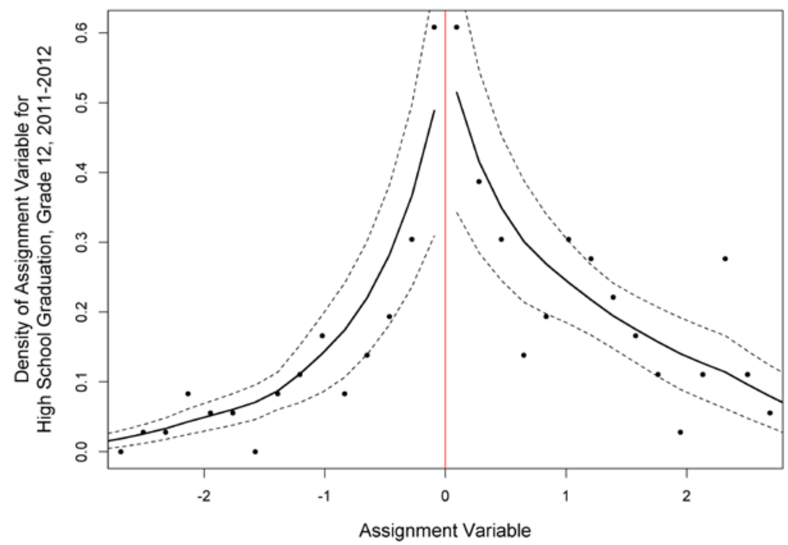


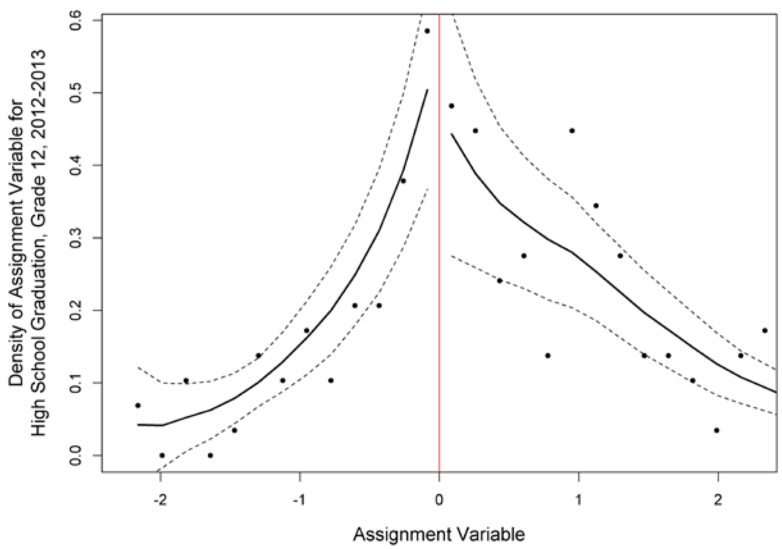
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.22. Density of the assignment variable for high school graduation, place-based analysis



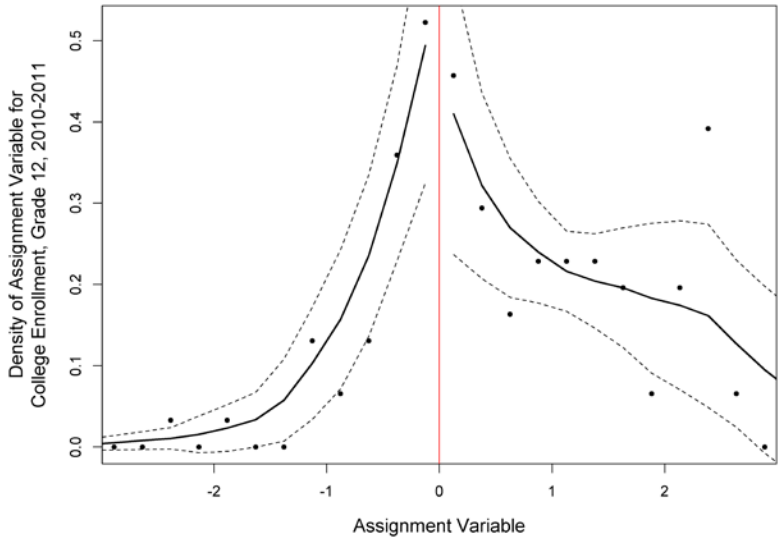


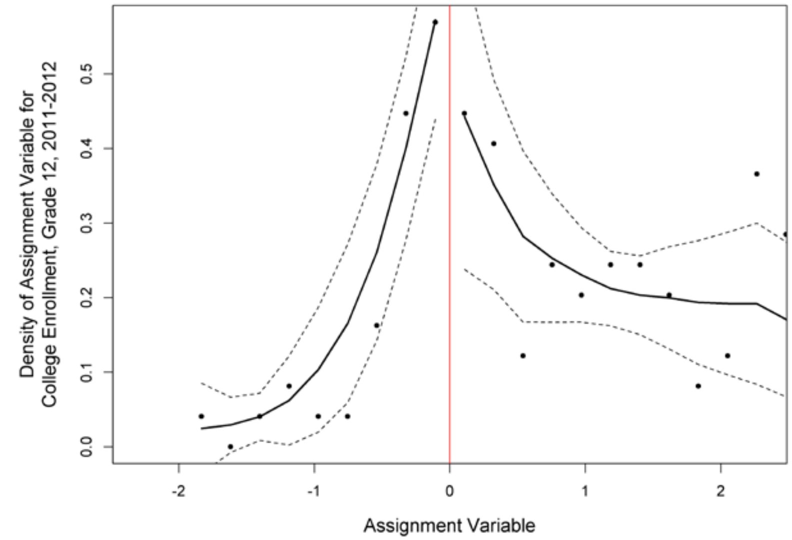


Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.23. Density of the assignment variable for college enrollment, place-based analysis

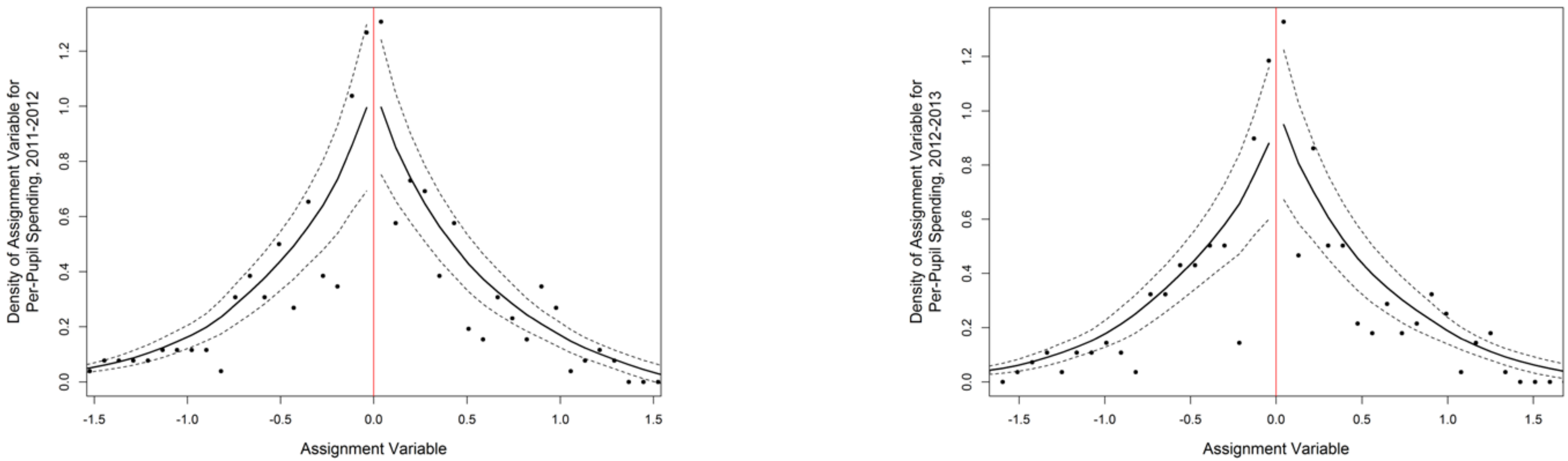




Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines). We omitted the figure for 2012–2013 because we were not able to estimate an impact for college enrollment in this year due to insufficient sample sizes.

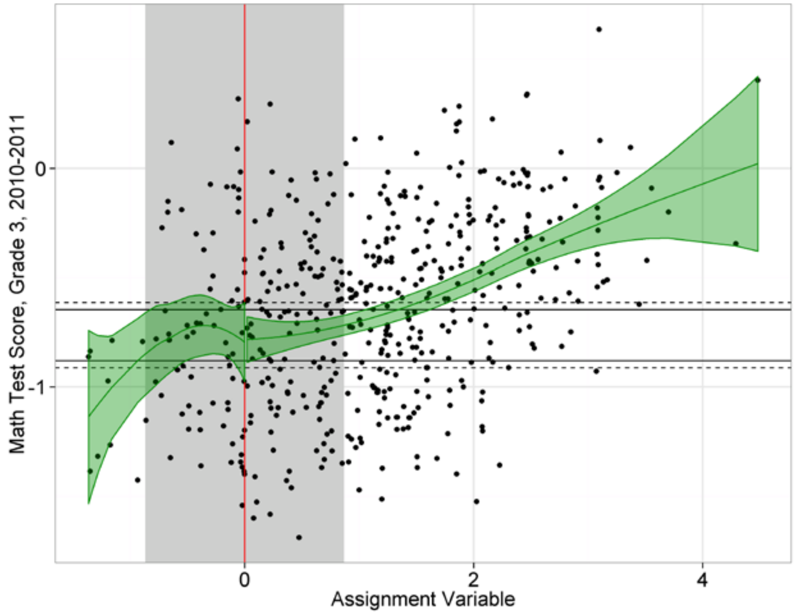
Figure A.24. Density of the assignment variable for per-pupil spending, place-based analysis

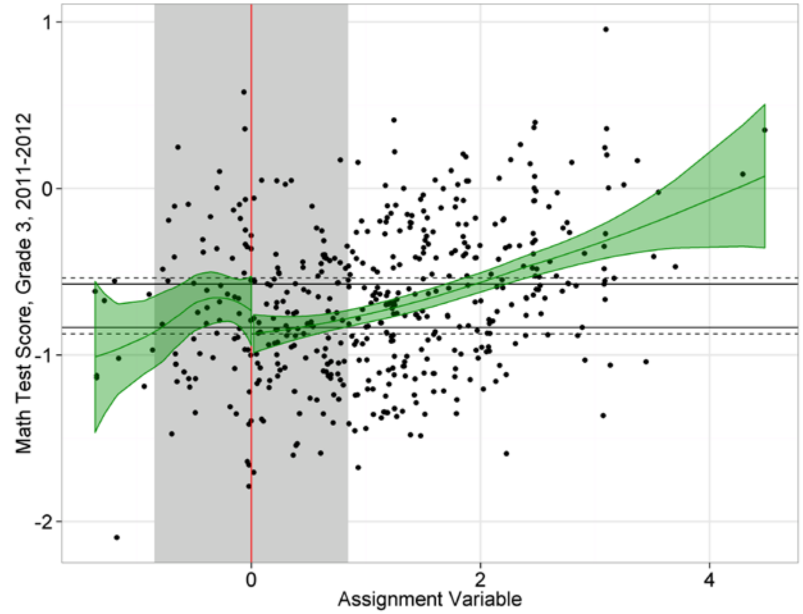


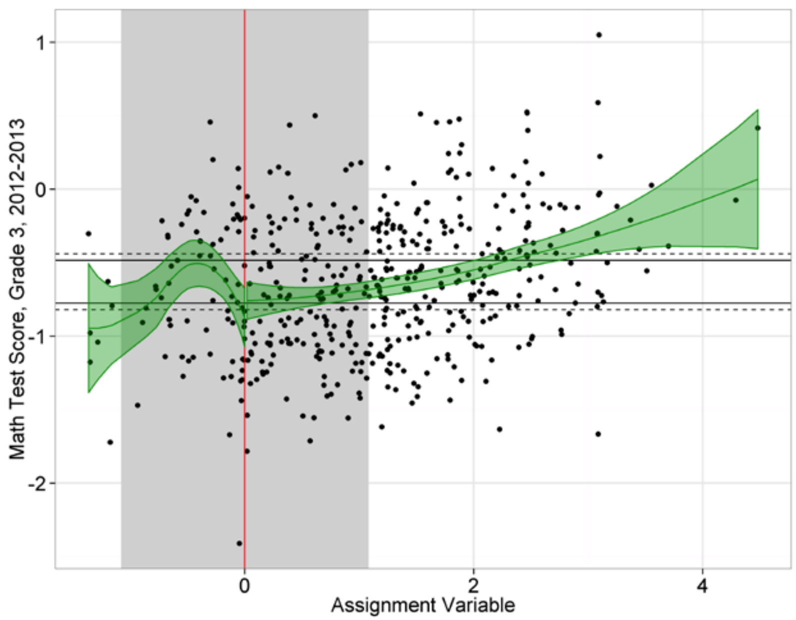
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.25. Math test score in grade 3, place-based analysis



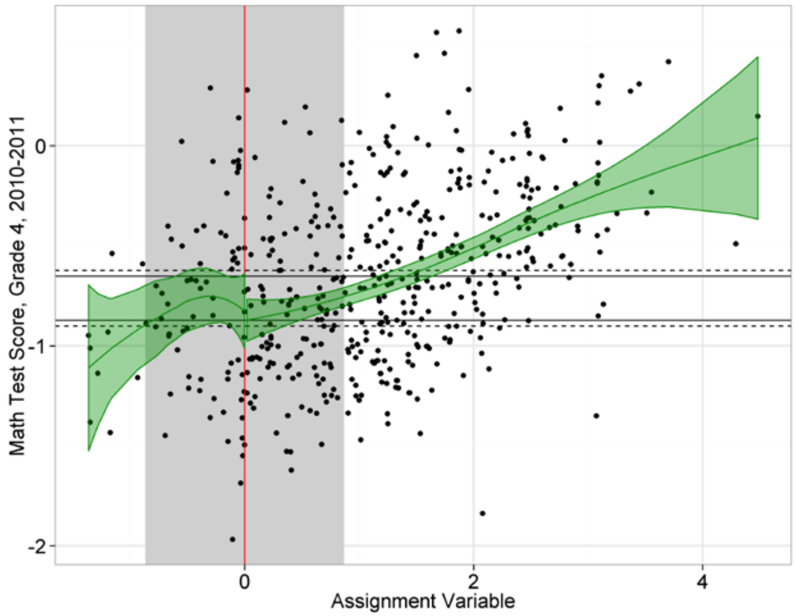


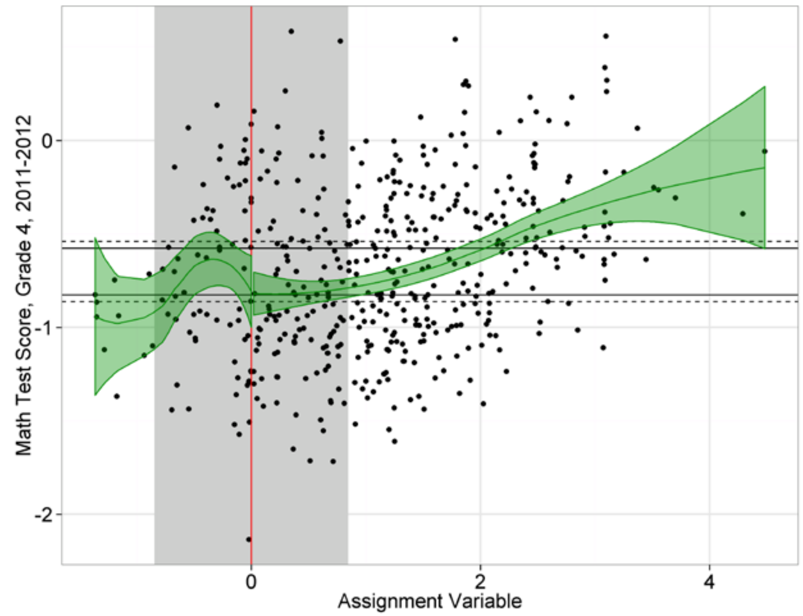


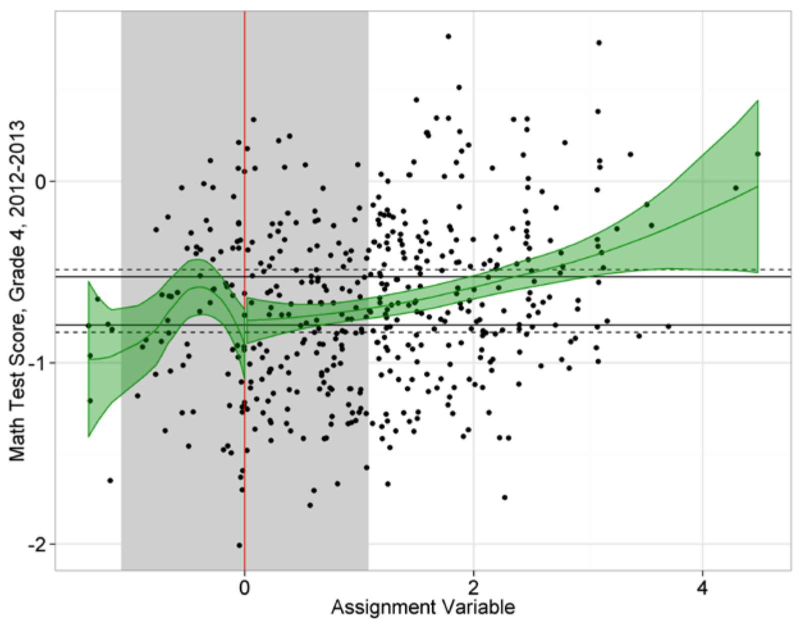
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.26. Math test score in grade 4, place-based analysis



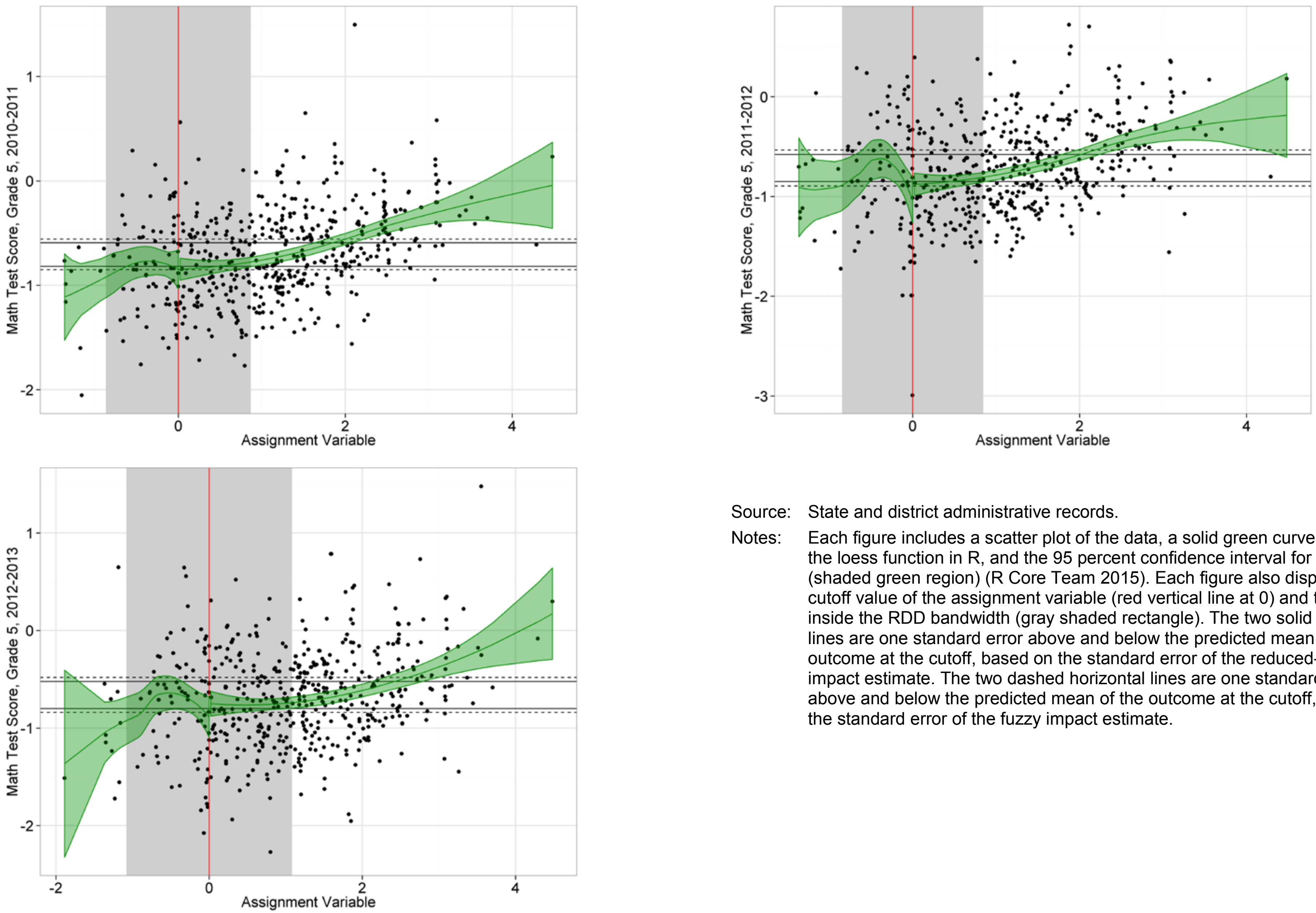




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.27. Math test score in grade 5, place-based analysis

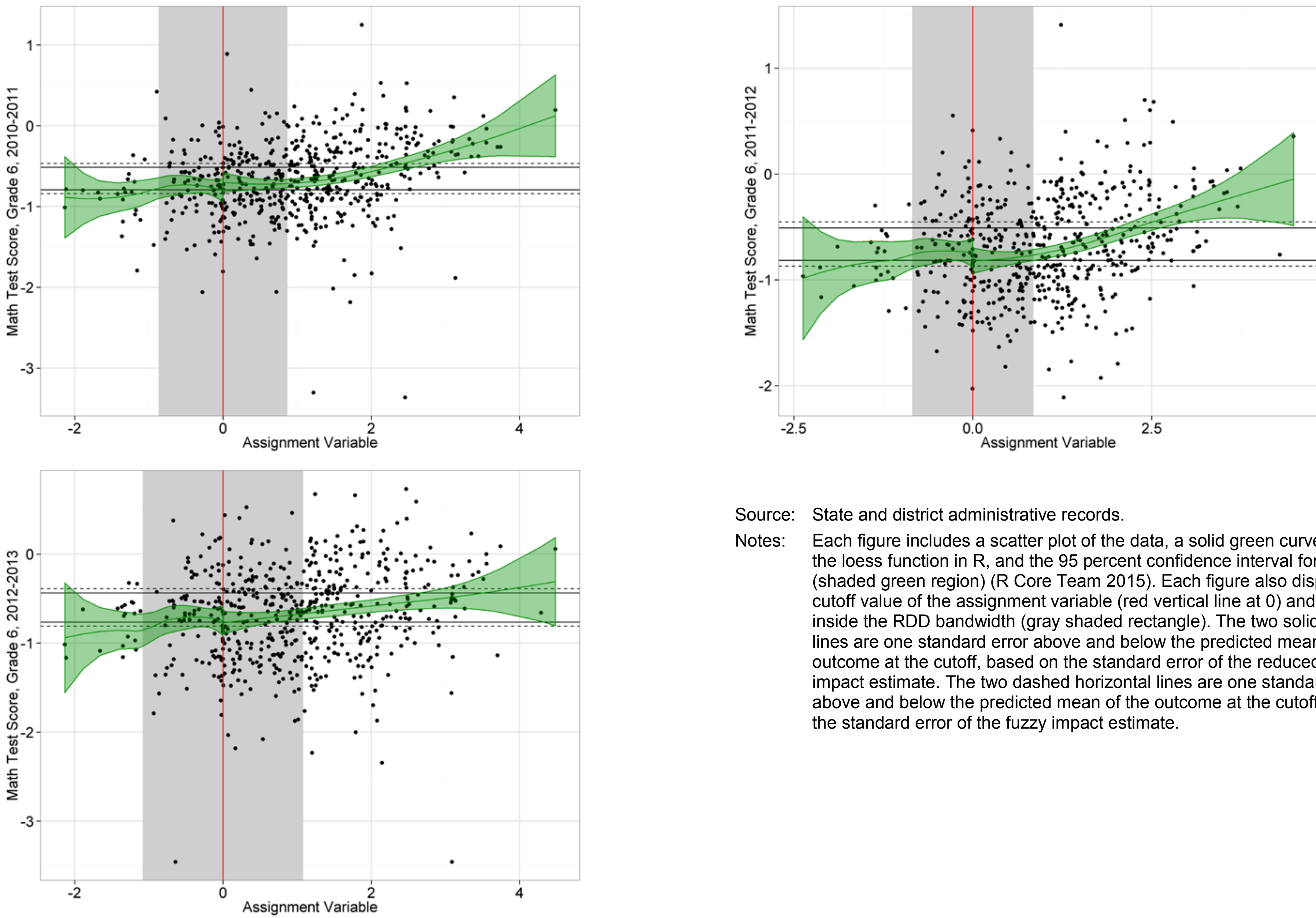


Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

.

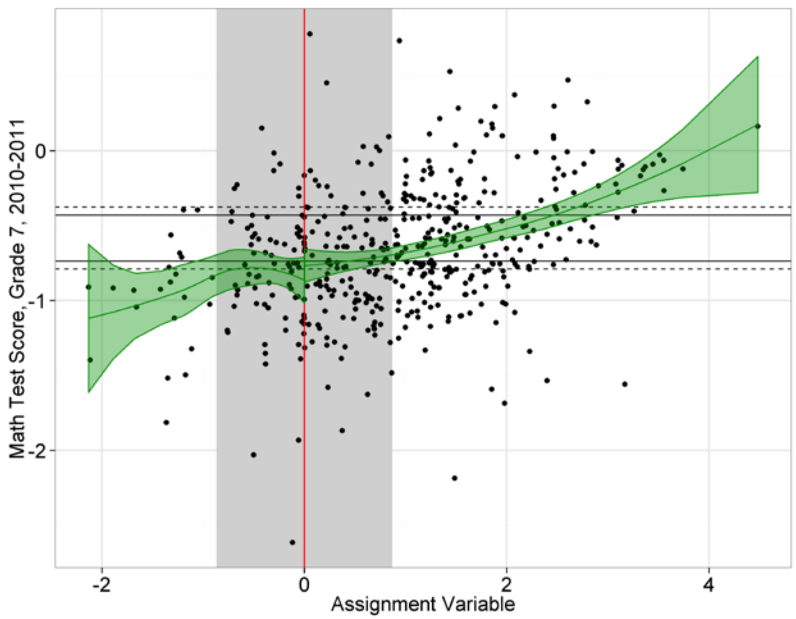
Figure A.28. Math test score in grade 6, place-based analysis

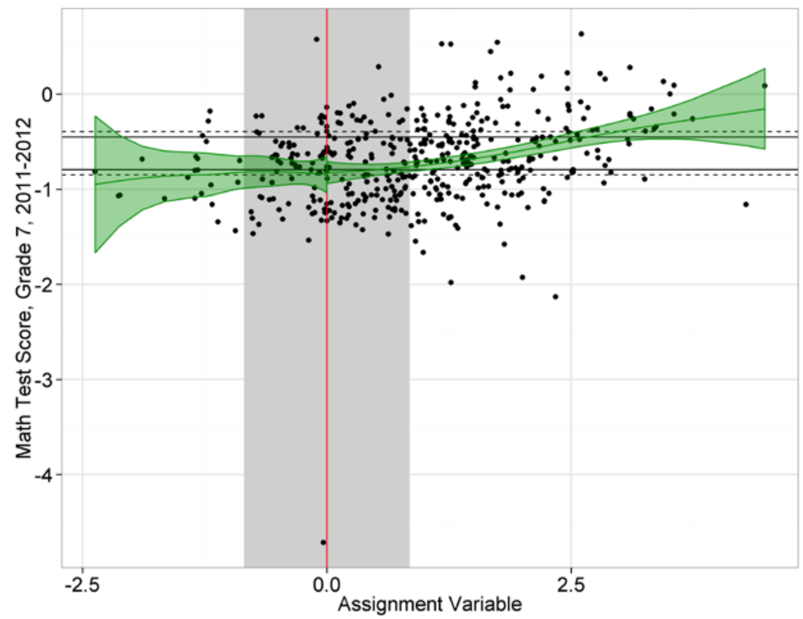


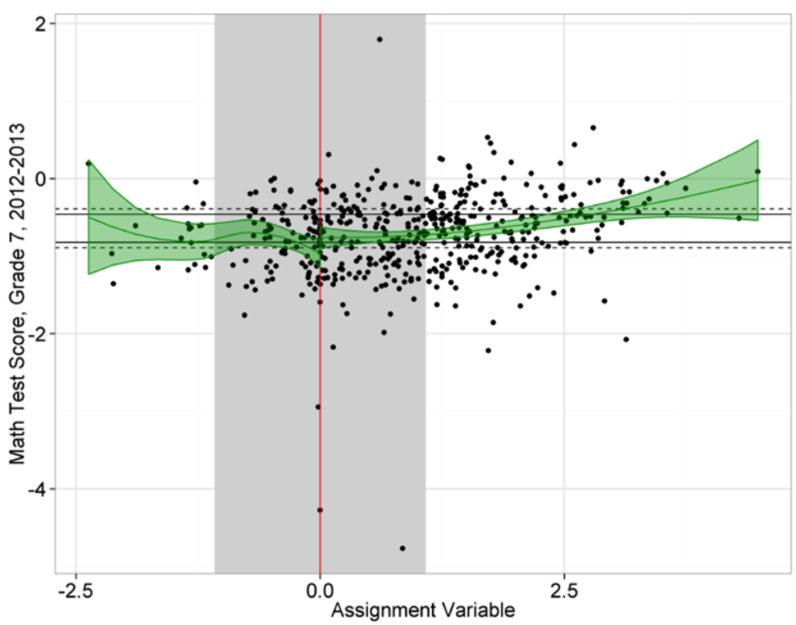
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.29. Math test score in grade 7, place-based analysis



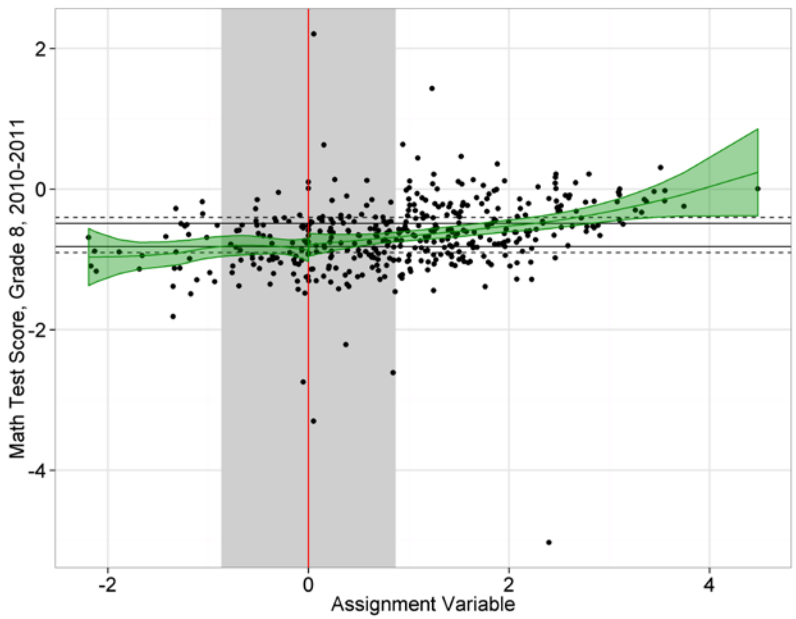


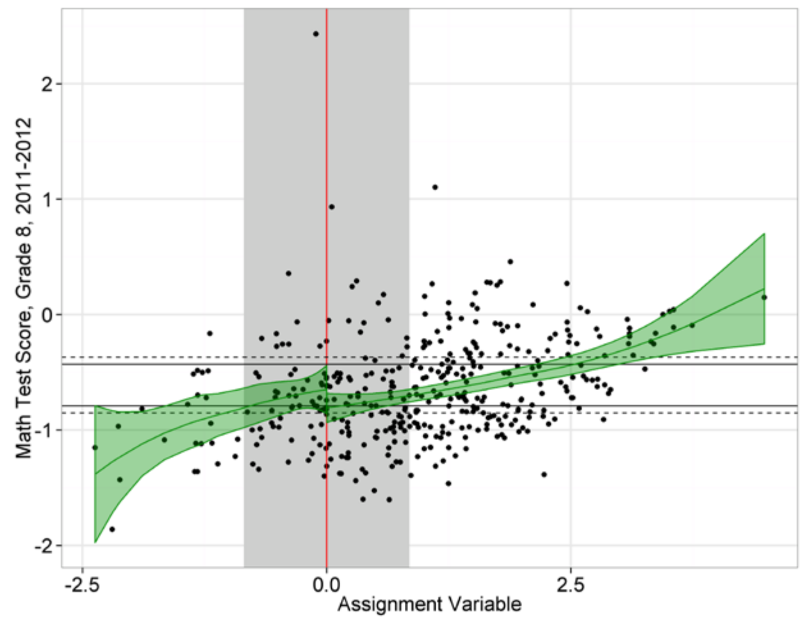


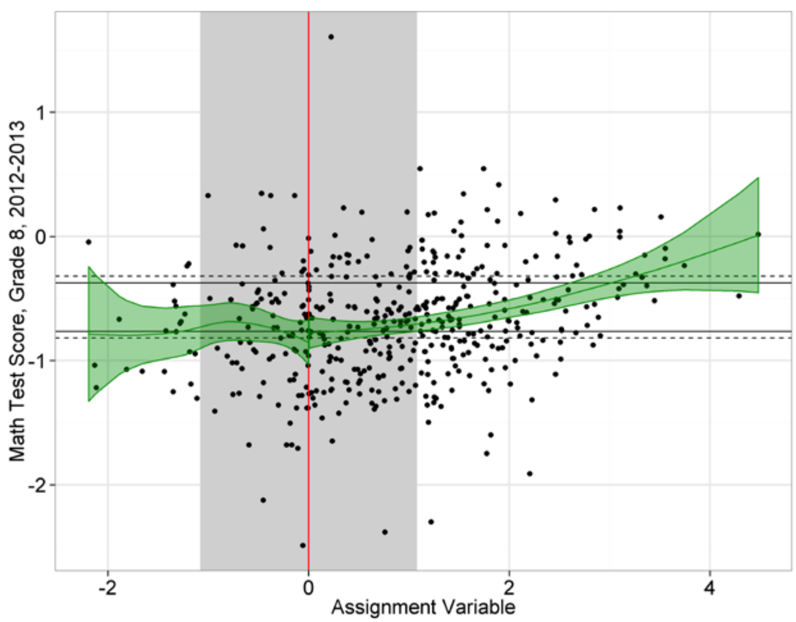
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.30. Math test score in grade 8, place-based analysis





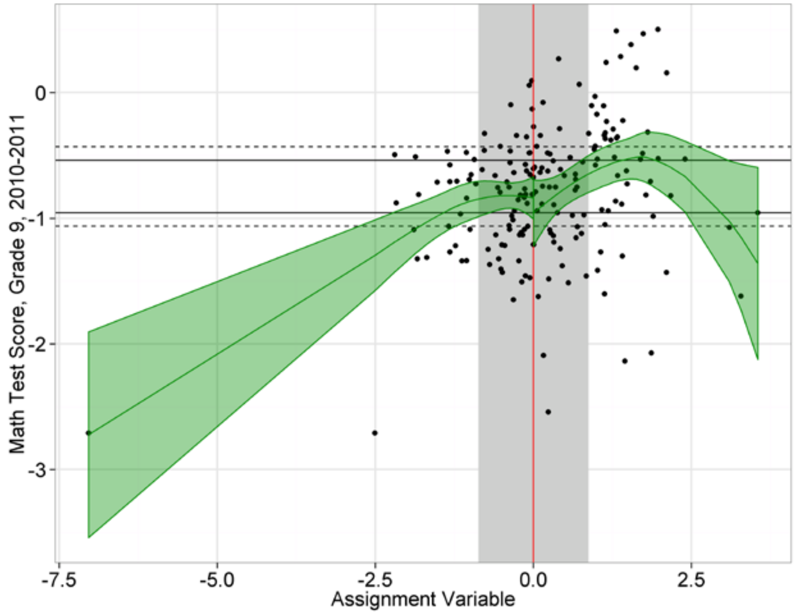


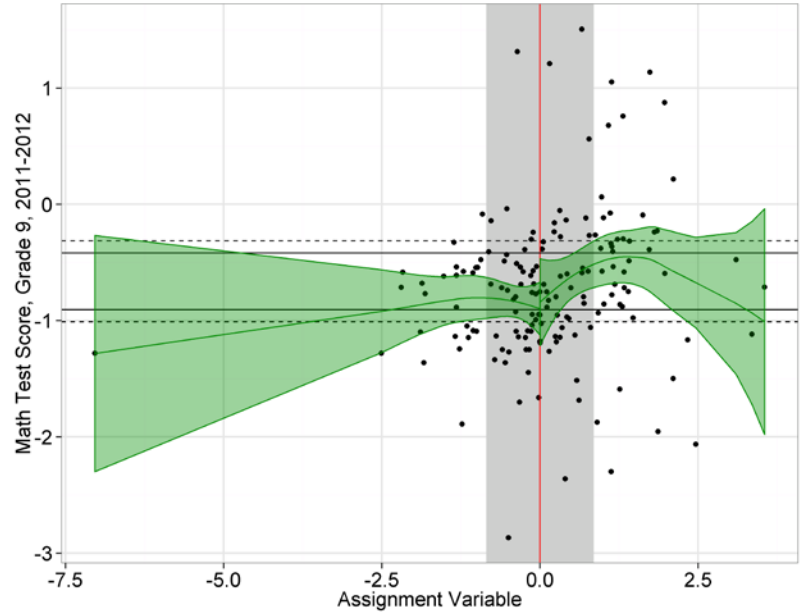
Source: State and district administrative records.

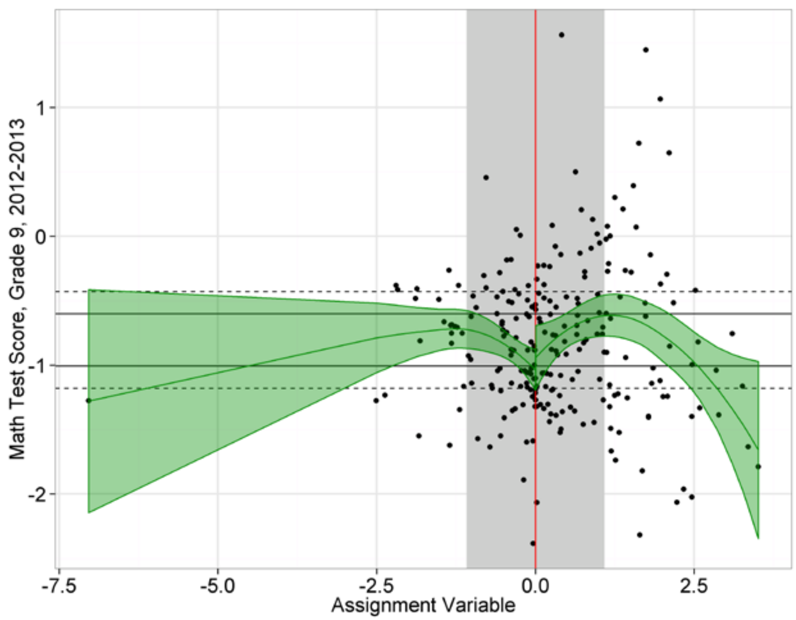
Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

.

Figure A.31. Math test score in grade 9, place-based analysis



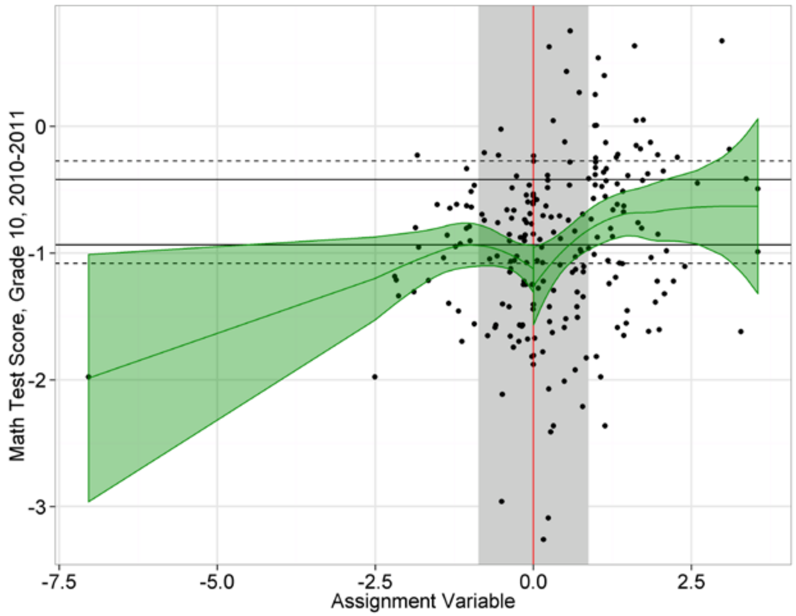


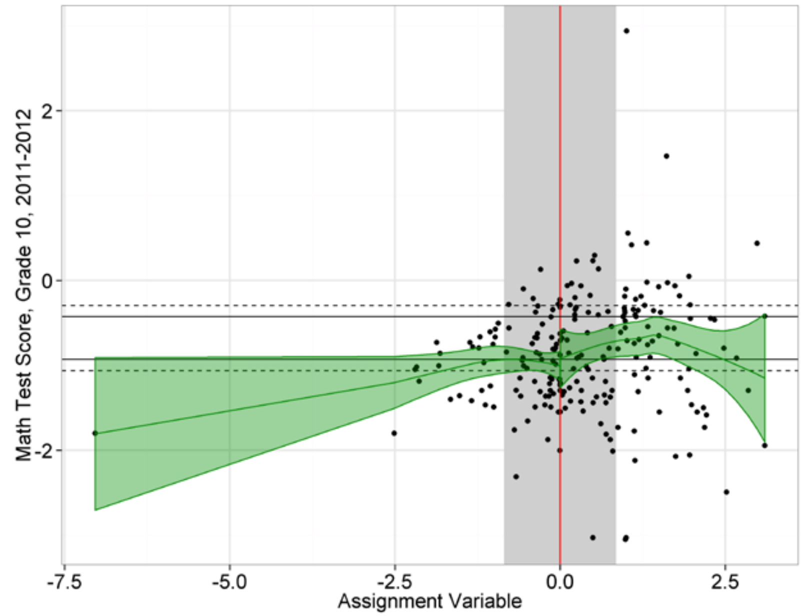


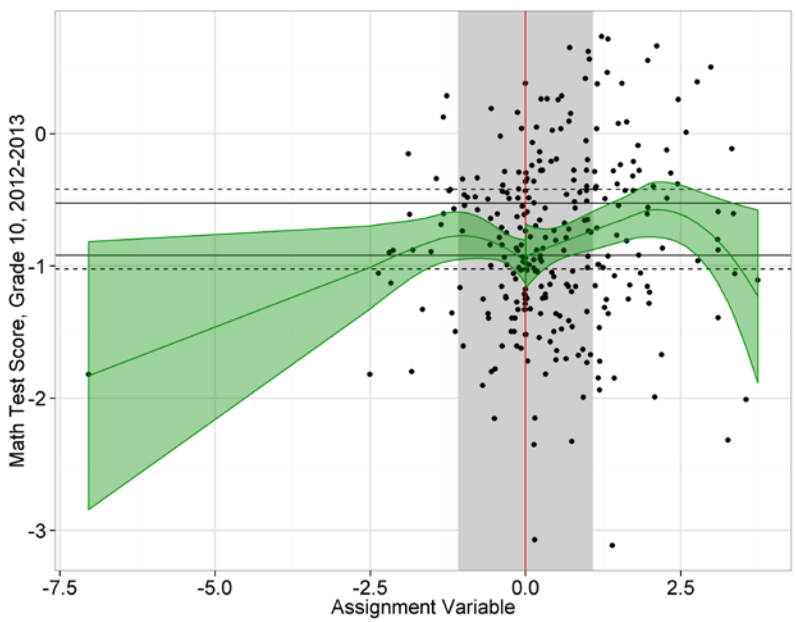
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.32. Math test score in grade 10, place-based analysis



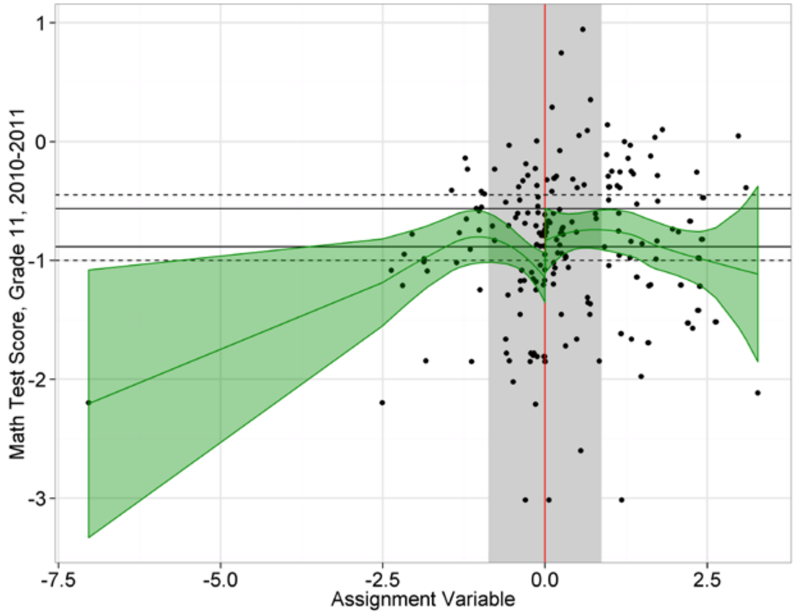


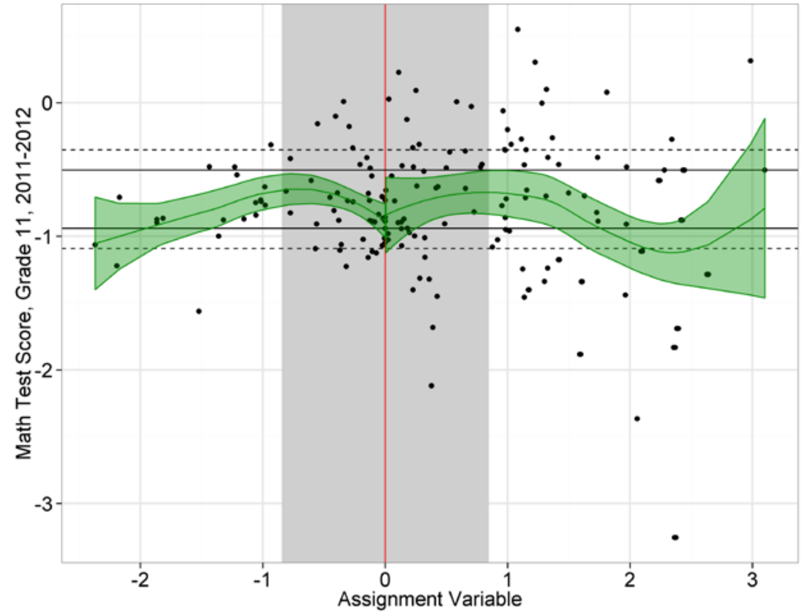


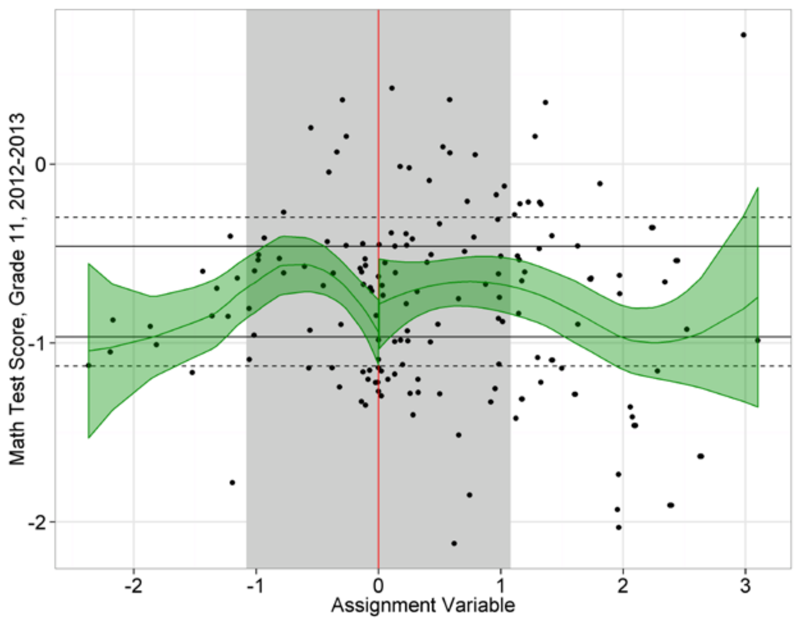
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.33. Math test score in grade 11, place-based analysis





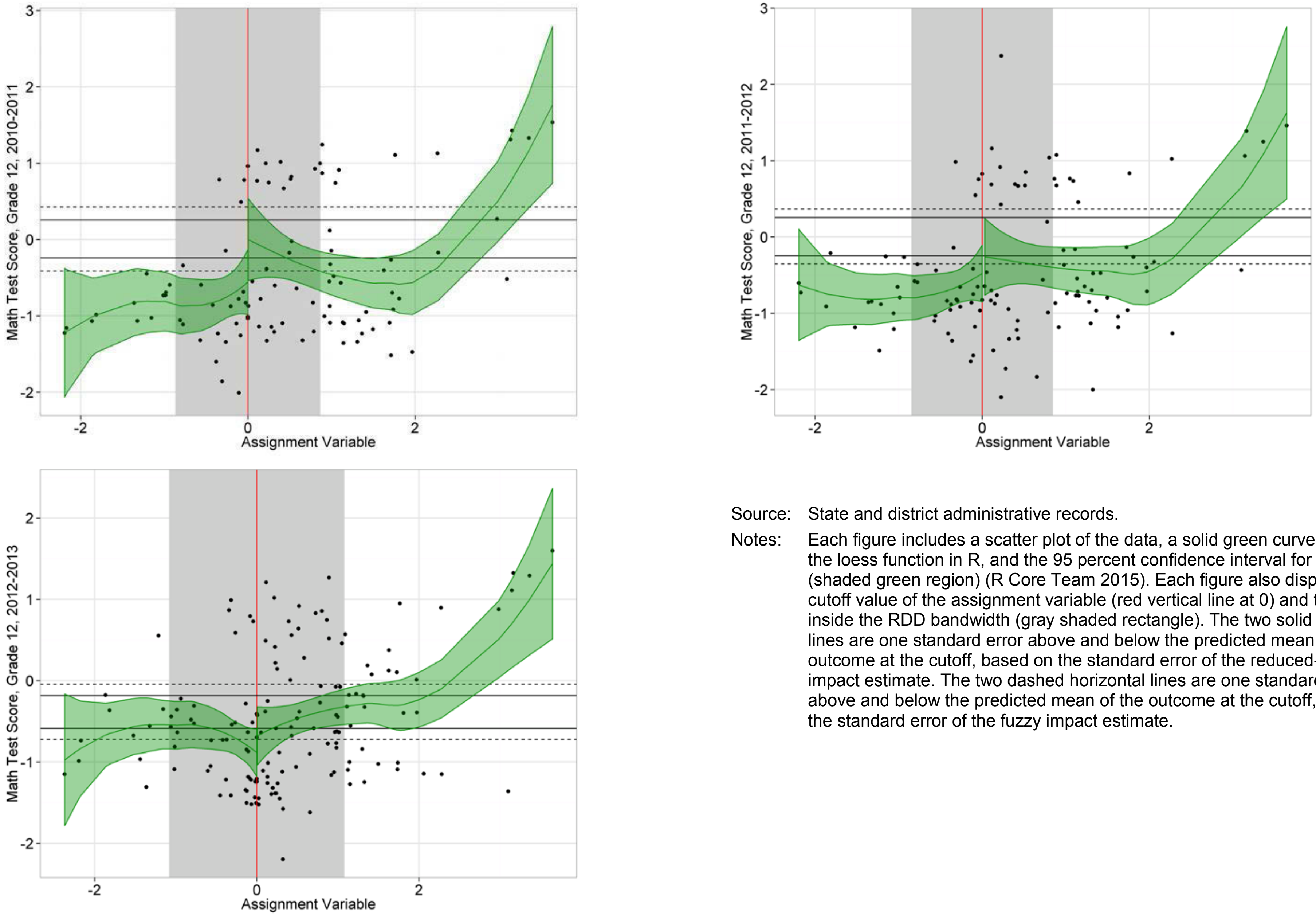


Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

.

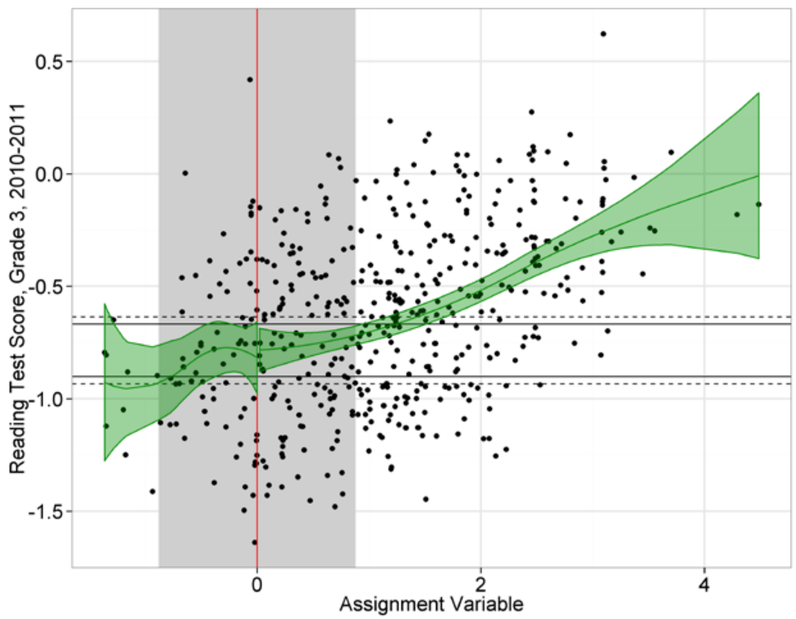
Figure A.34. Math test score in grade 12, place-based analysis

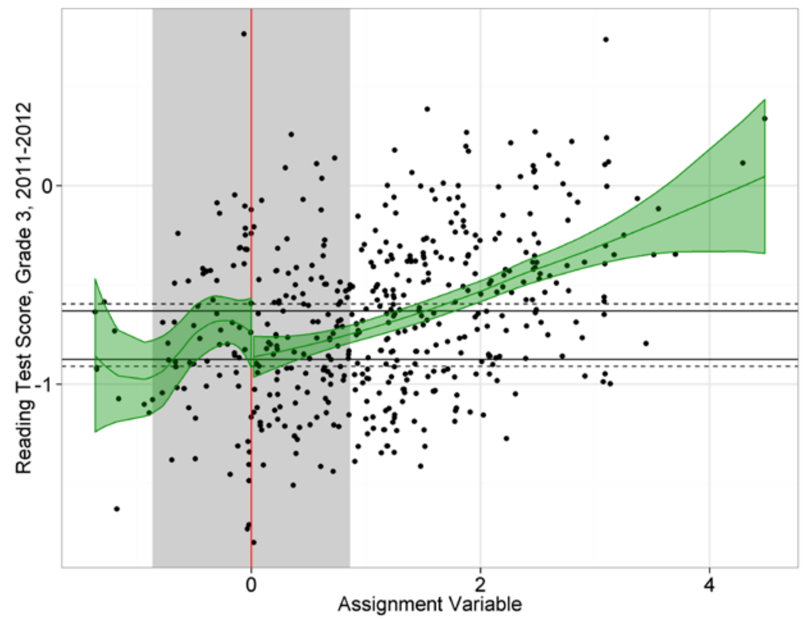


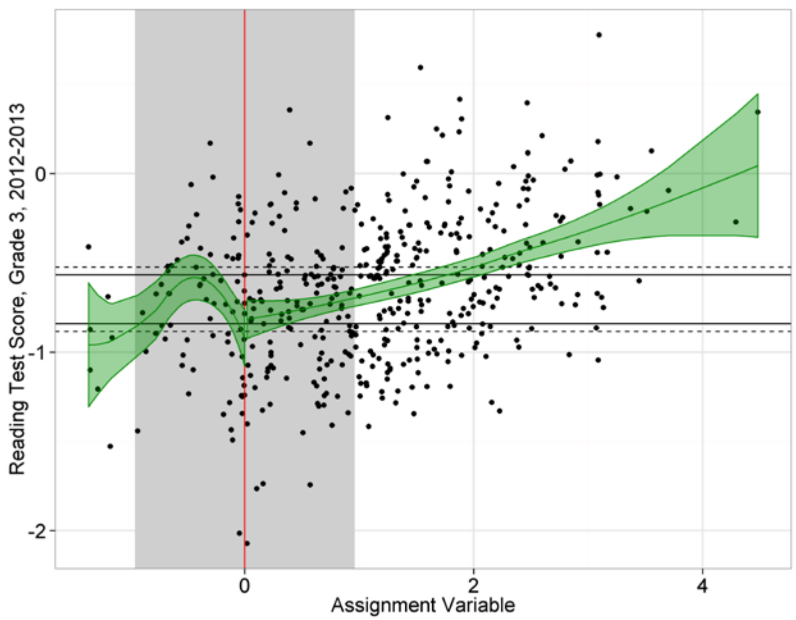
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.35. Reading test score in grade 3, place-based analysis



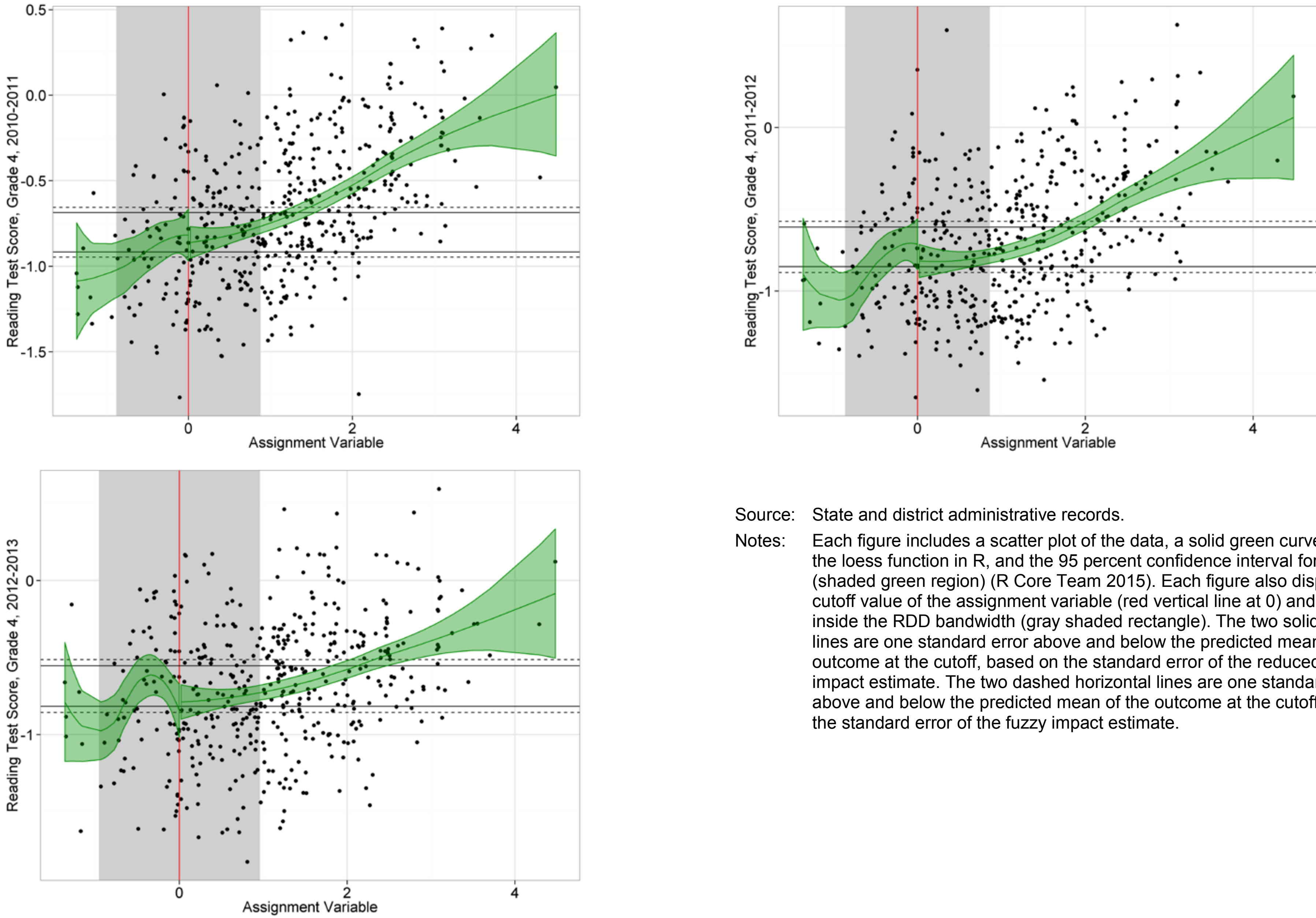




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

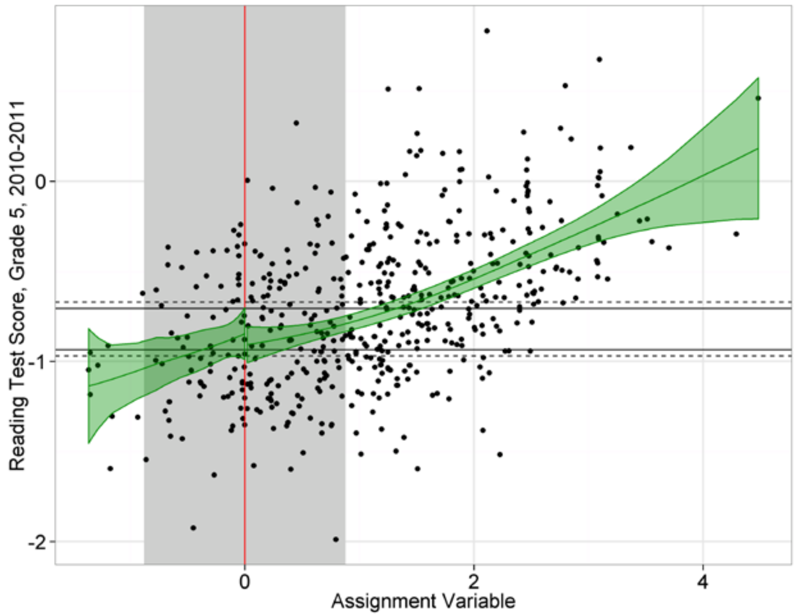
Figure A.36. Reading test score in grade 4, place-based analysis

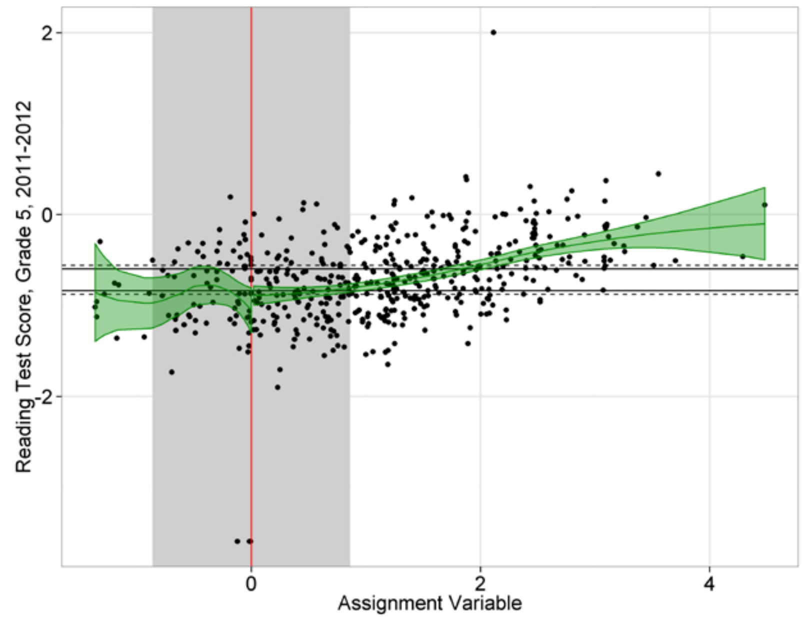


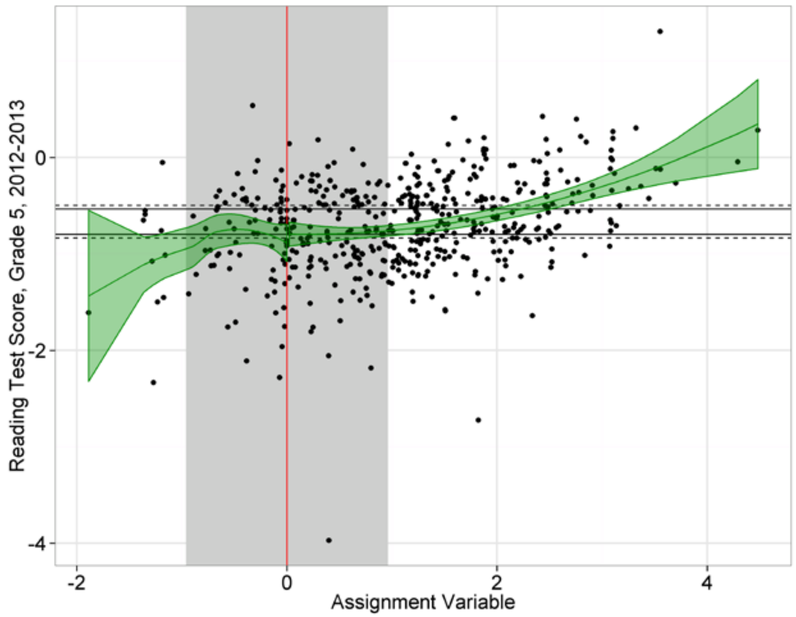
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.37. Reading test score in grade 5, place-based analysis



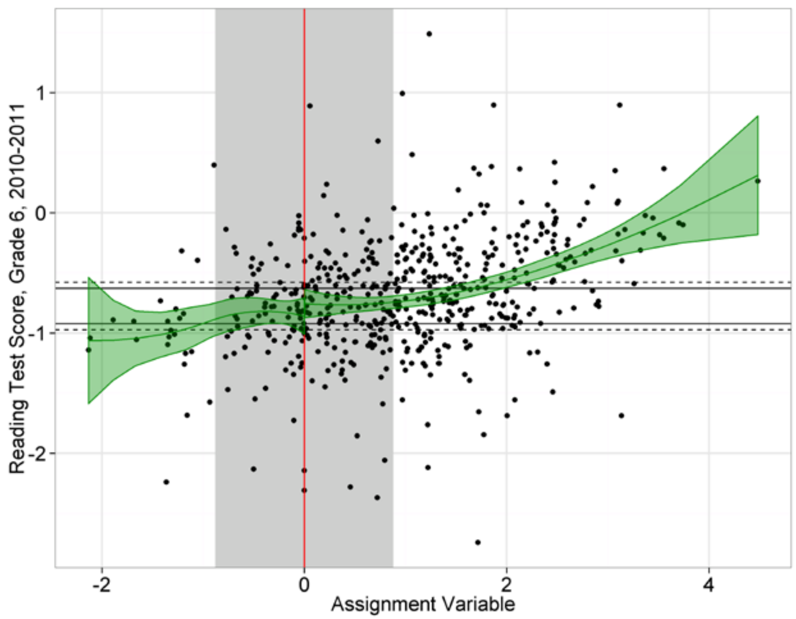


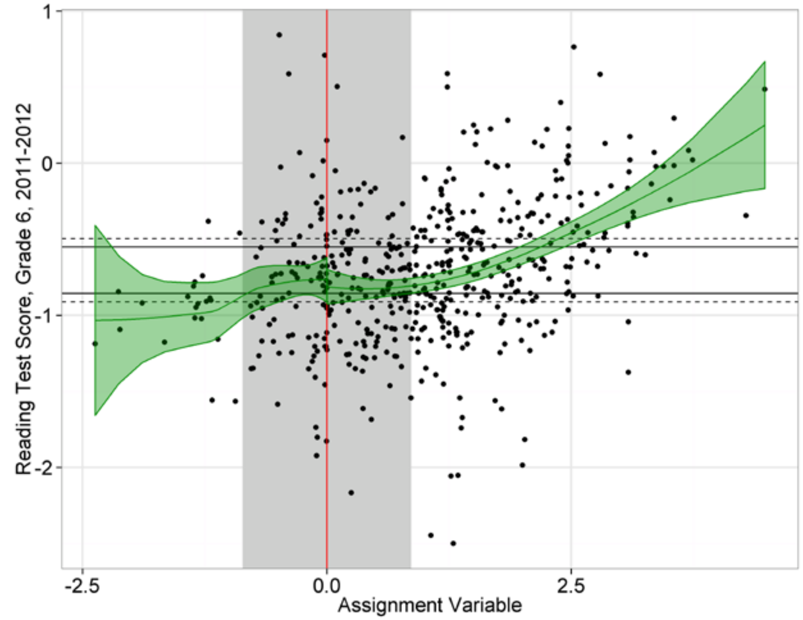


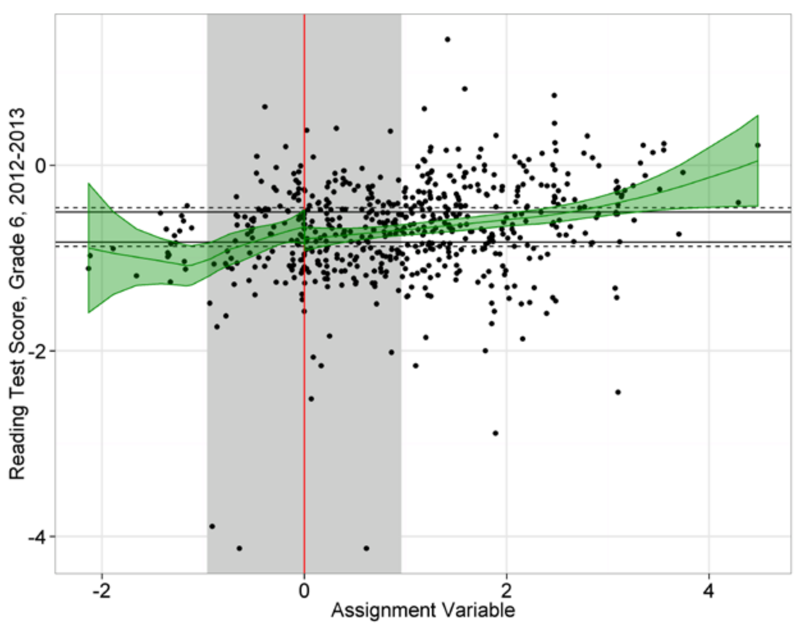
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.38. Reading test score in grade 6, place-based analysis



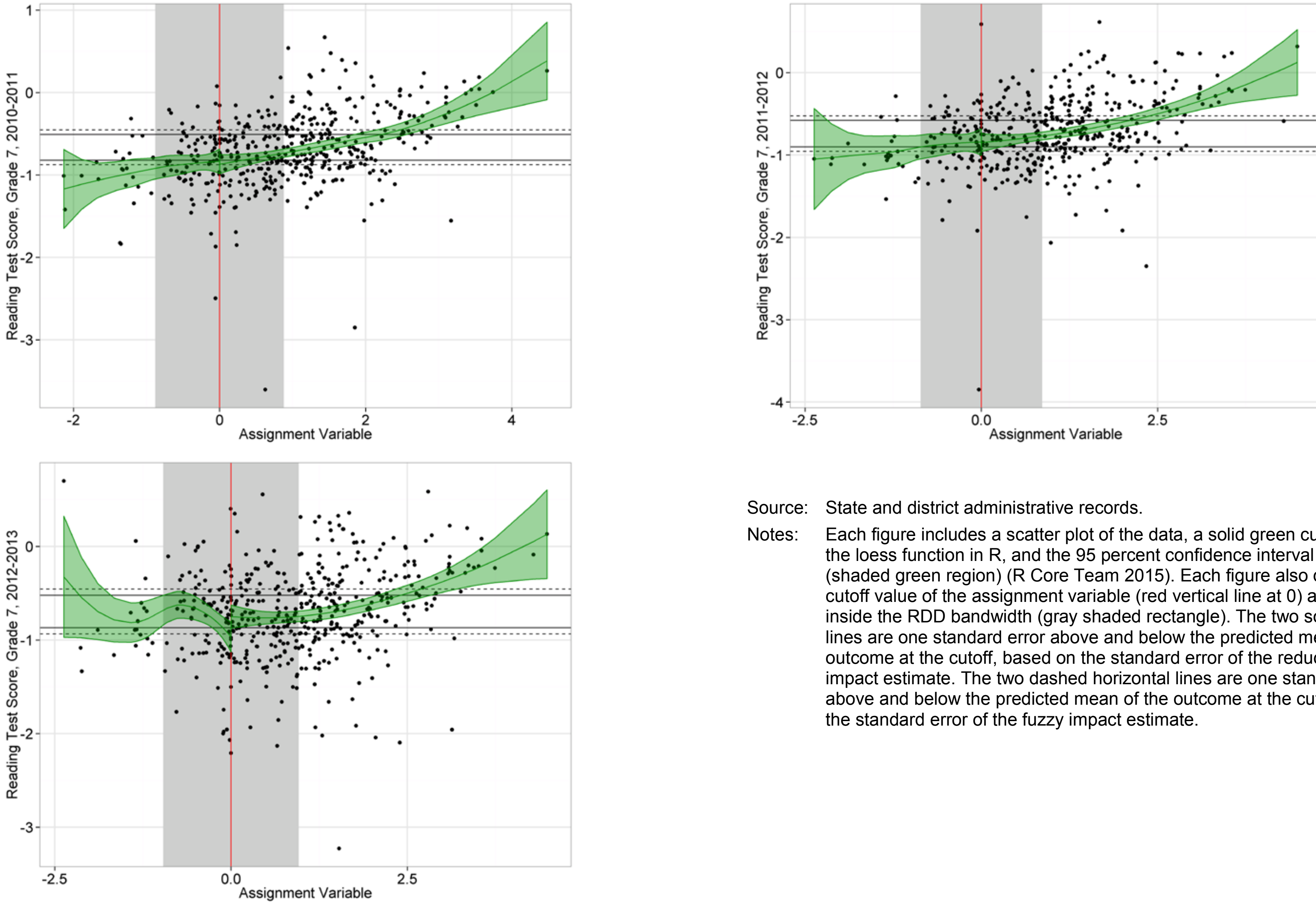




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

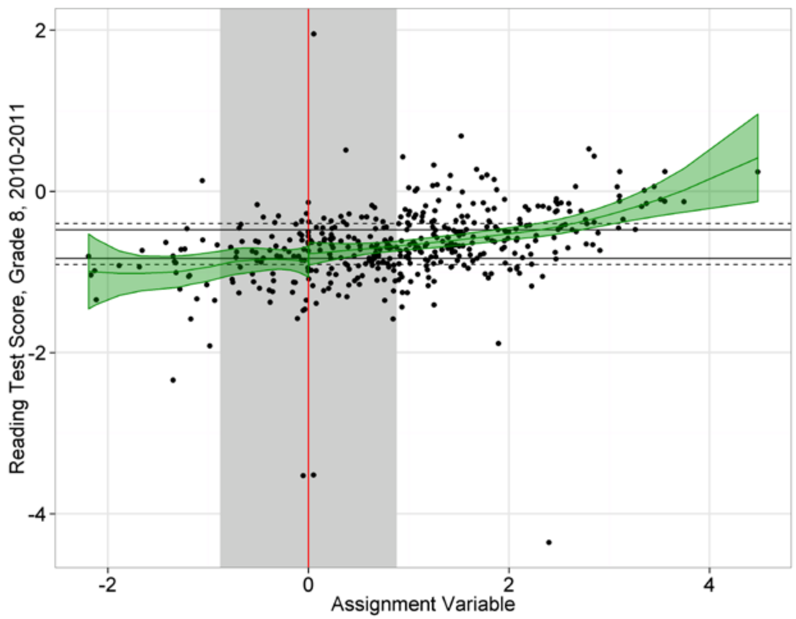
Figure A.39. Reading test score in grade 7, place-based analysis

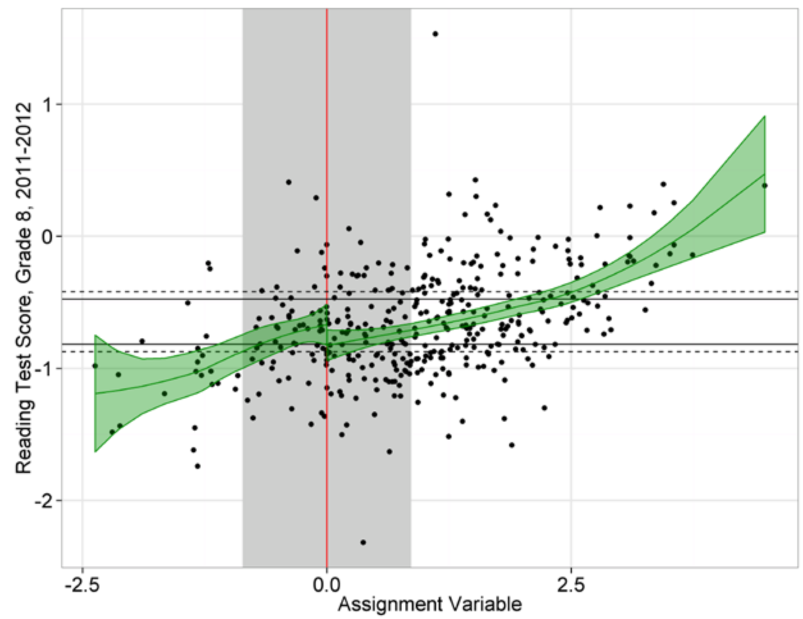


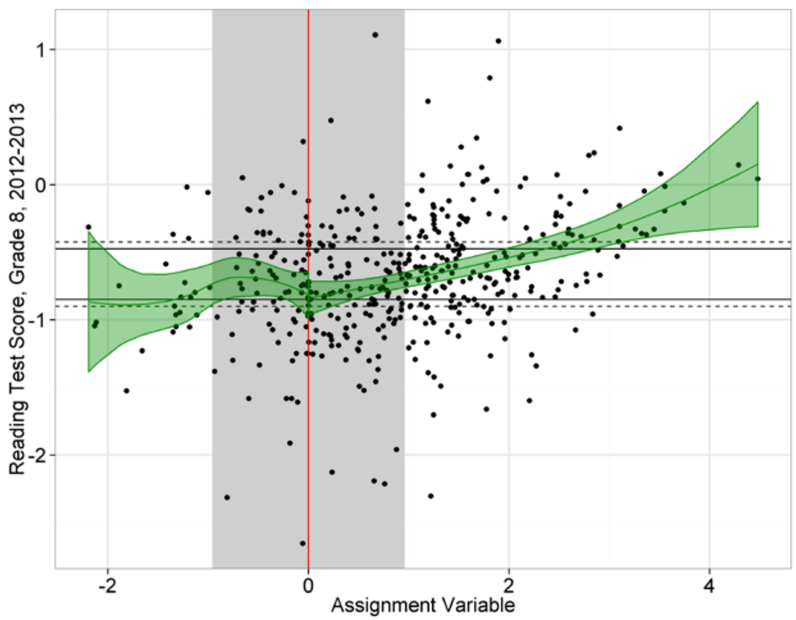
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.40. Reading test score in grade 8, place-based analysis



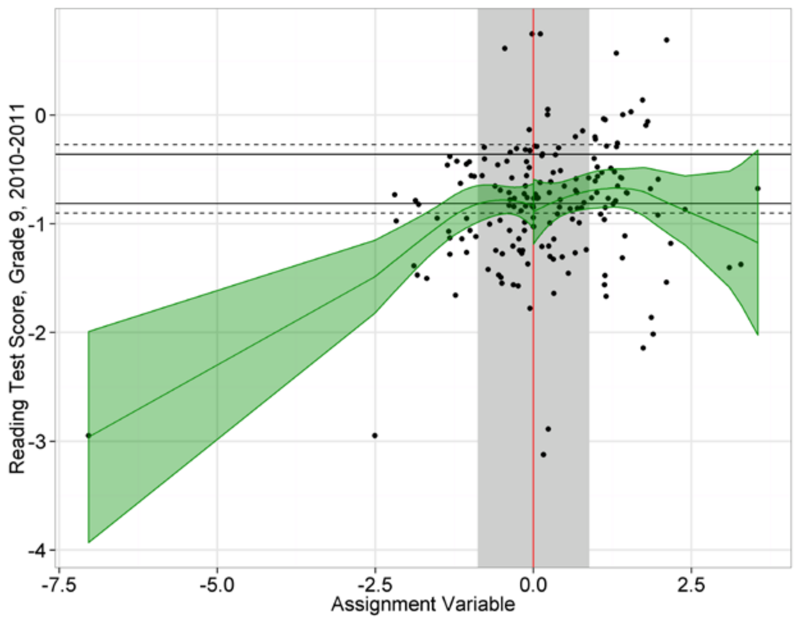


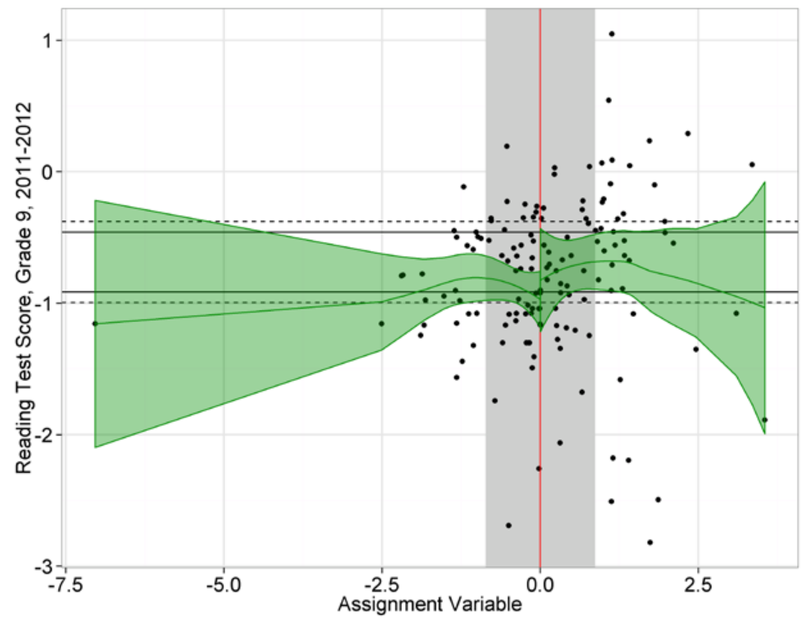


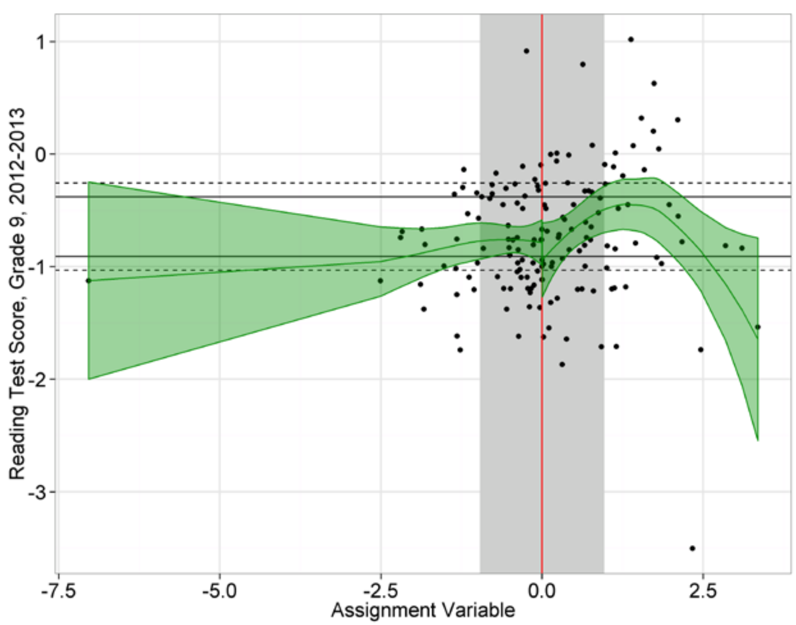
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.41. Reading test score in grade 9, place-based analysis



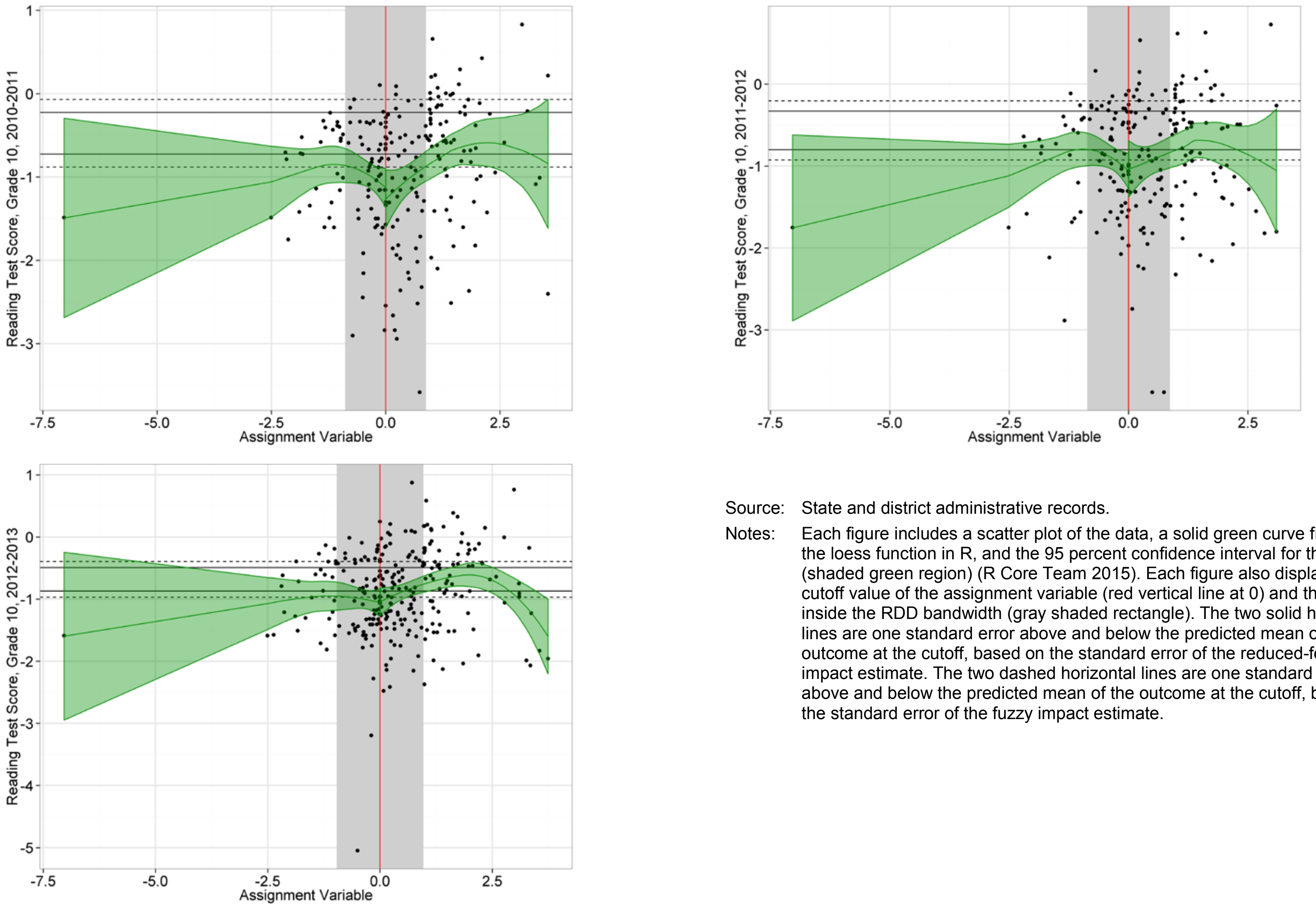




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

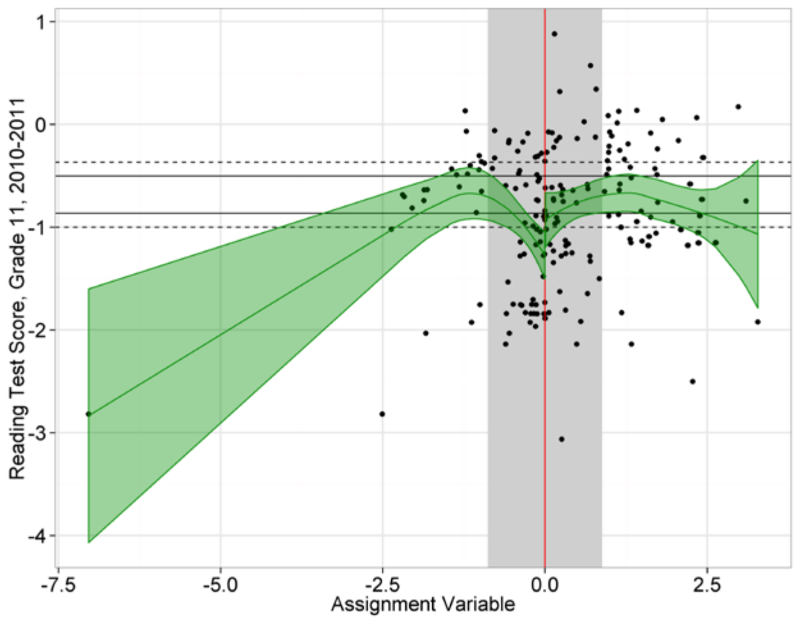
Figure A.42. Reading test score in grade 10, place-based analysis

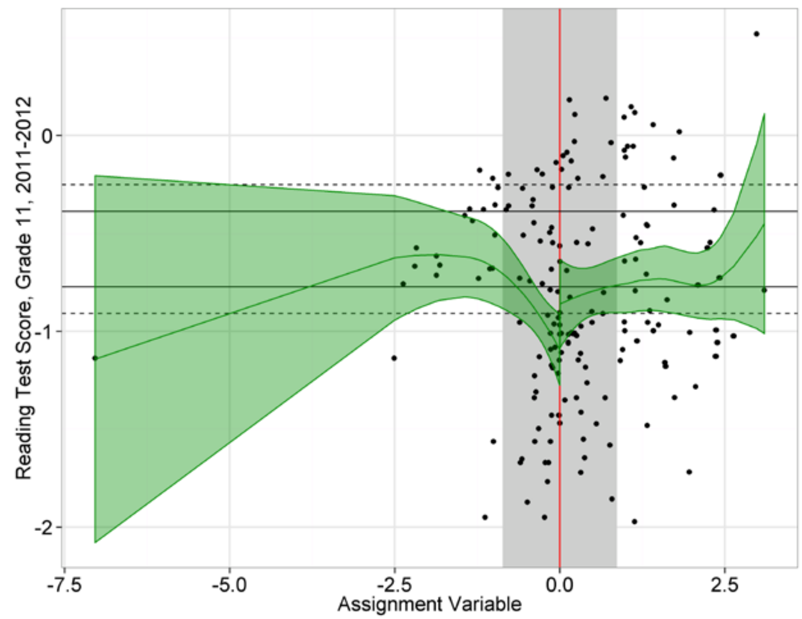


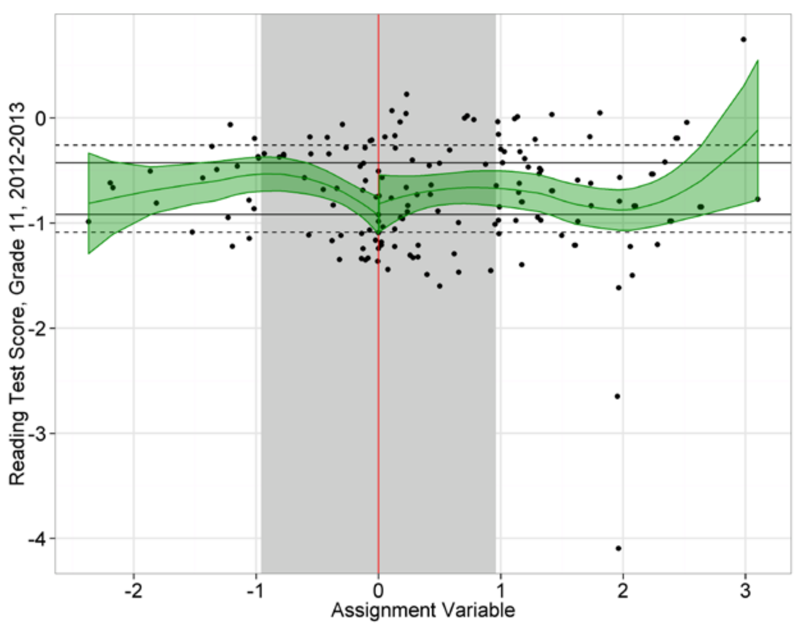
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.43. Reading test score in grade 11, place-based analysis



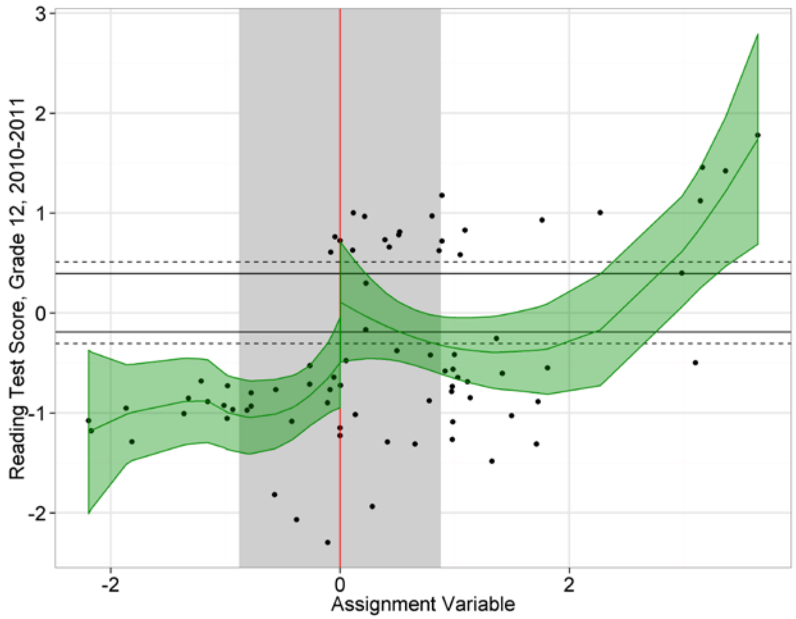


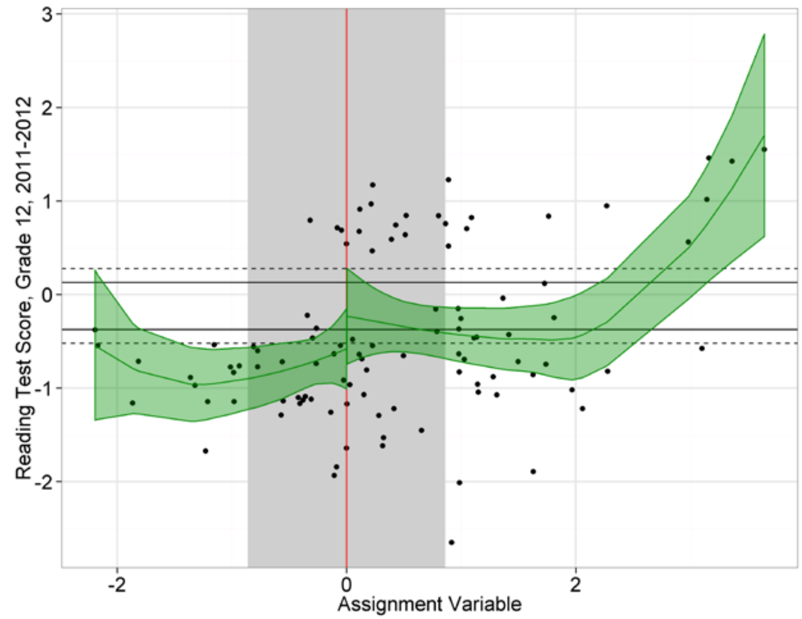


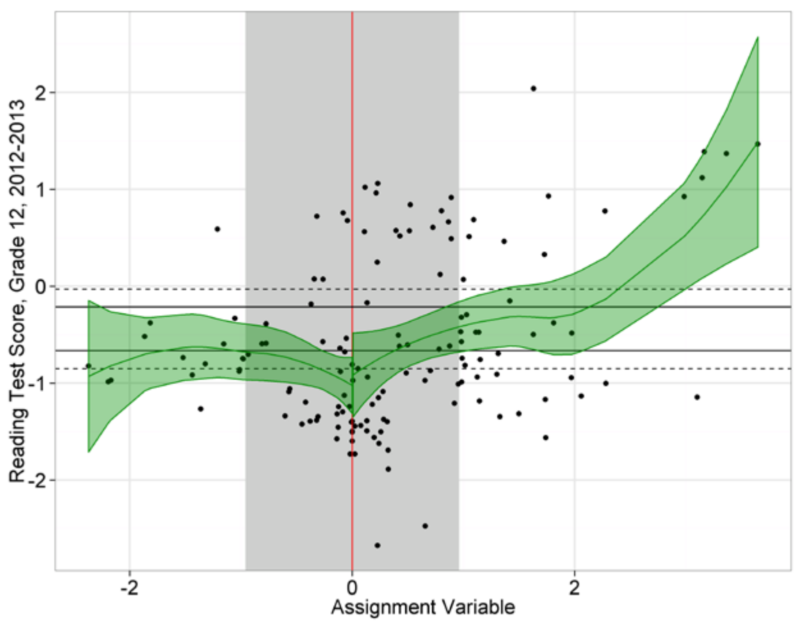
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.44. Reading test score in grade 12, place-based analysis



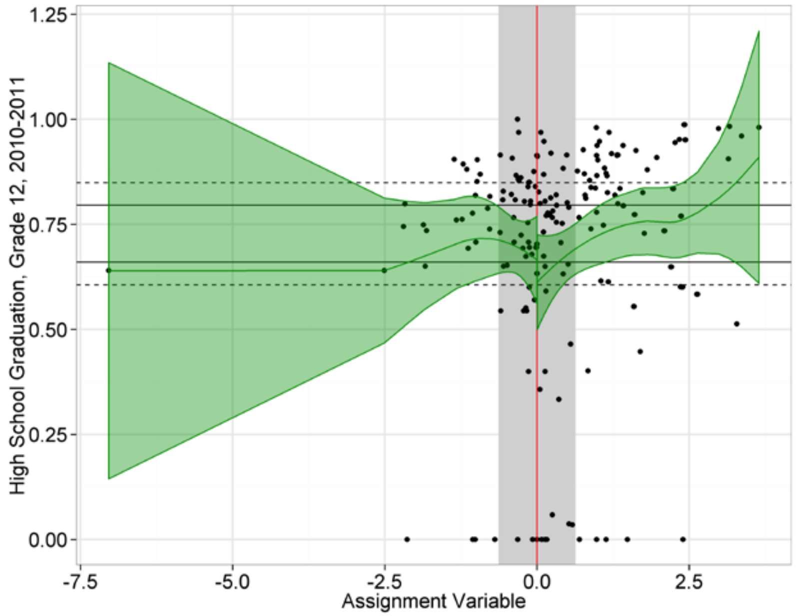


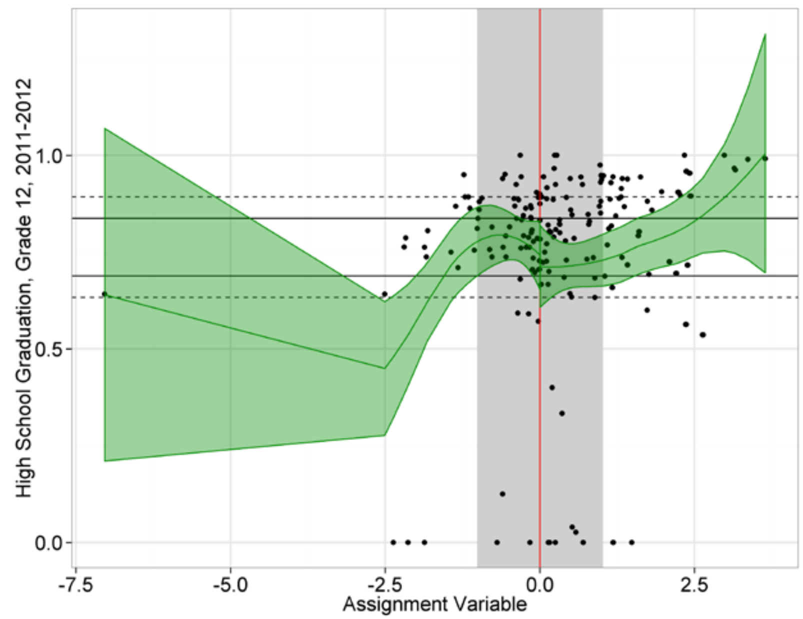


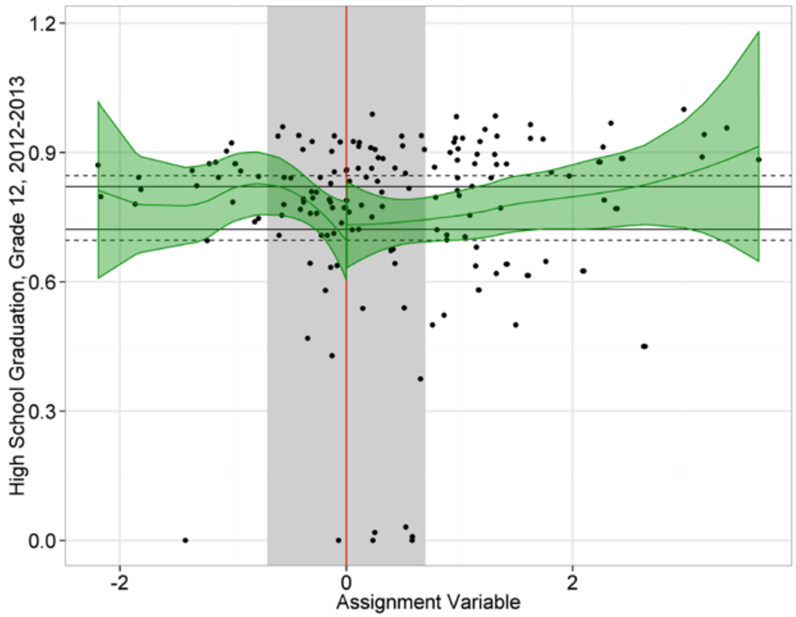
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.45. High school graduation, place-based analysis



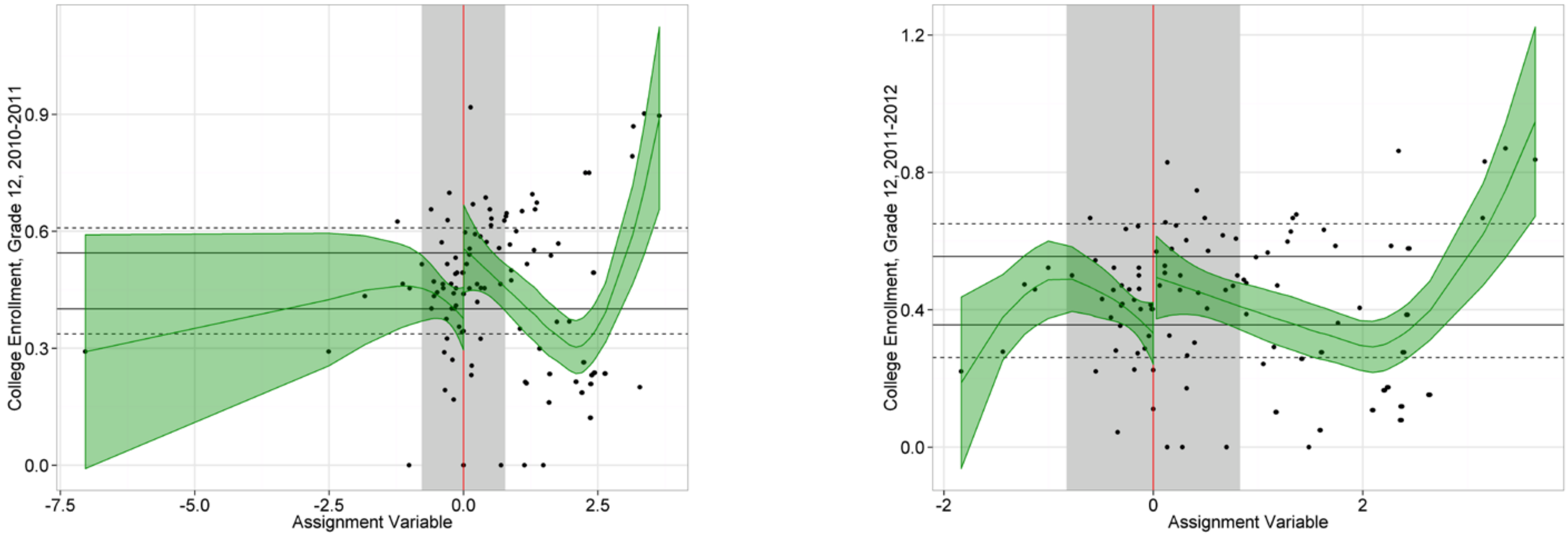




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

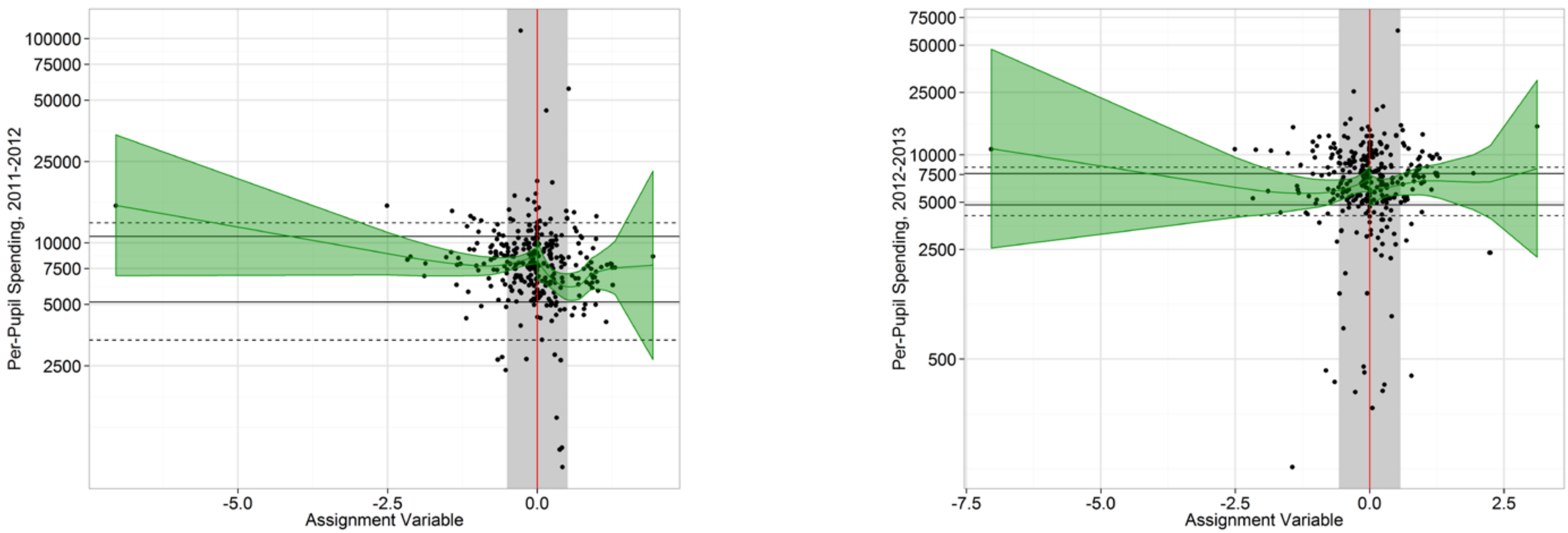
Figure A.46. College enrollment, place-based analysis



Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate. We were not able to estimate an impact on college enrollment for 2012–2013 due to insufficient sample sizes.

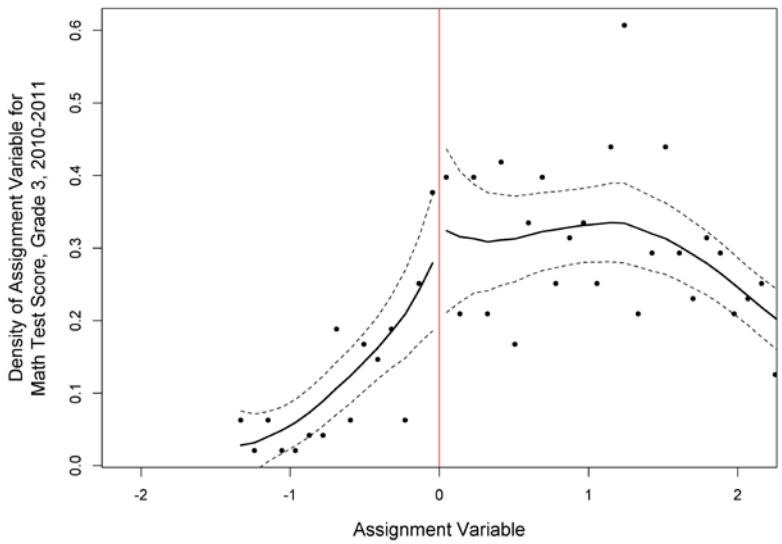
Figure A.47. Per-pupil spending, place-based analysis

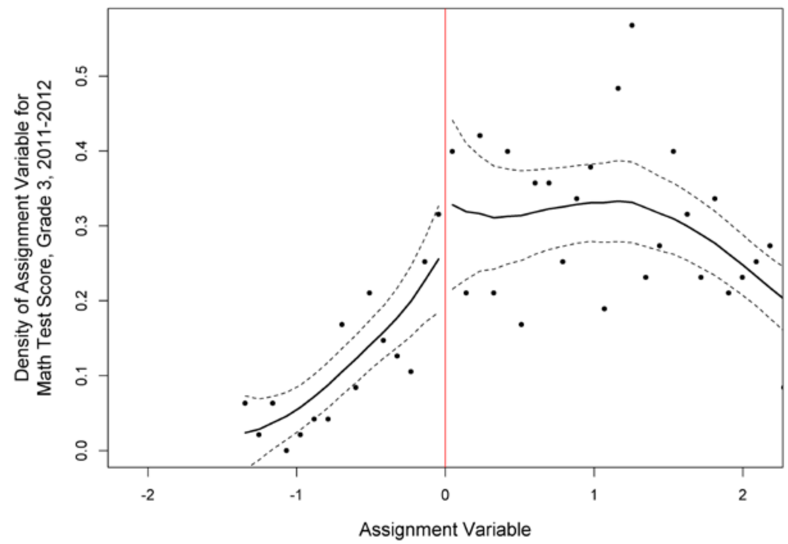


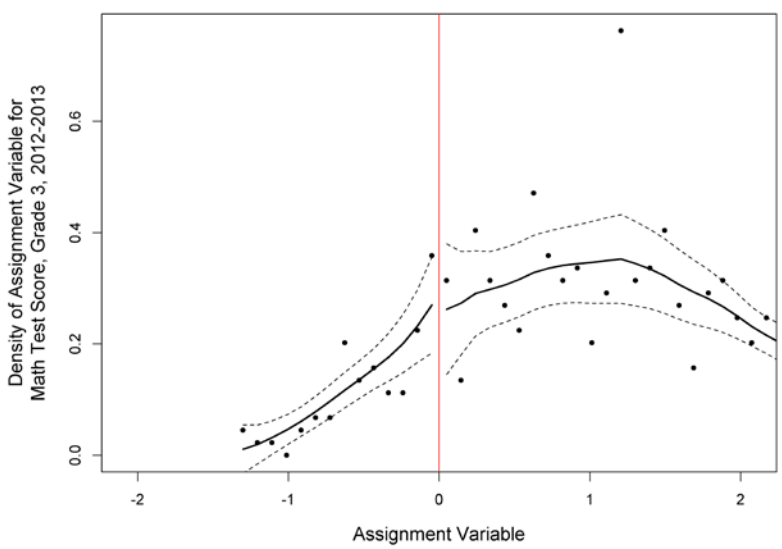
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.48. Density of the assignment variable for math test score in grade 3, accounting for student mobility



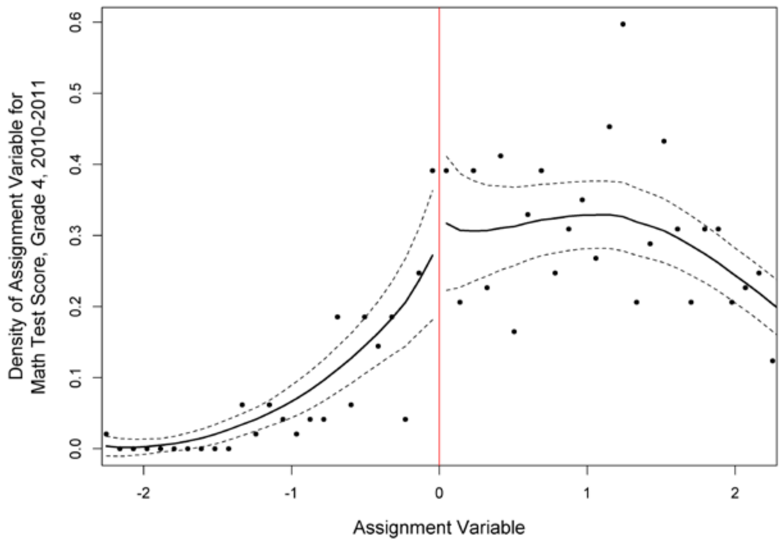


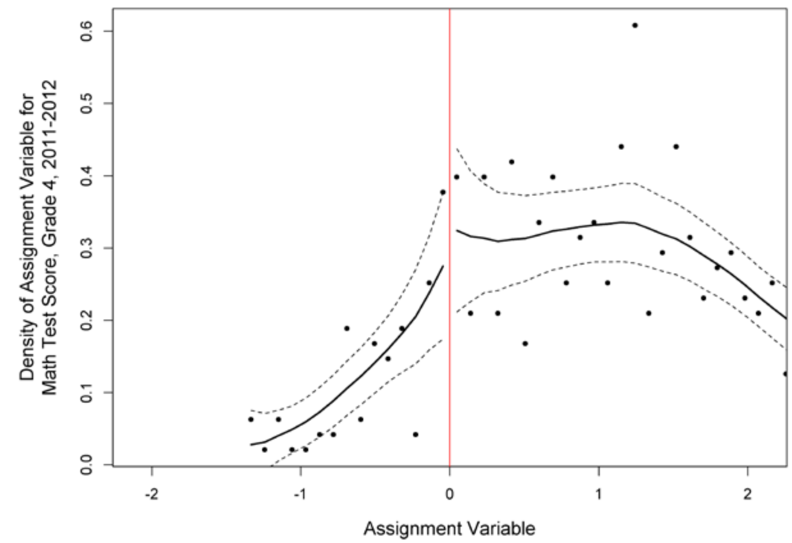


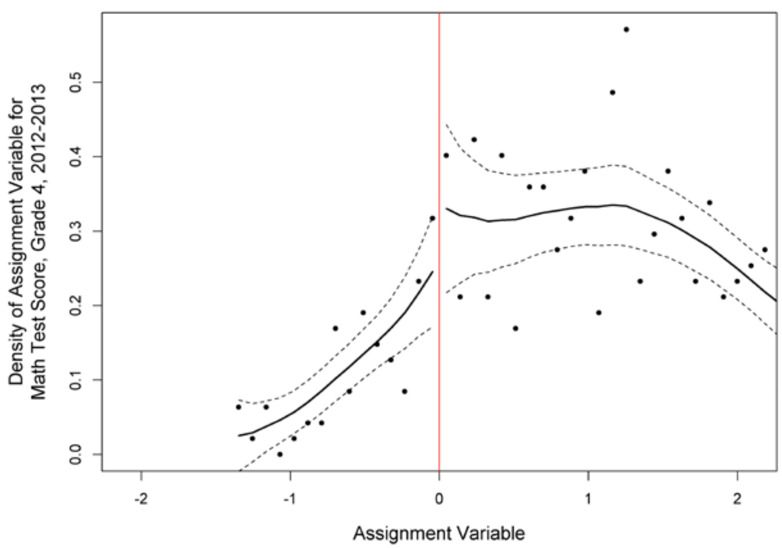
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.49. Density of the assignment variable for math test score in grade 4, accounting for student mobility



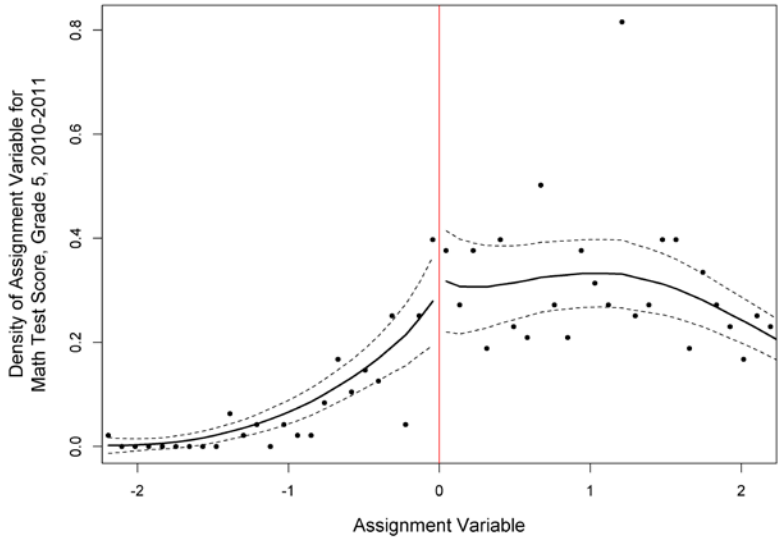


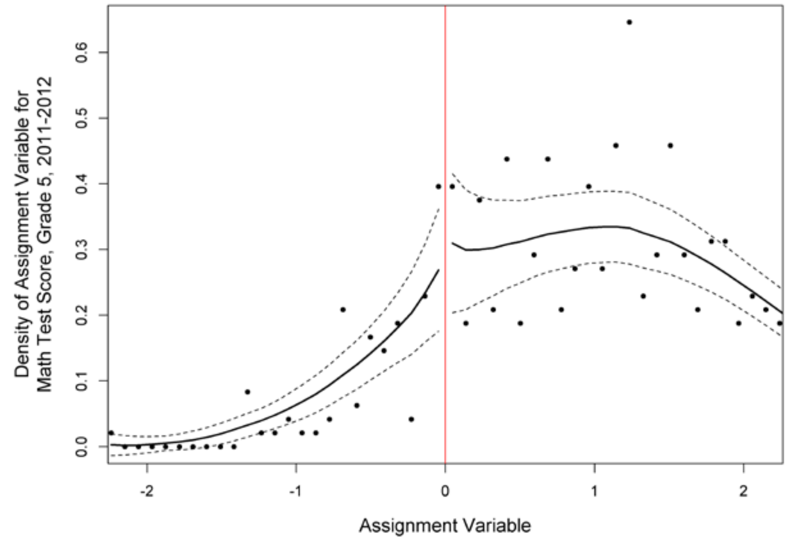


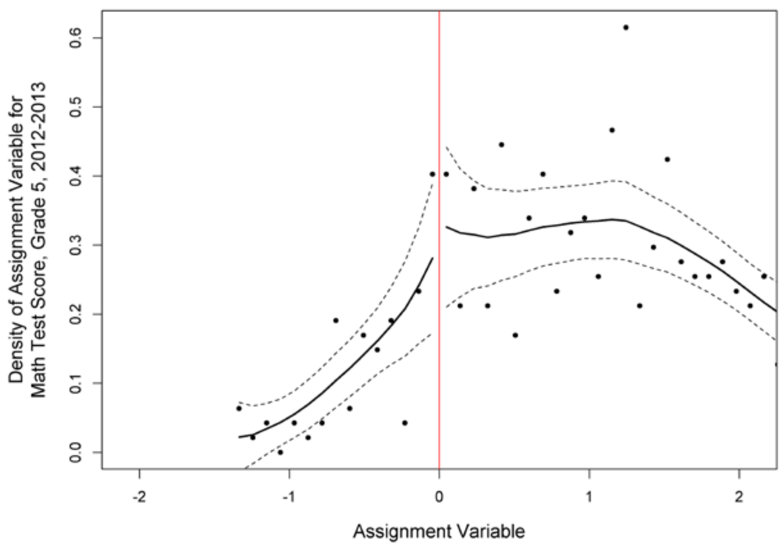
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.50. Density of the assignment variable for math test score in grade 5, accounting for student mobility



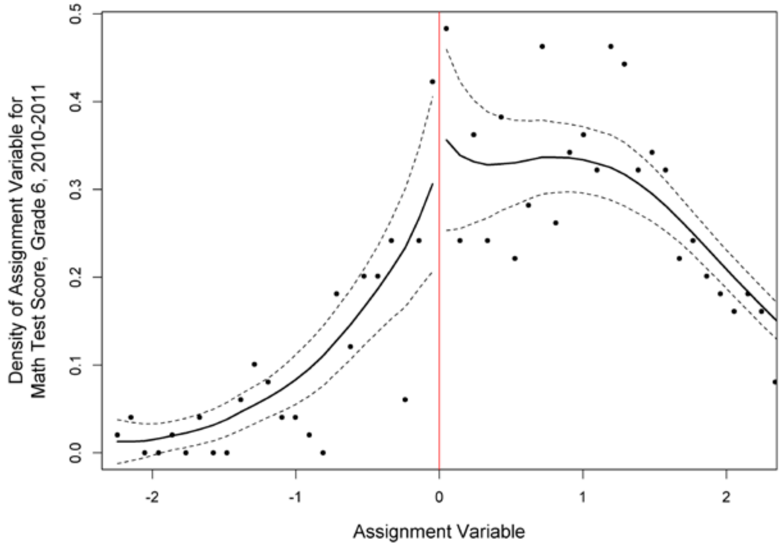


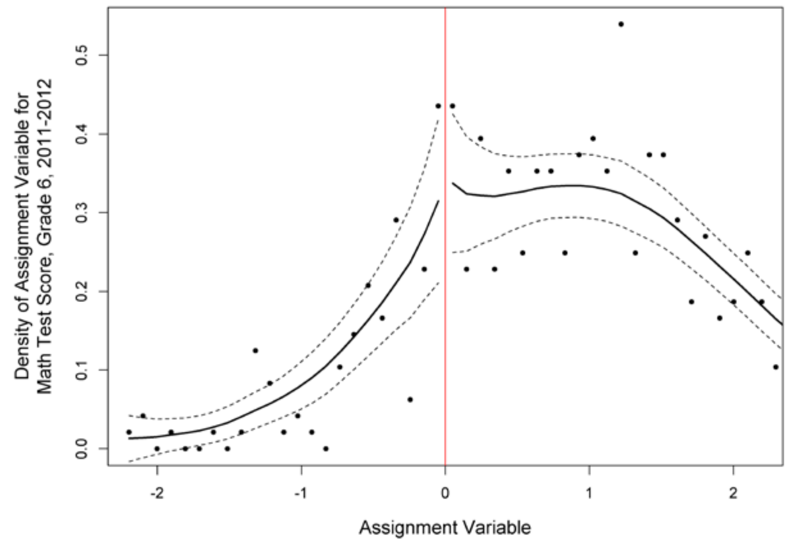


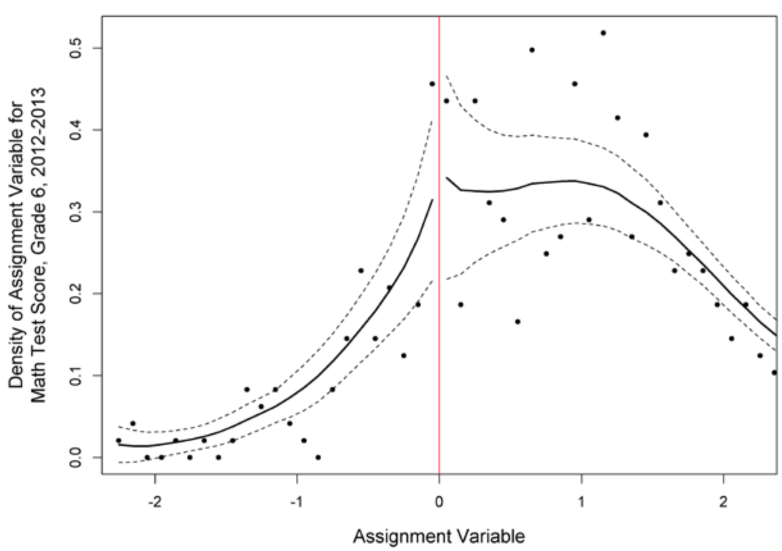
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.51. Density of the assignment variable for math test score in grade 6, accounting for student mobility



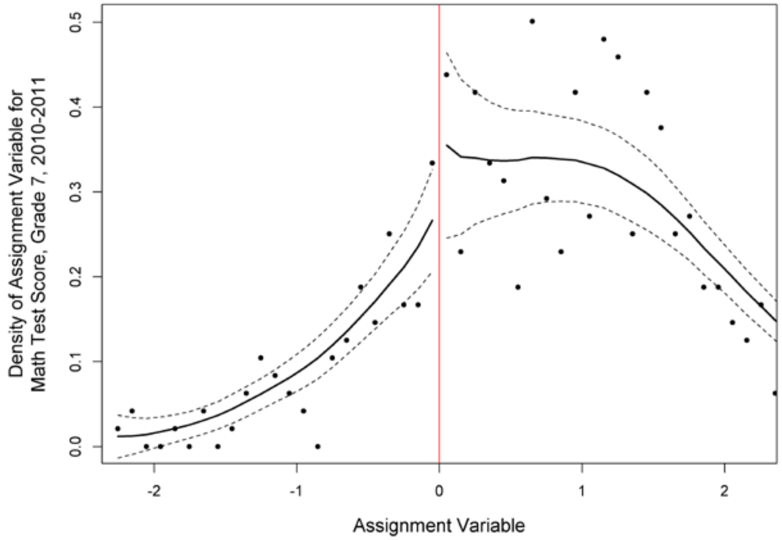


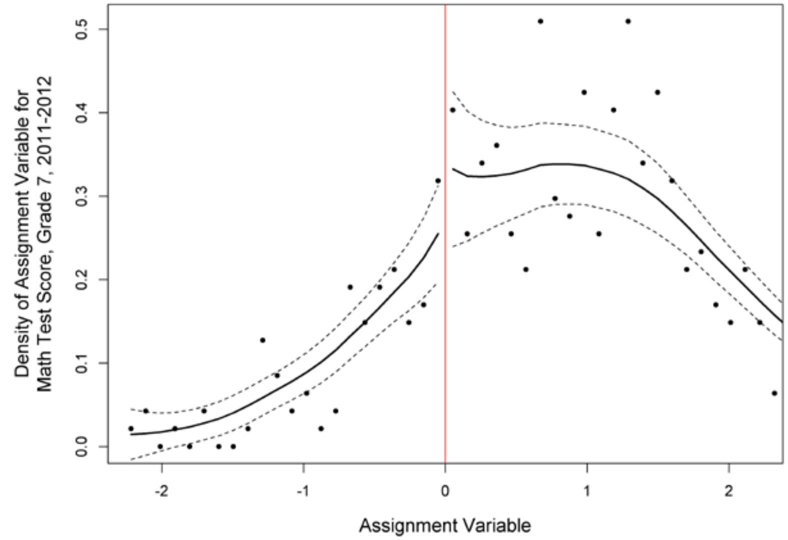


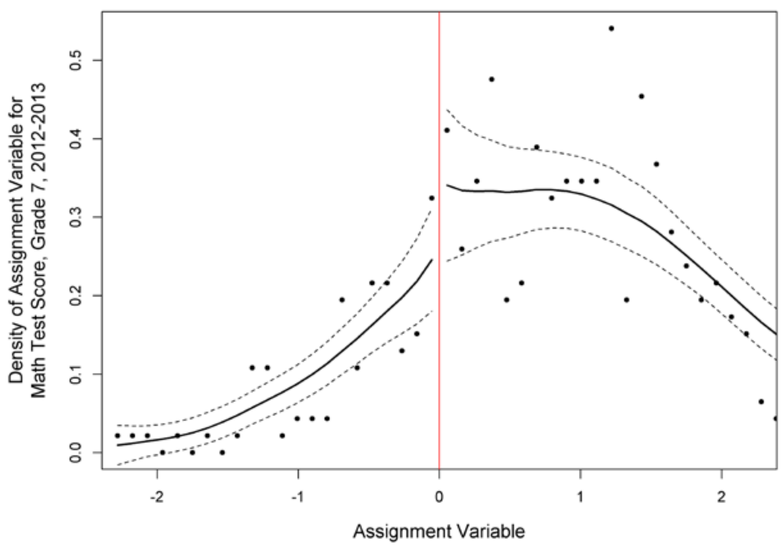
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.52. Density of the assignment variable for math test score in grade 7, accounting for student mobility



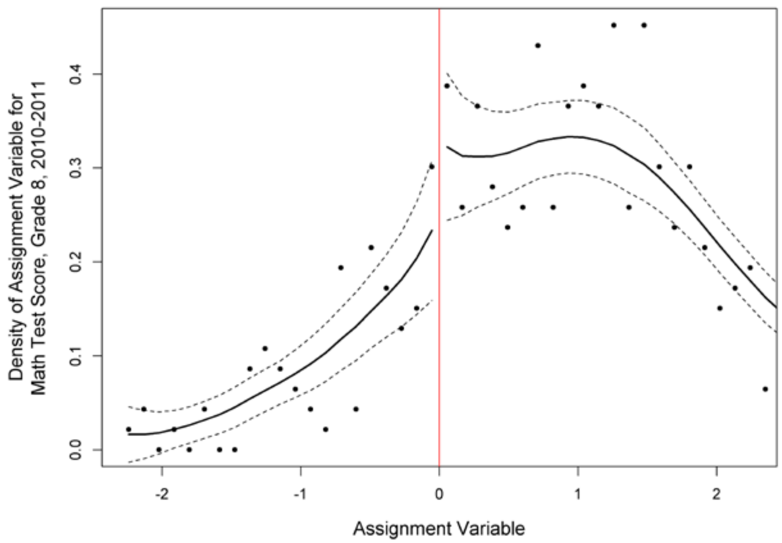


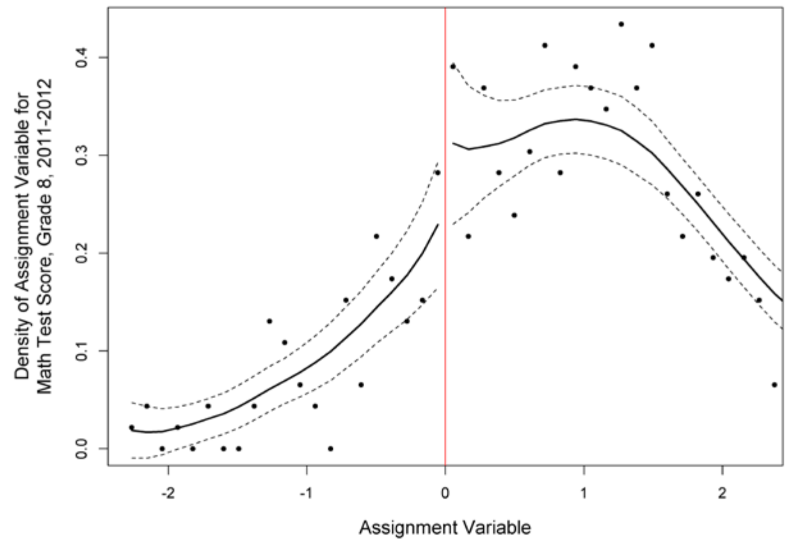


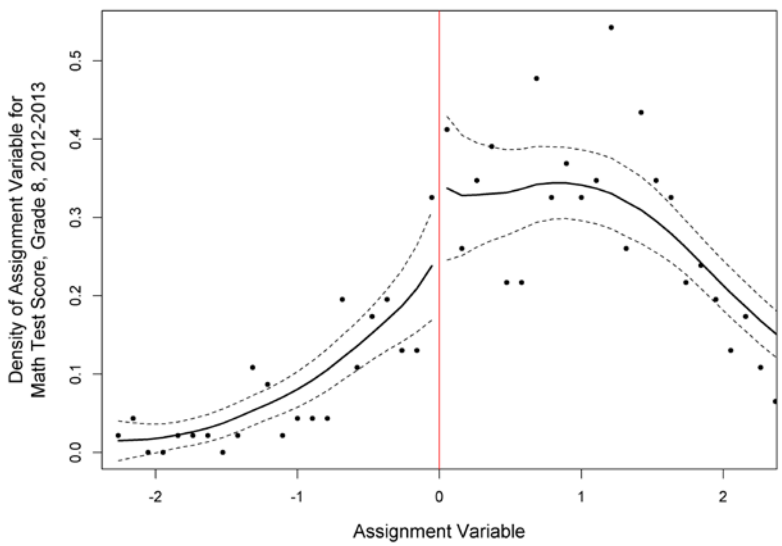
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.53. Density of the assignment variable for math test score in grade 8, accounting for student mobility



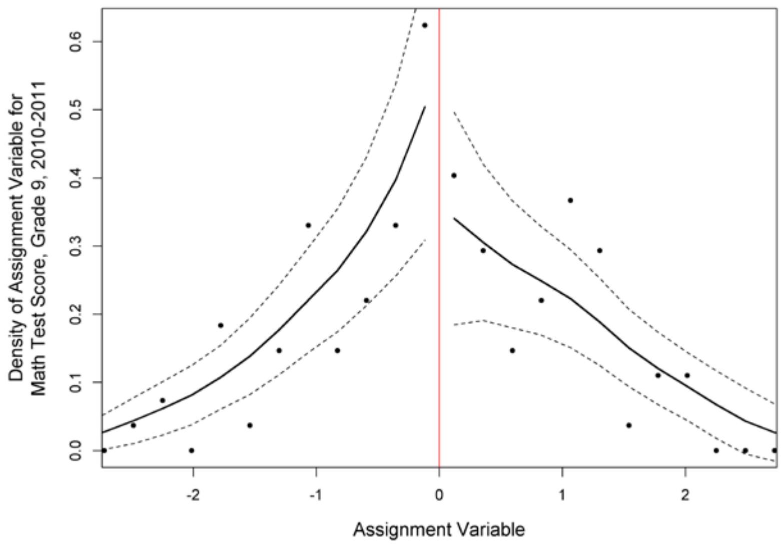


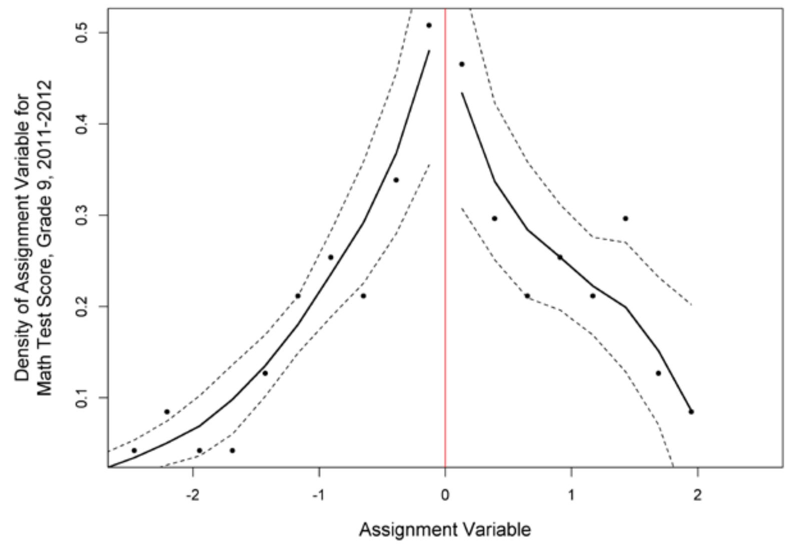


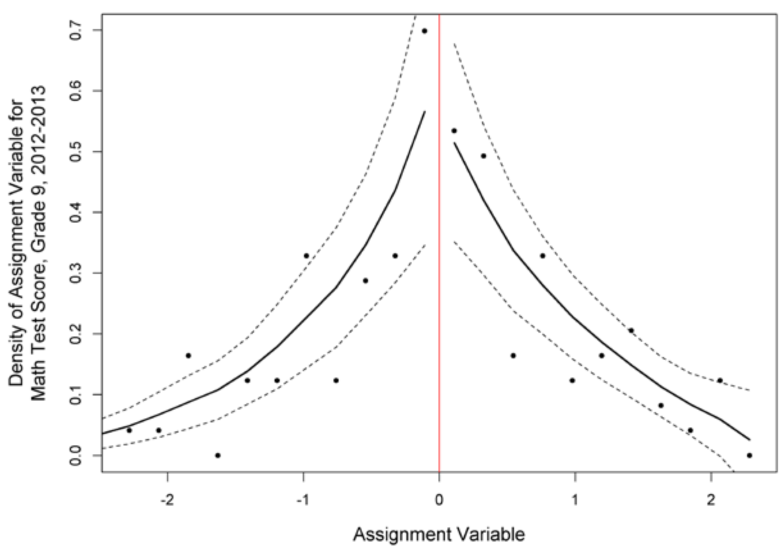
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.54. Density of the assignment variable for math test score in grade 9, accounting for student mobility



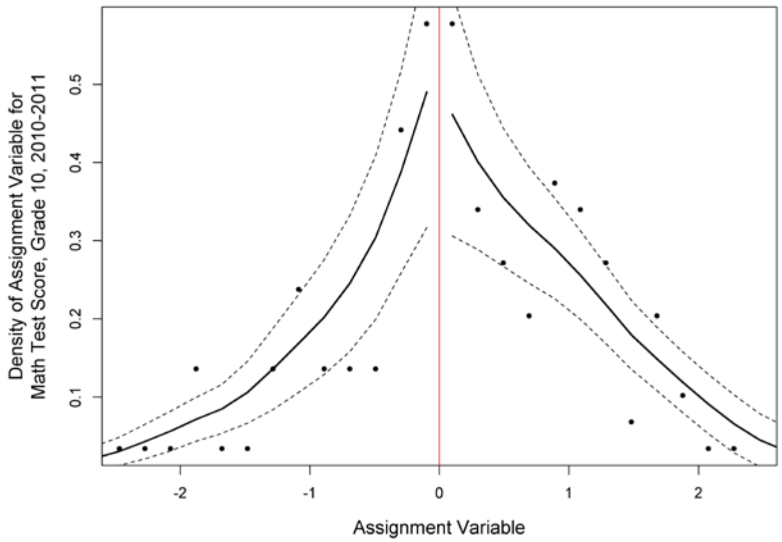


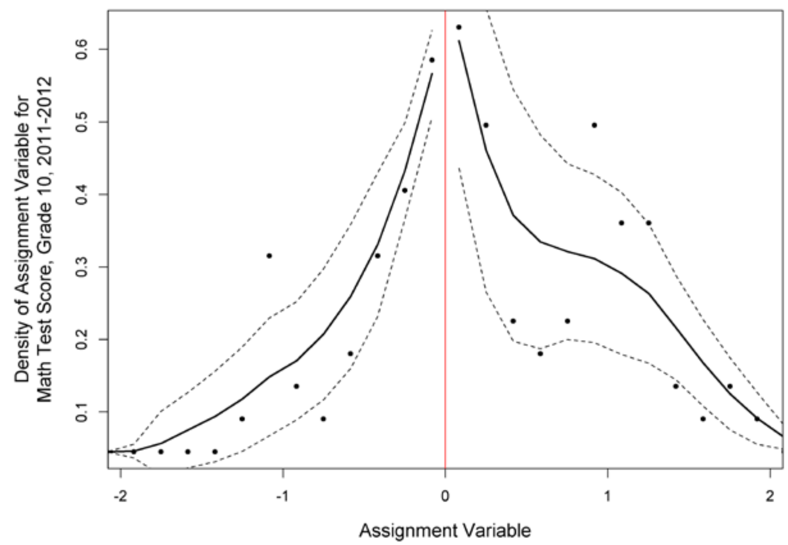


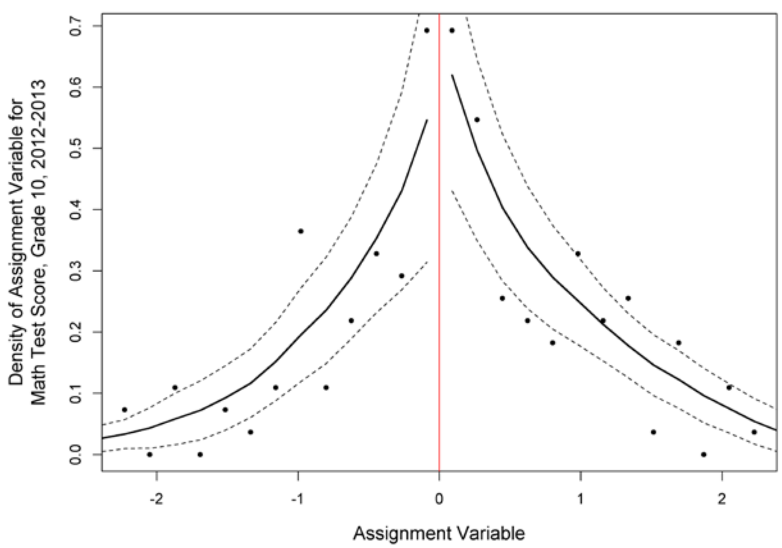
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.55. Density of the assignment variable for math test score in grade 10, accounting for student mobility



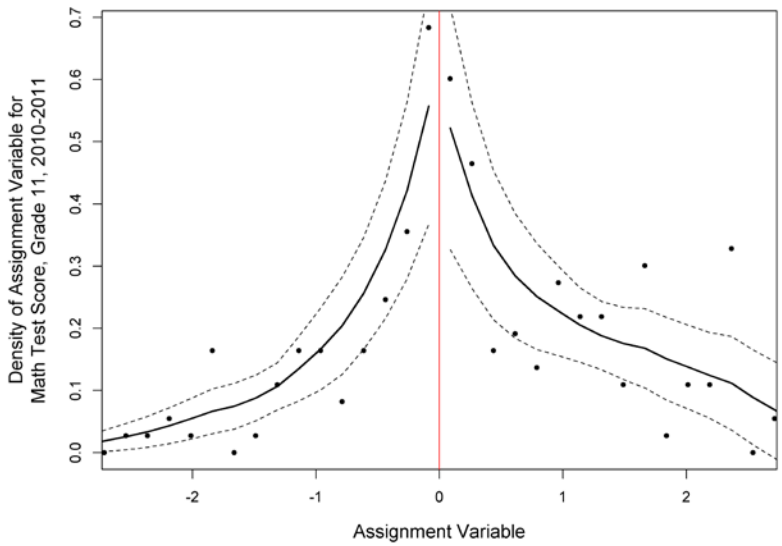


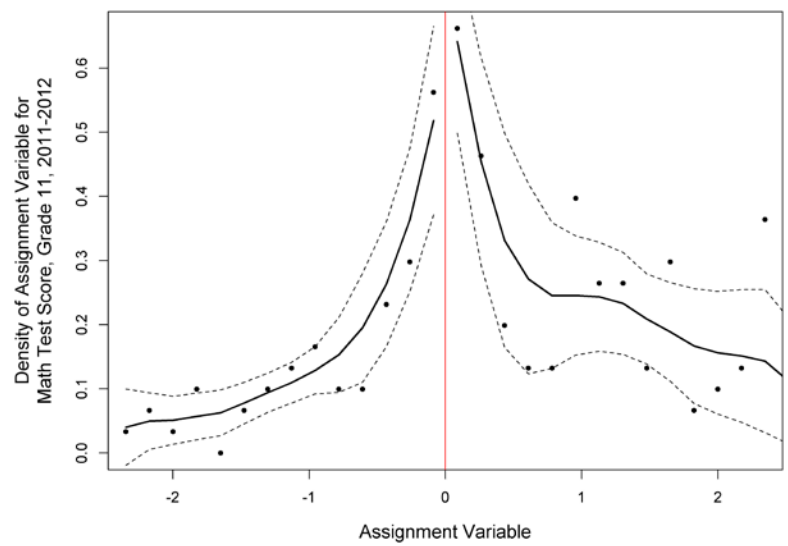


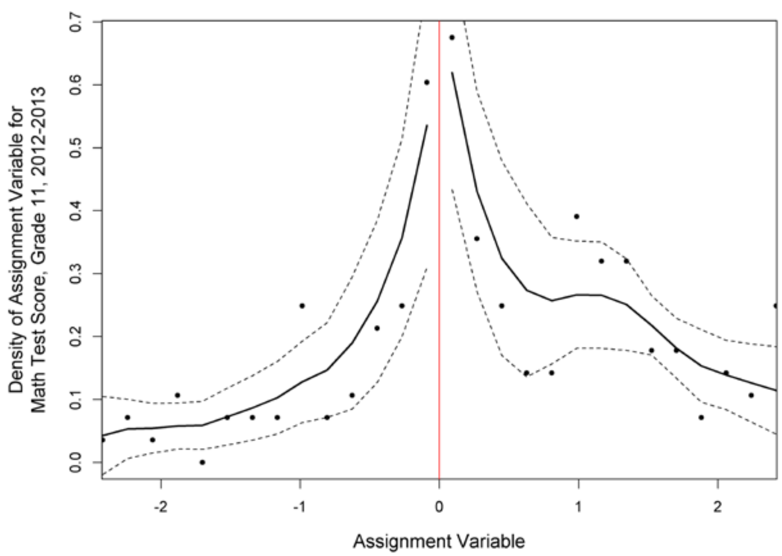
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.56. Density of the assignment variable for math test score in grade 11, accounting for student mobility



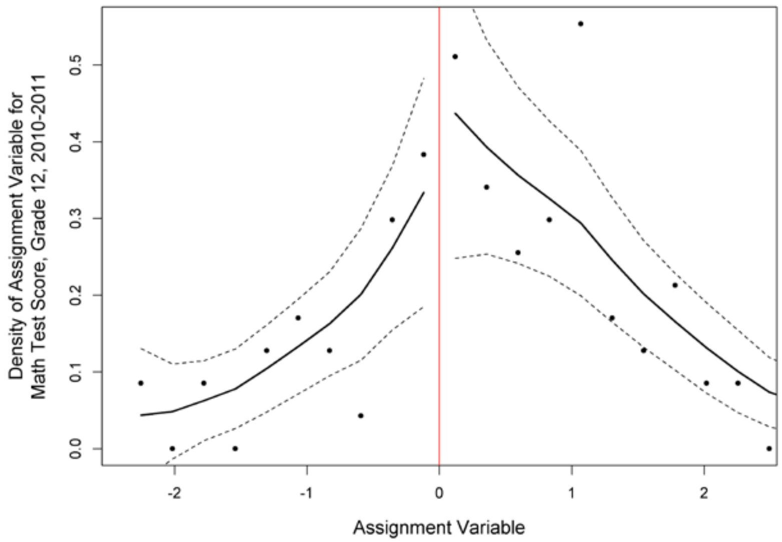


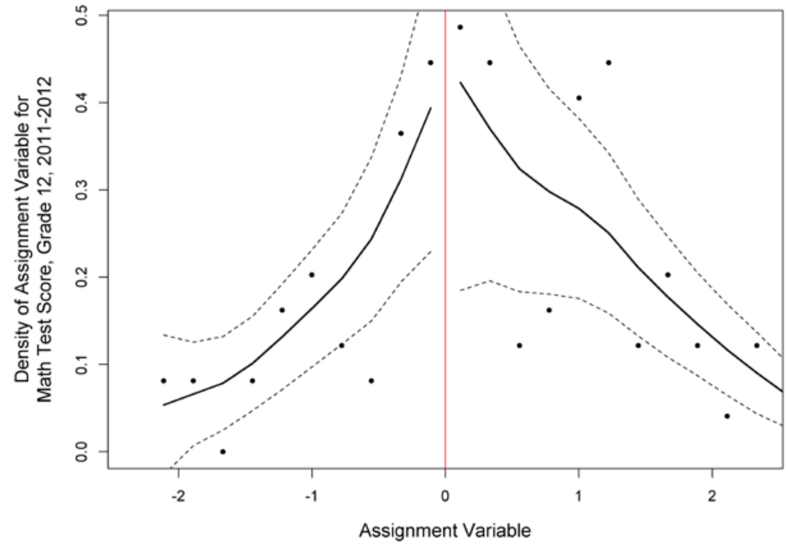


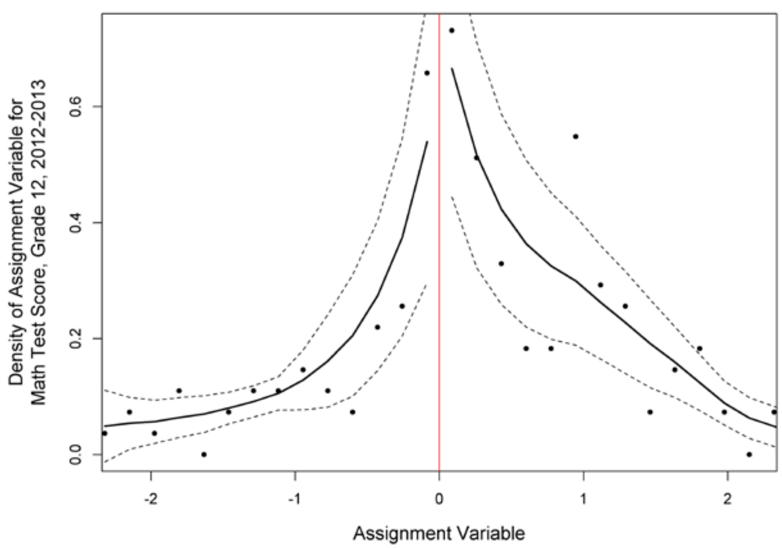
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.57. Density of the assignment variable for math test score in grade 12, accounting for student mobility



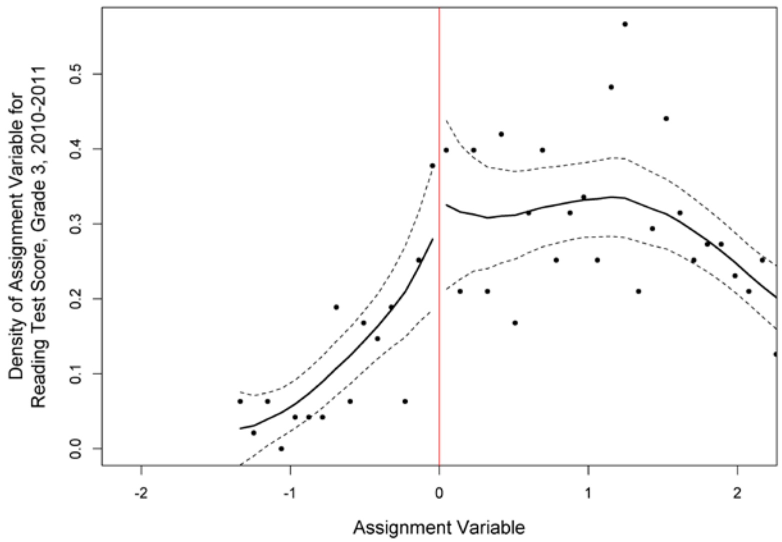


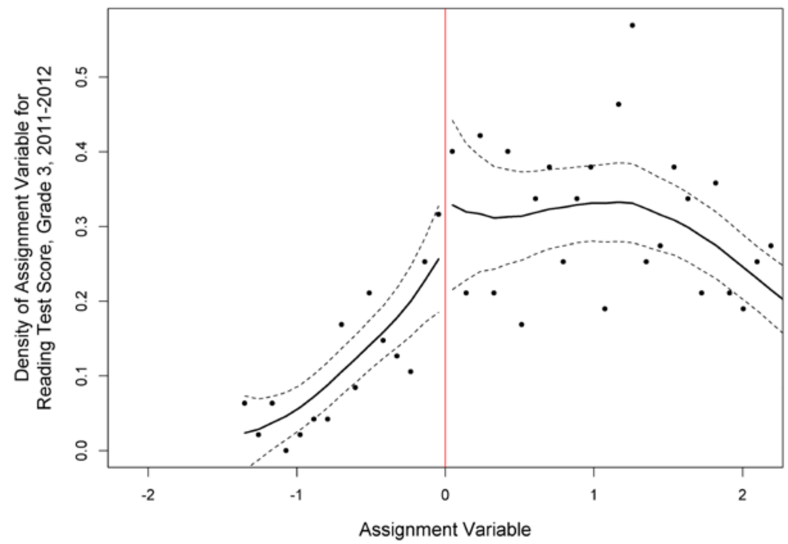


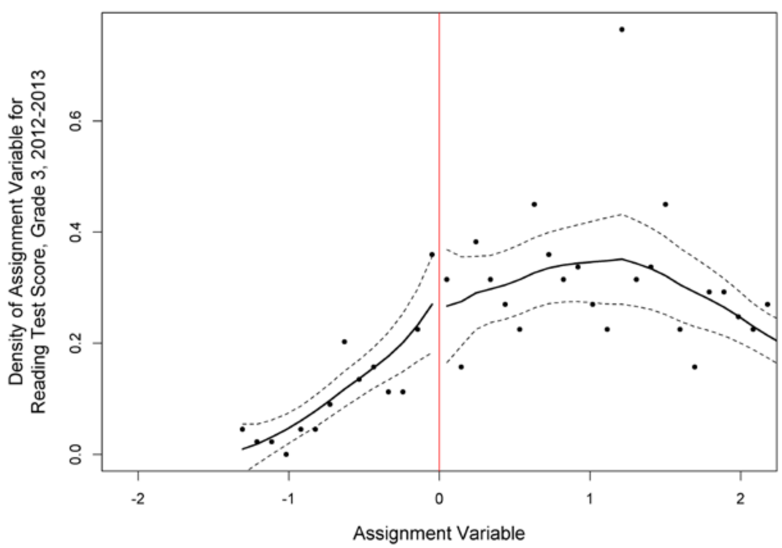
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.58. Density of the assignment variable for reading test score in grade 3, accounting for student mobility



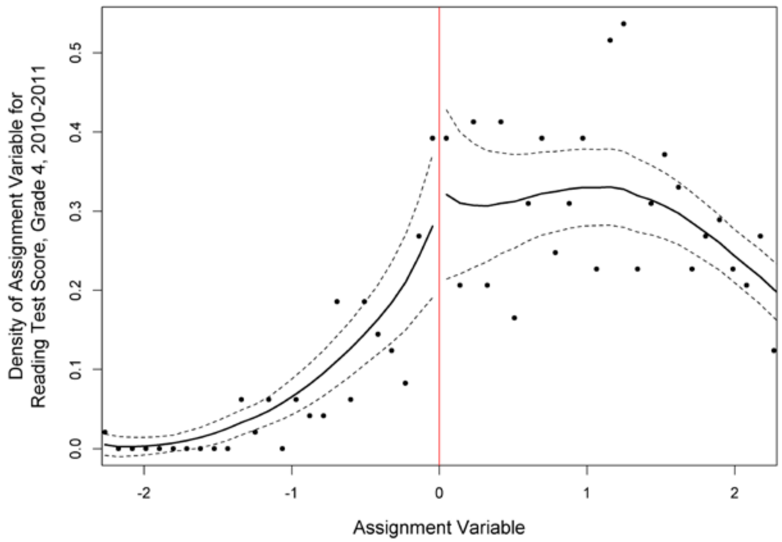


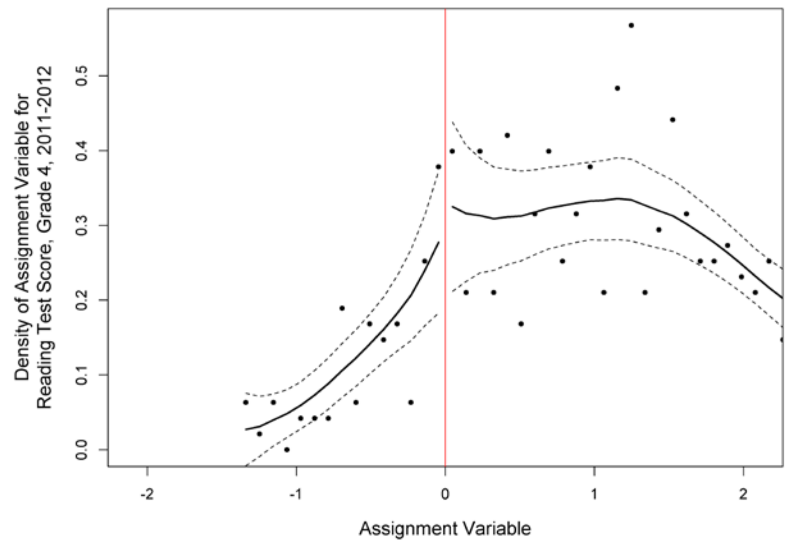


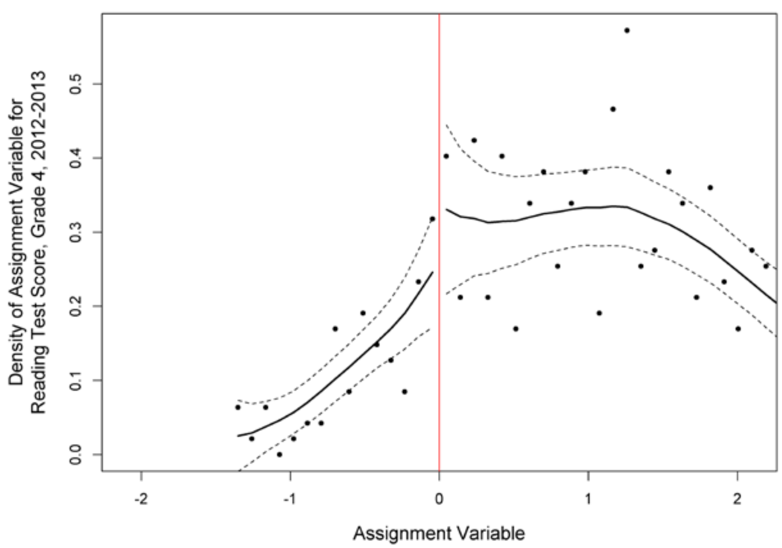
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.59. Density of the assignment variable for reading test score in grade 4, accounting for student mobility



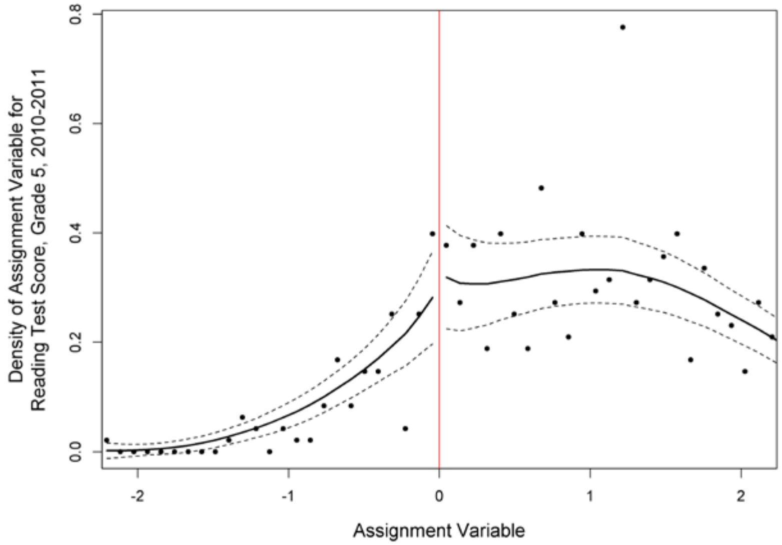


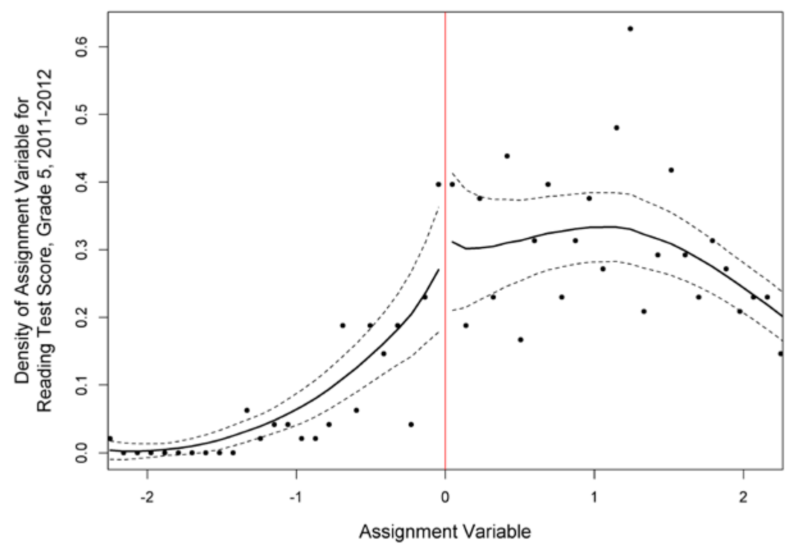


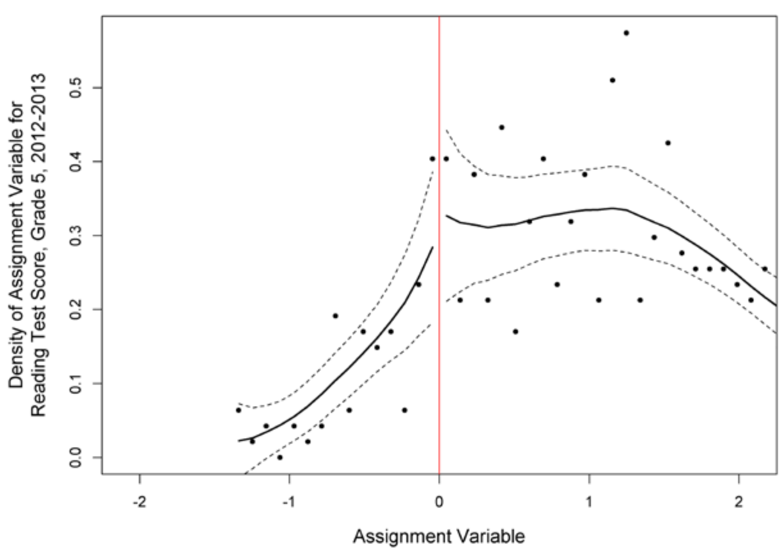
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.60. Density of the assignment variable for reading test score in grade 5, accounting for student mobility



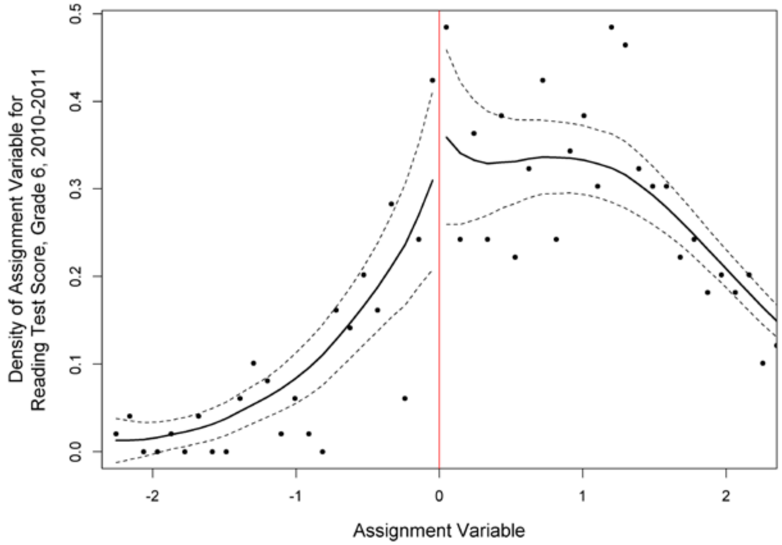


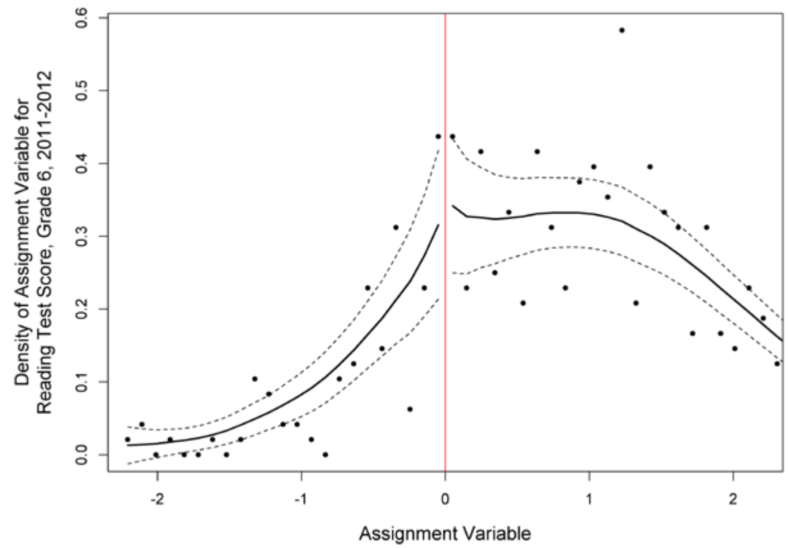


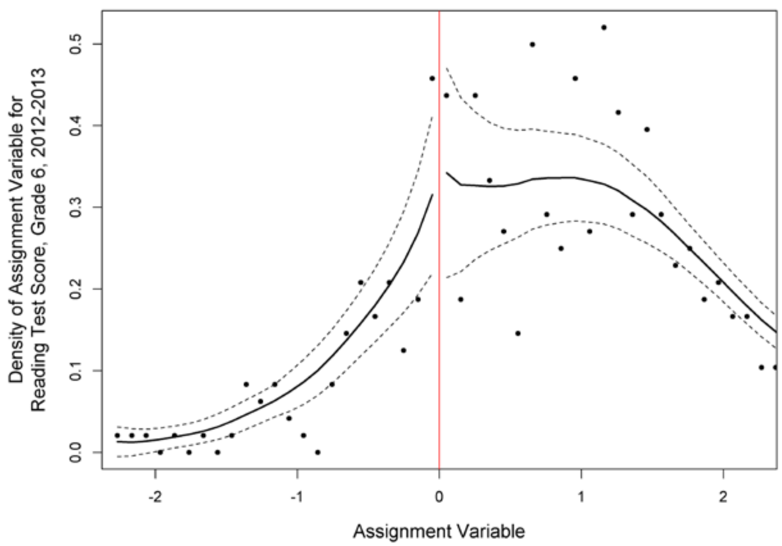
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.61. Density of the assignment variable for reading test score in grade 6, accounting for student mobility



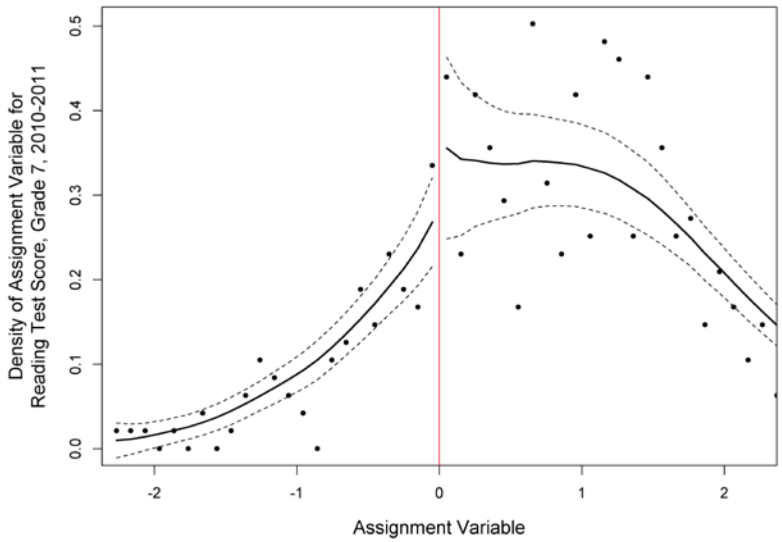


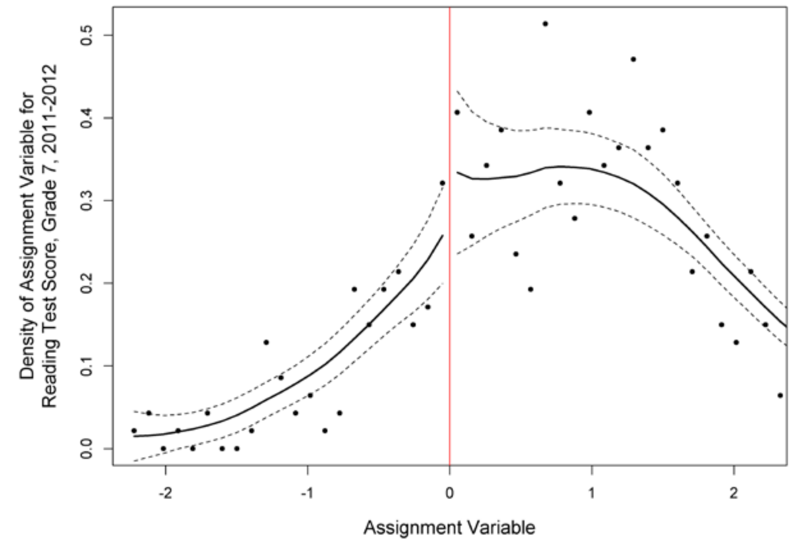


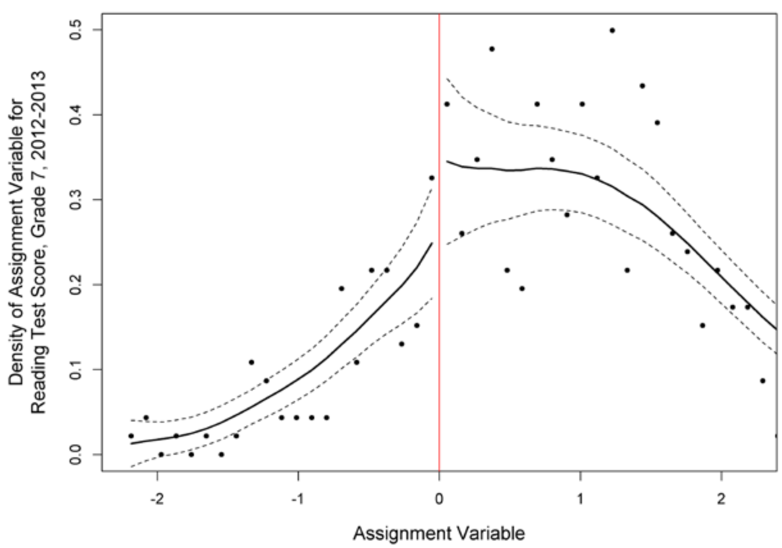
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.62. Density of the assignment variable for reading test score in grade 7, accounting for student mobility



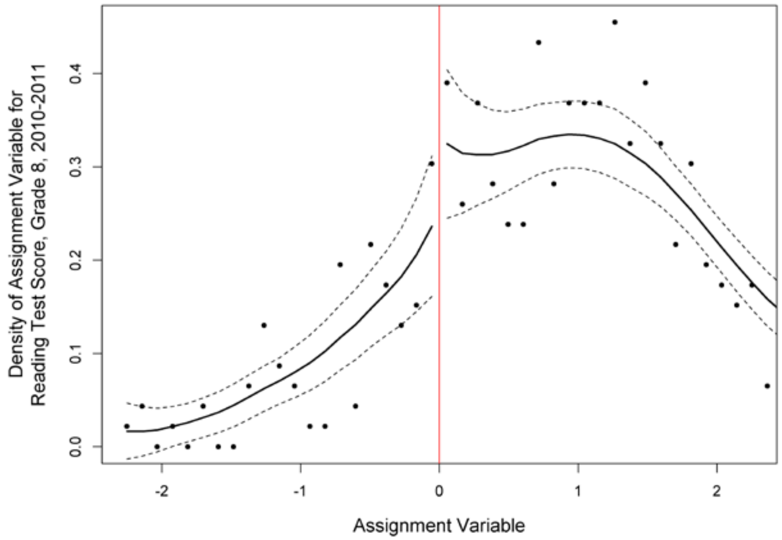


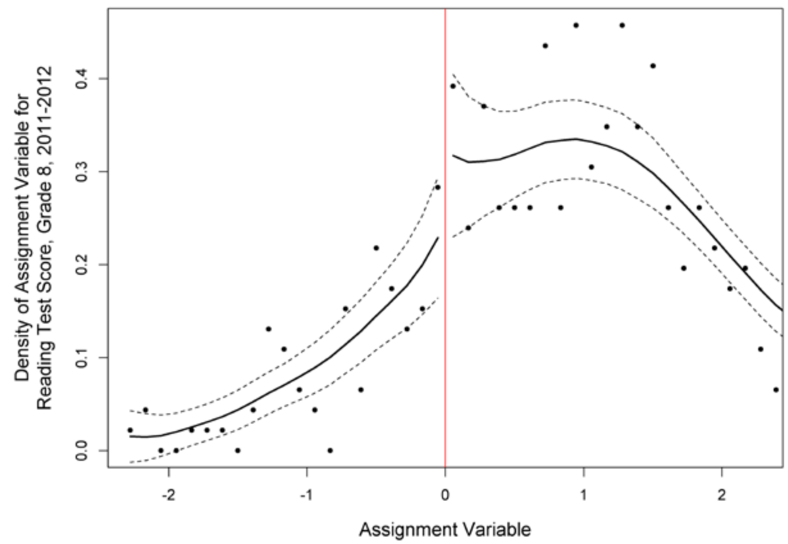


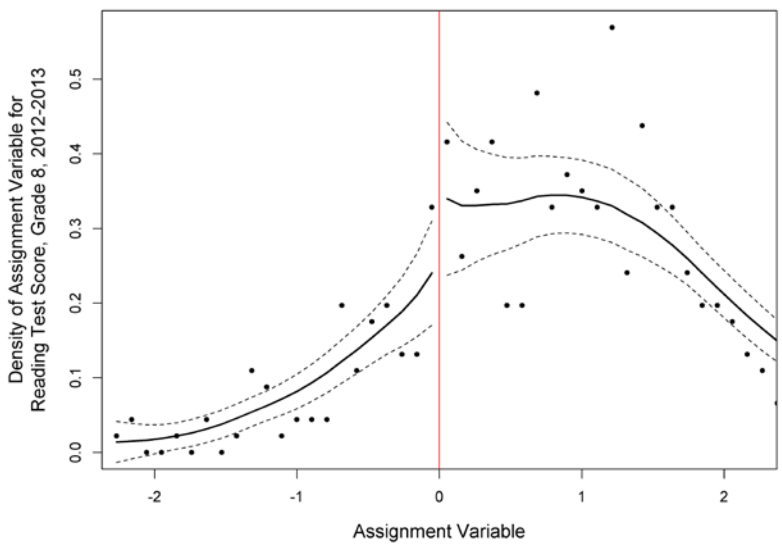
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.63. Density of the assignment variable for reading test score in grade 8, accounting for student mobility



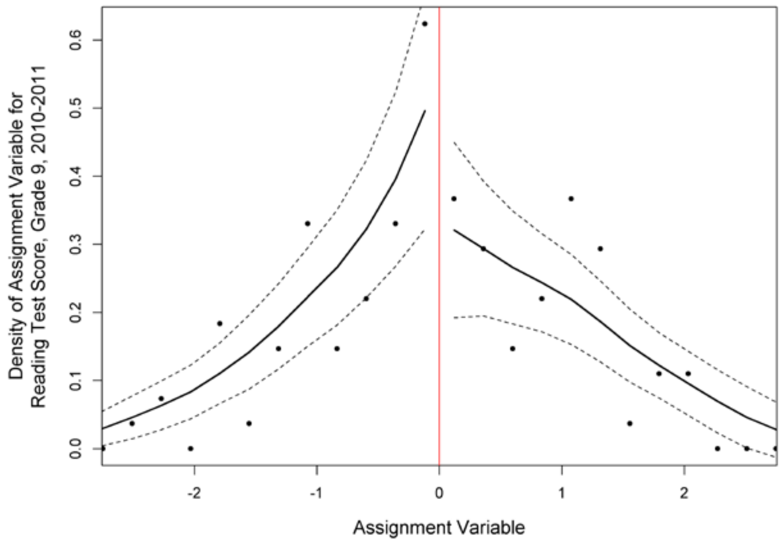


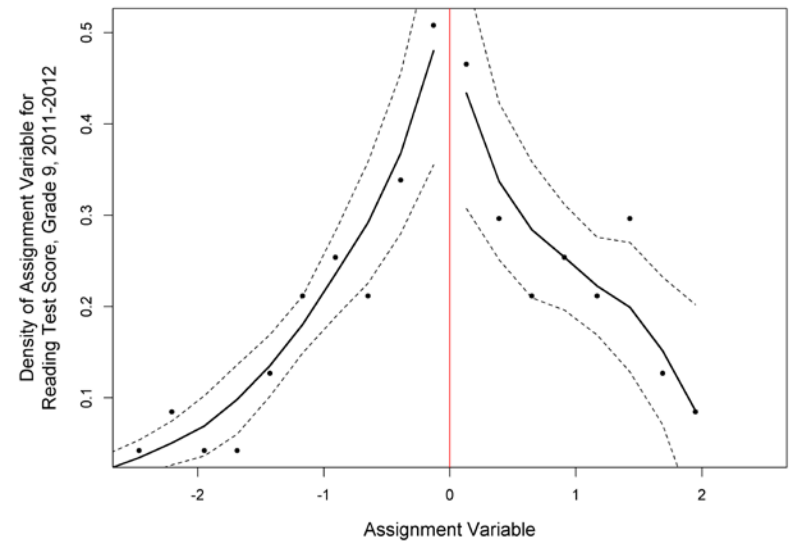


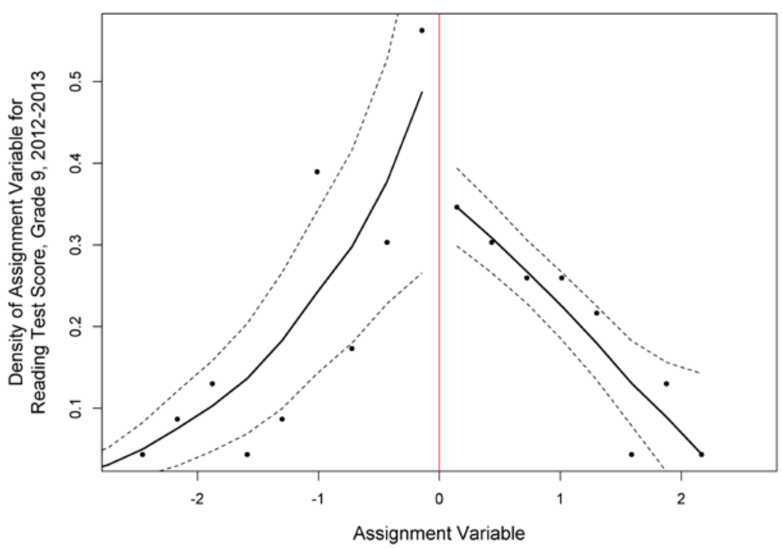
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.64. Density of the assignment variable for reading test score in grade 9, accounting for student mobility



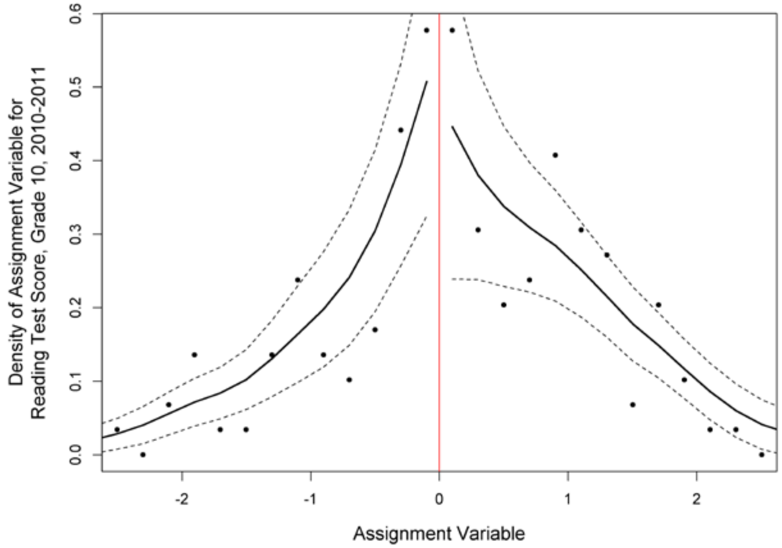


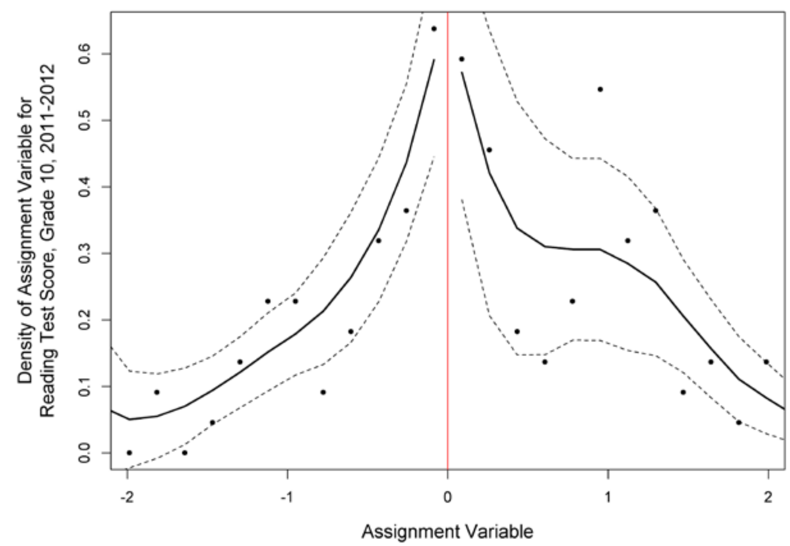


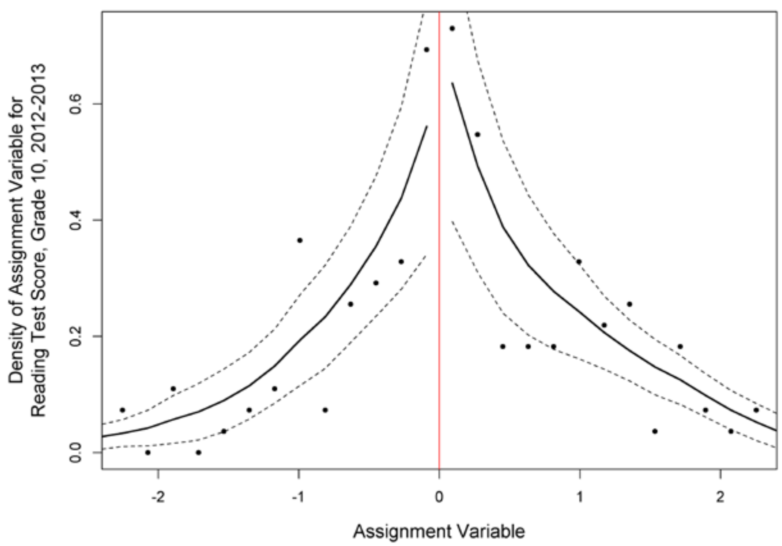
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.65. Density of the assignment variable for reading test score in grade 10, accounting for student mobility



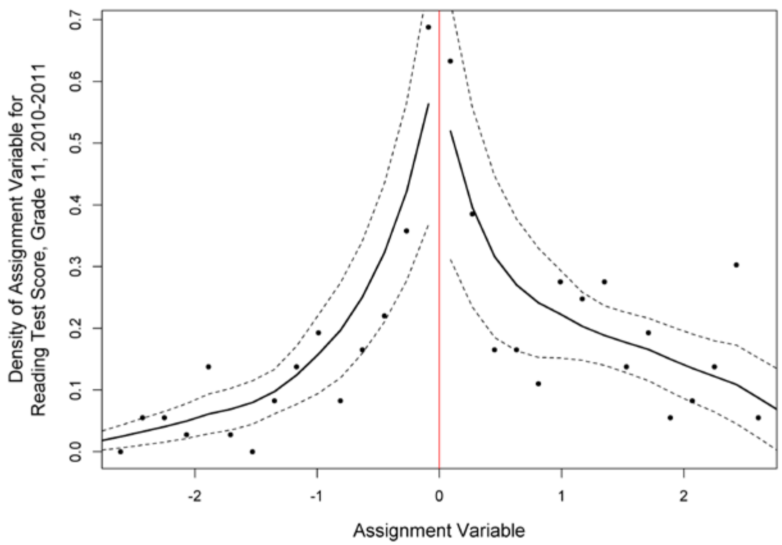


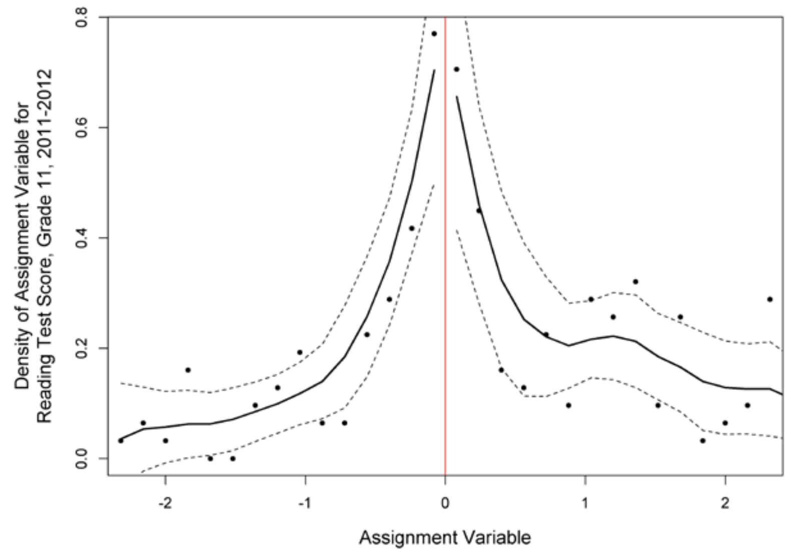


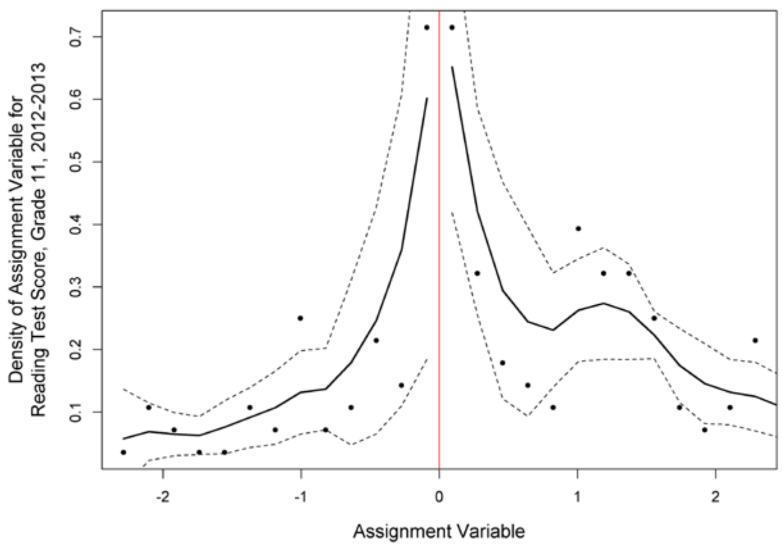
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.66. Density of the assignment variable for reading test score in grade 11, accounting for student mobility



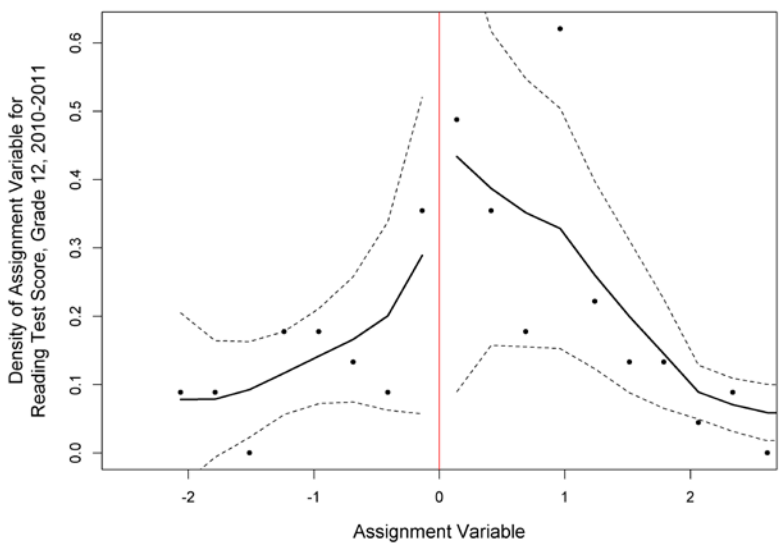


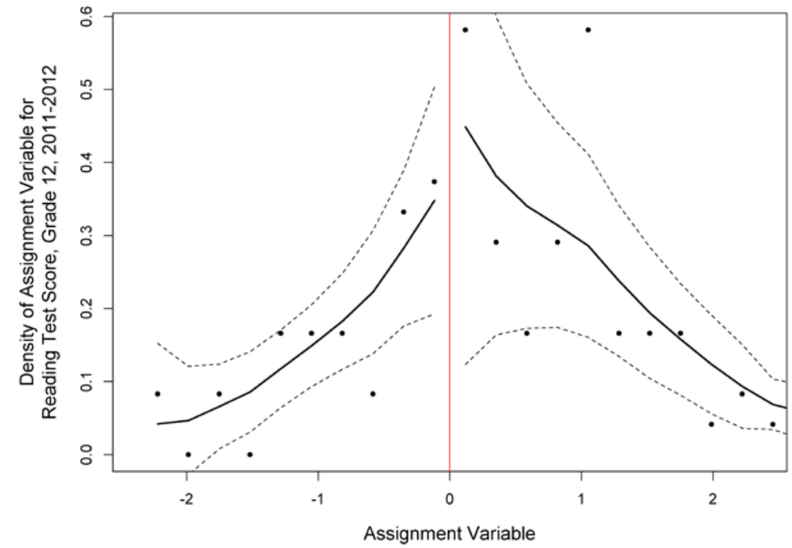


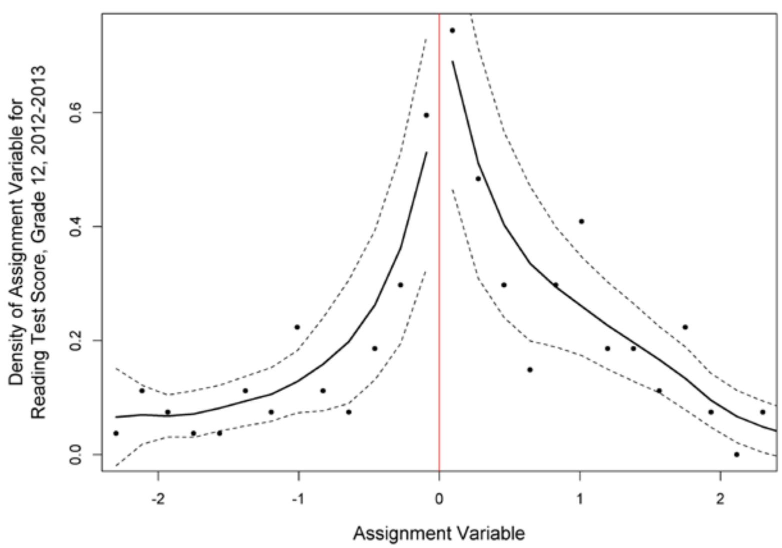
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.67. Density of the assignment variable for reading test score in grade 12, accounting for student mobility



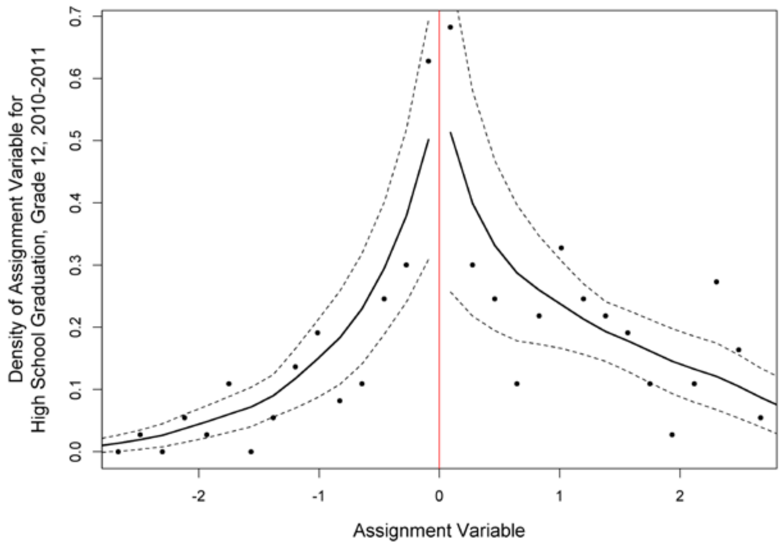


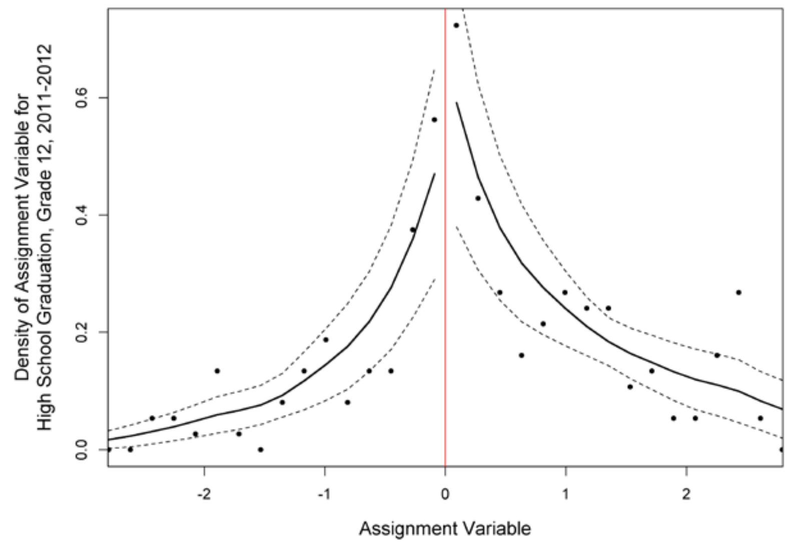


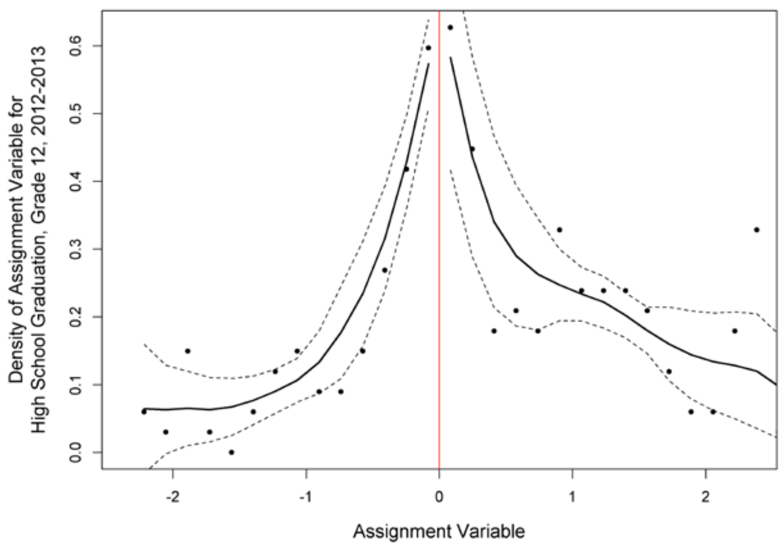
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.68. Density of the assignment variable for high school graduation, accounting for student mobility



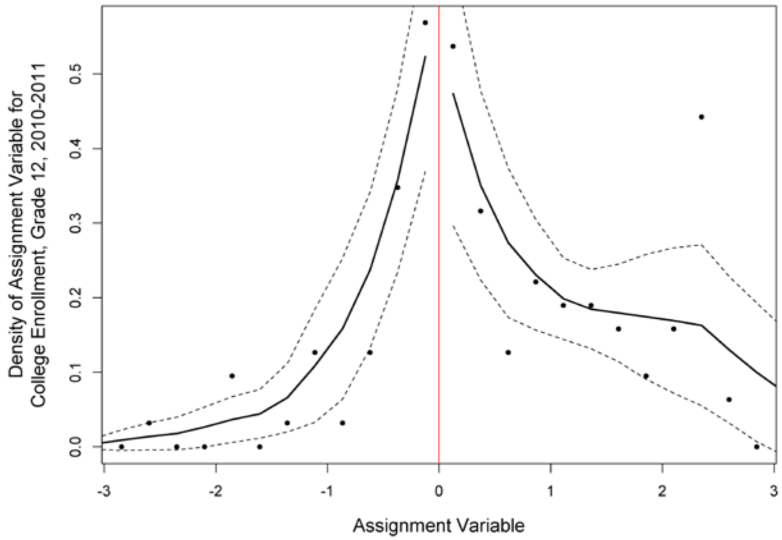


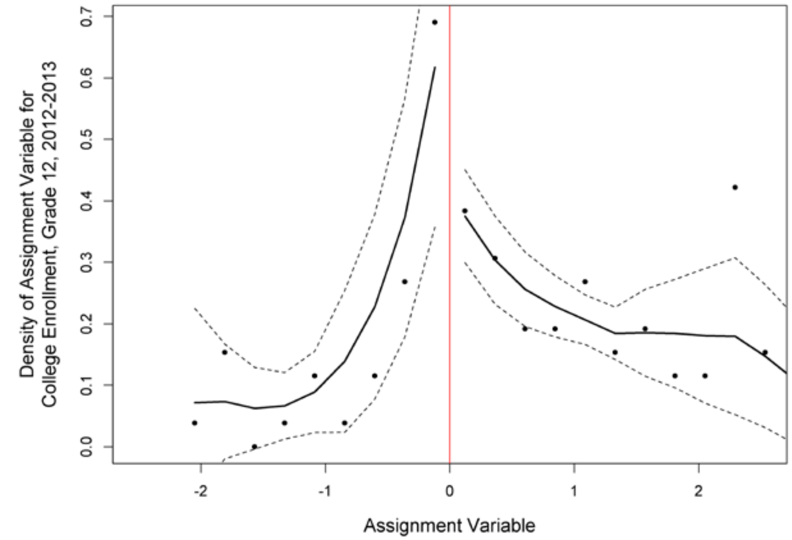


Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines).

Figure A.69. Density of the assignment variable for college enrollment, accounting for student mobility

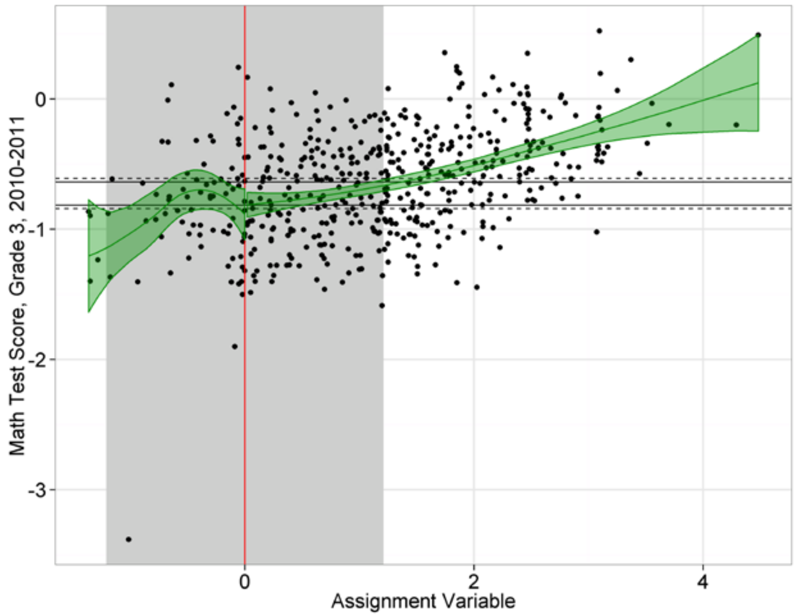


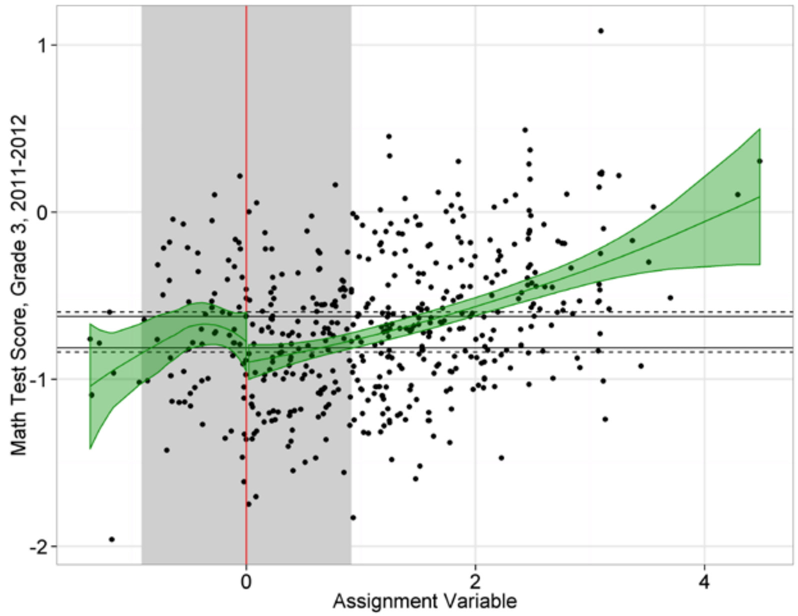


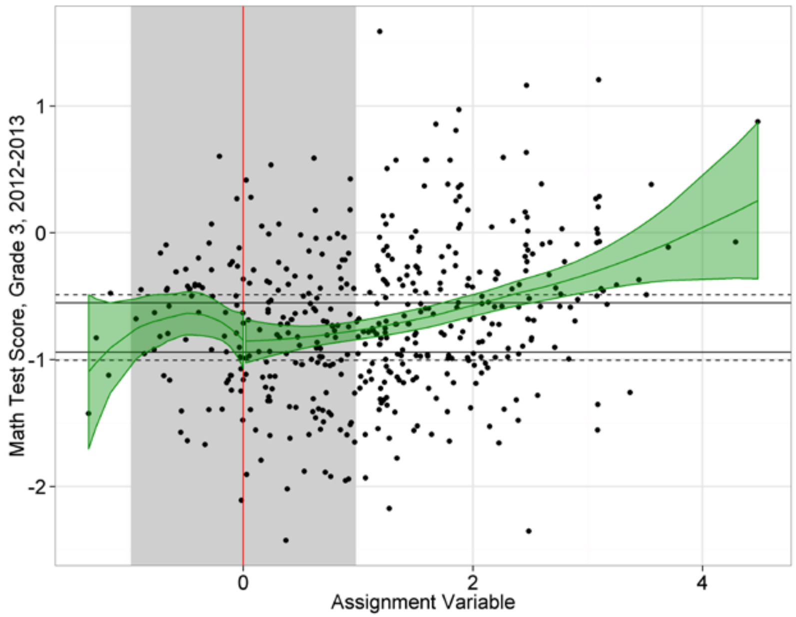
Source: State and district administrative records.

Notes: We estimated the density of the assignment variable using the DCdensity function from the rdd package in R (Dimmery 2013). Each figure displays the cutoff value of the assignment value (vertical red line at 0), estimated densities for each value of the assignment variable (solid black lines), and upper and lower confidence bounds (dashed lines). We omitted the figure for 2011–2012 because we were not able to estimate an impact for college enrollment in this year due to insufficient sample sizes.

Figure A.70. Math test score in grade 3, accounting for student mobility



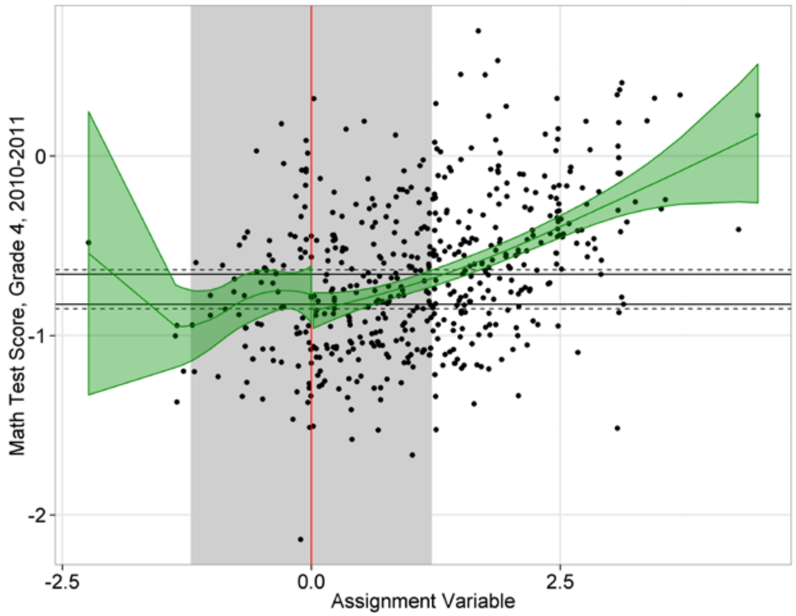


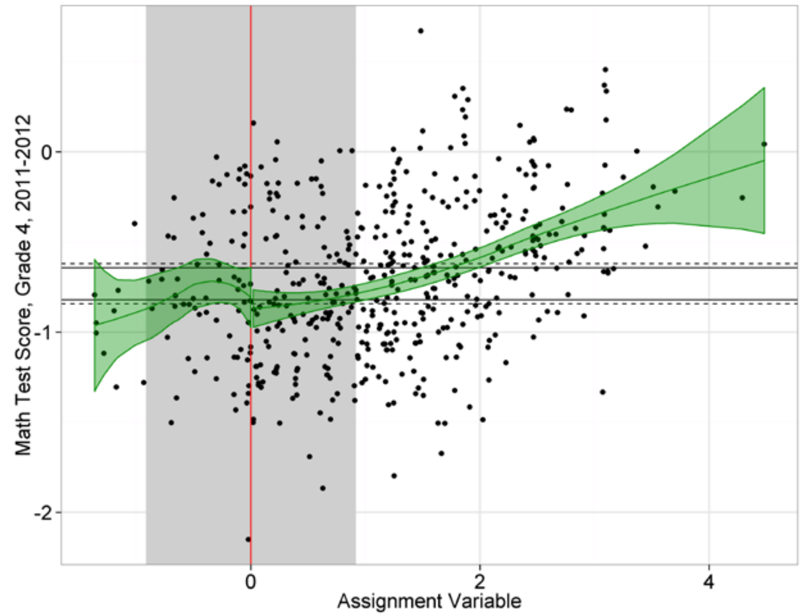


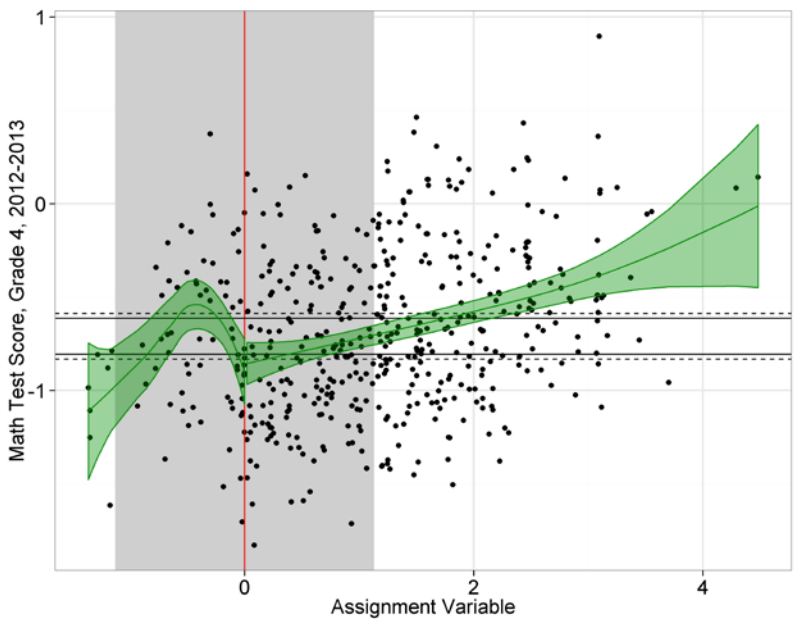
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.71. Math test score in grade 4, accounting for student mobility



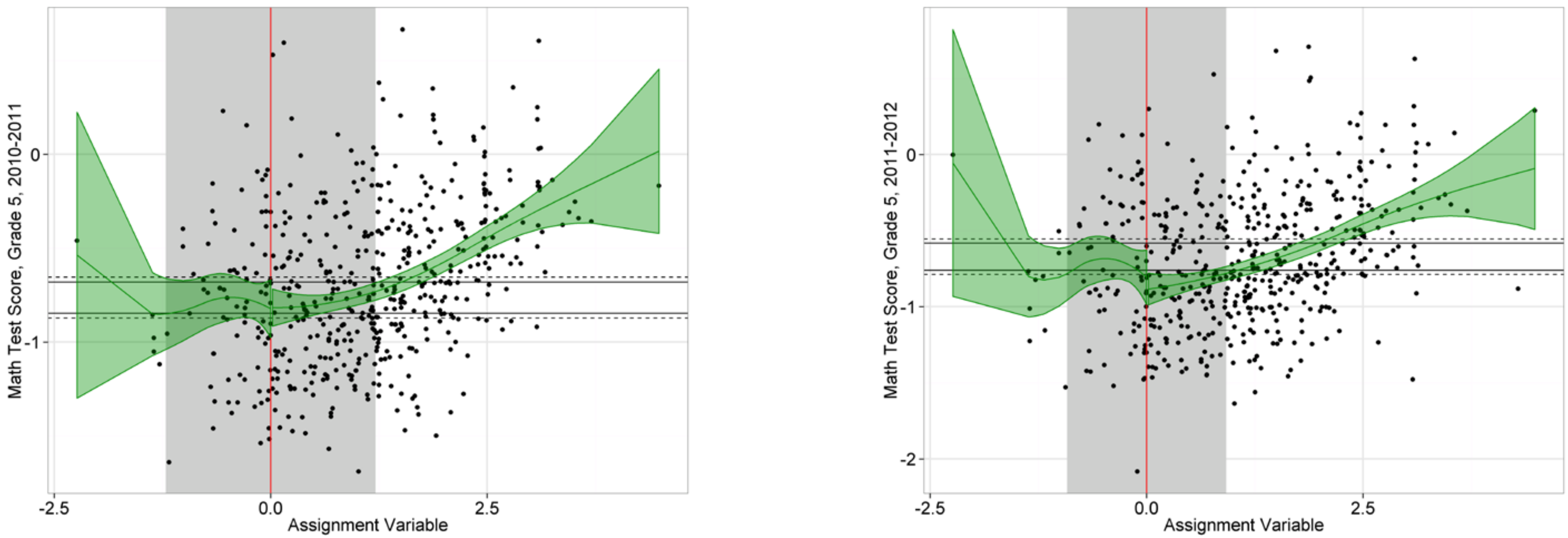


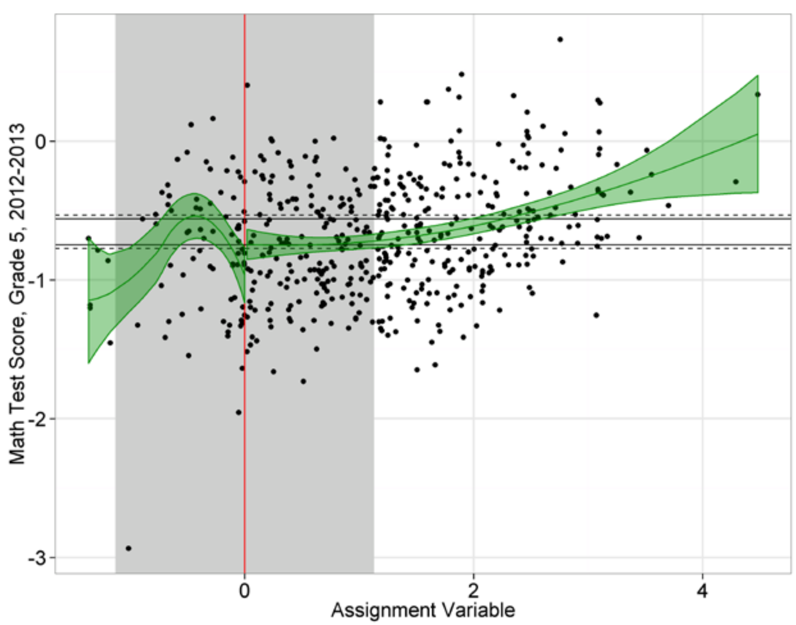


Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.72. Math test score in grade 5, accounting for student mobility

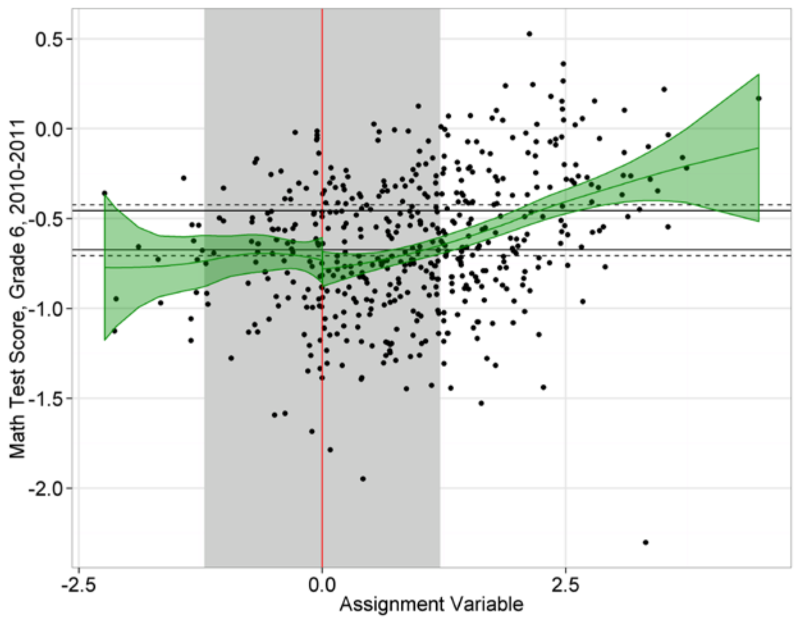


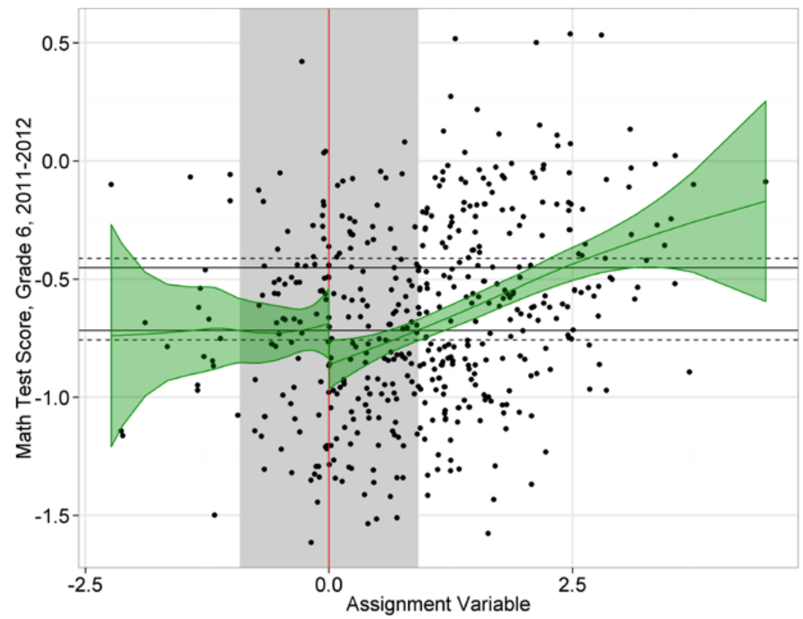


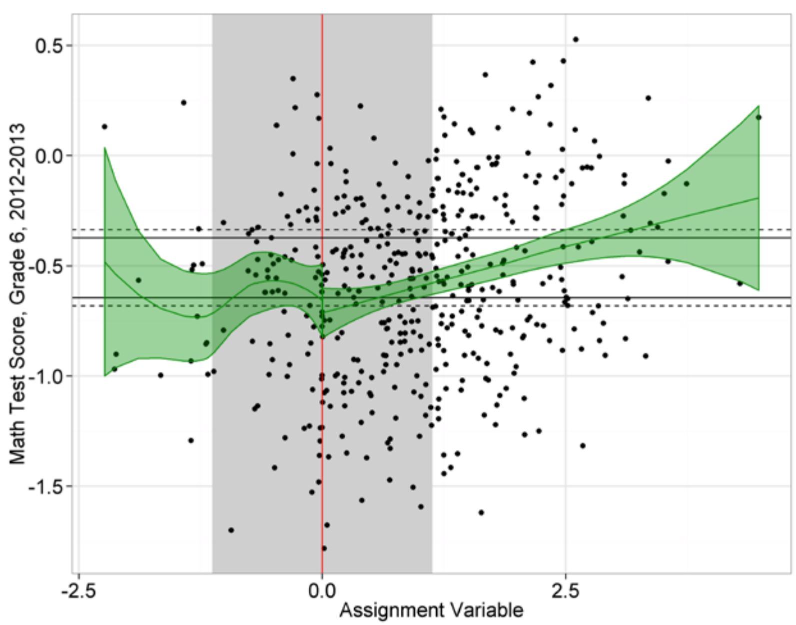
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.73. Math test score in grade 6, accounting for student mobility



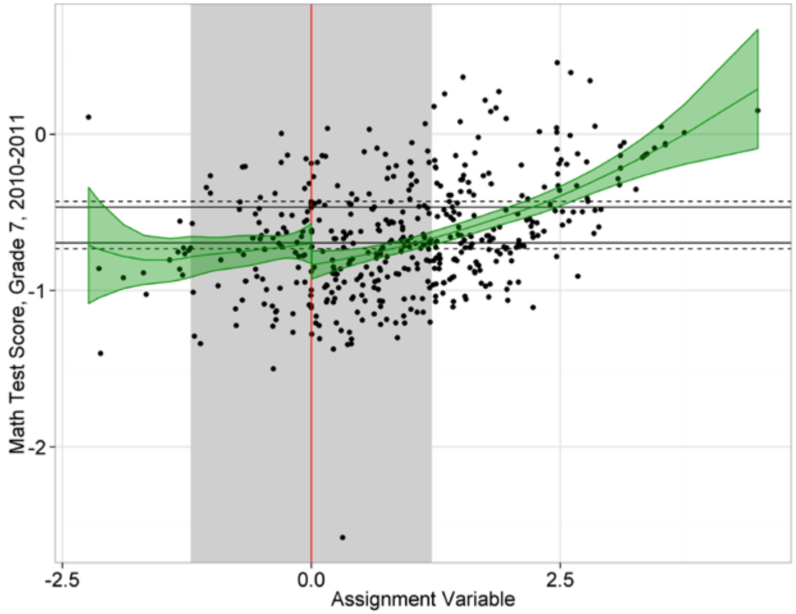


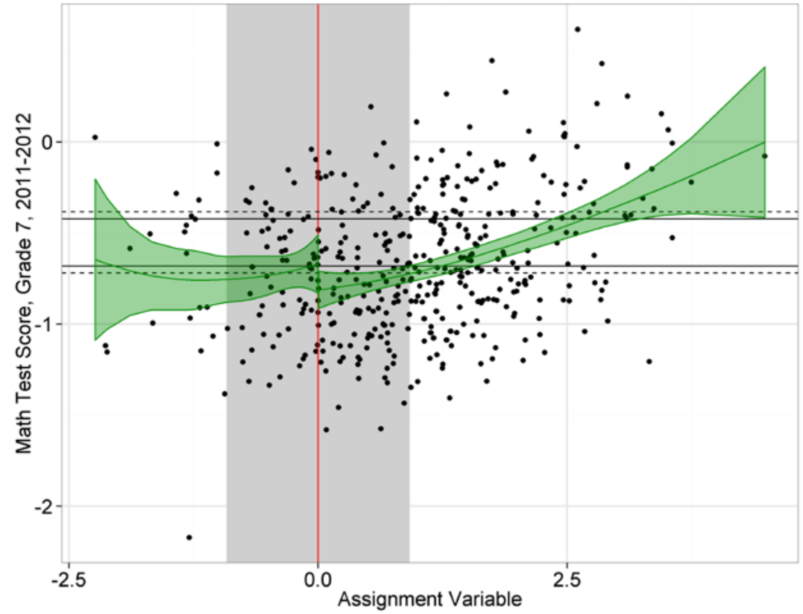


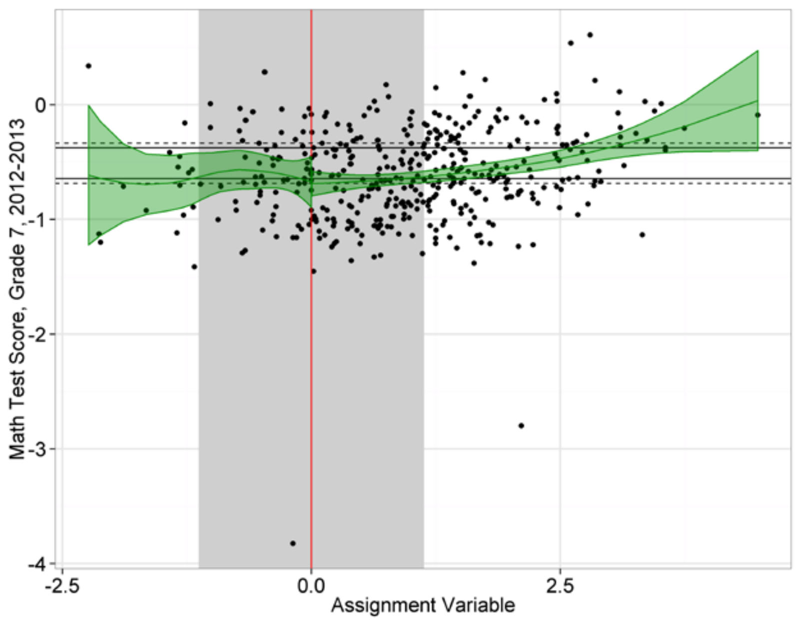
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.74. Math test score in grade 7, accounting for student mobility



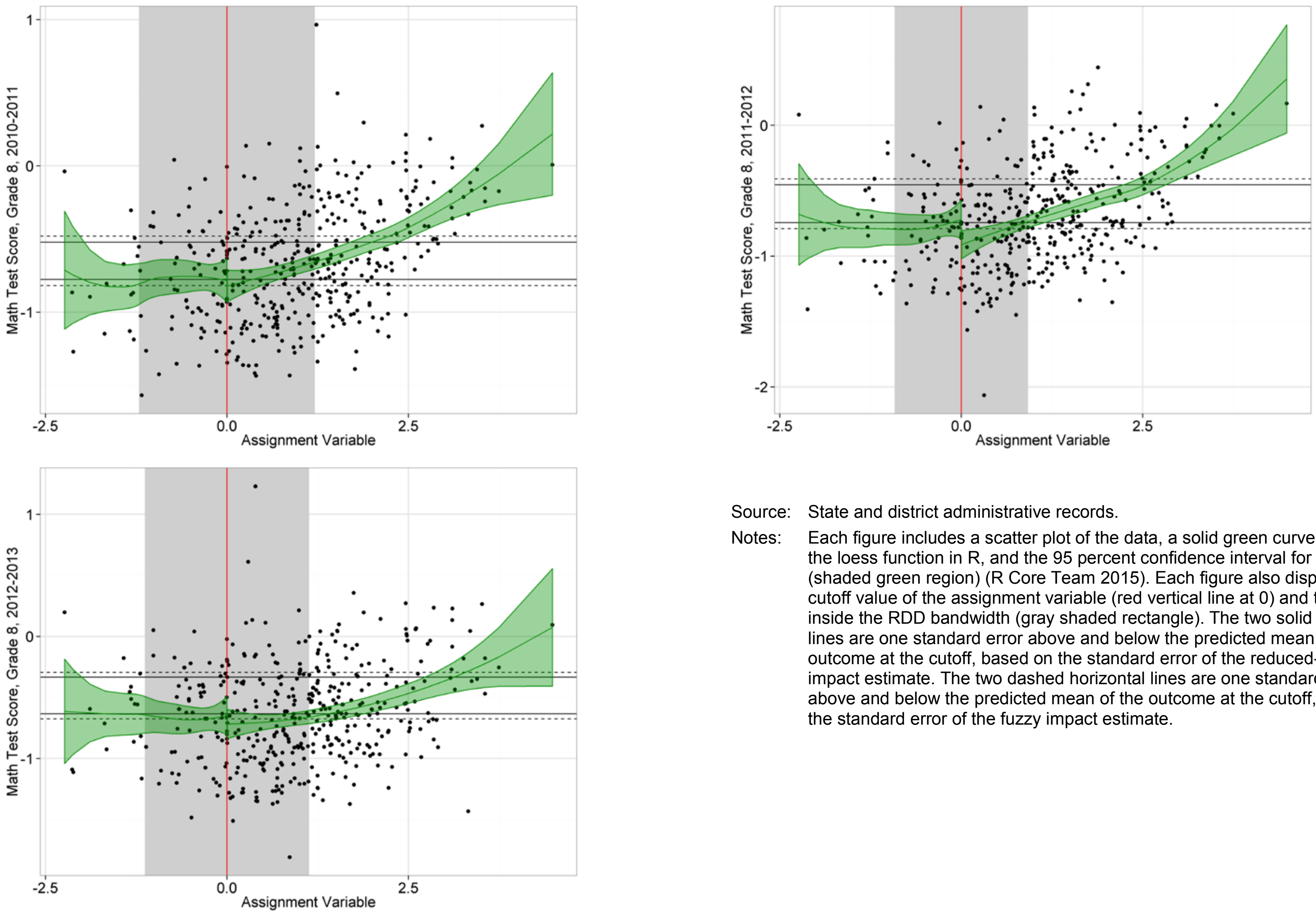




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

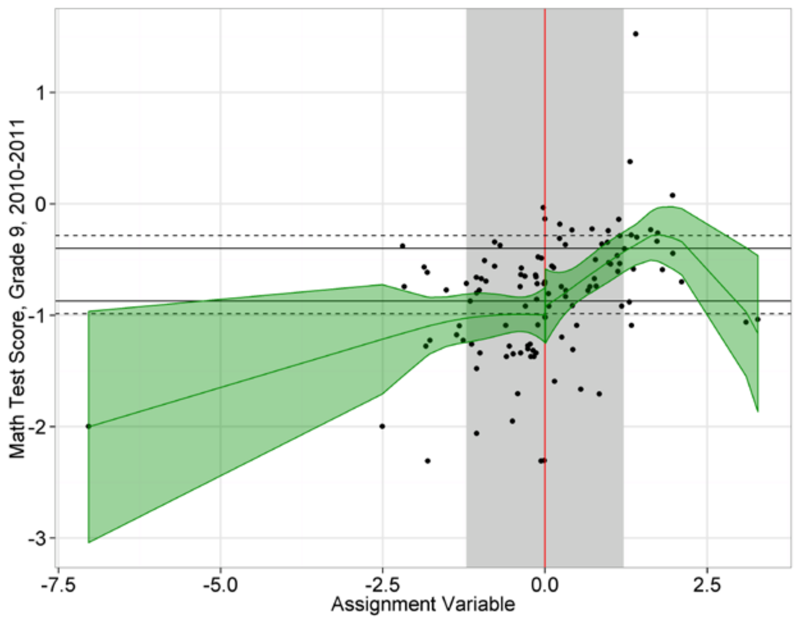
Figure A.75. Math test score in grade 8, accounting for student mobility

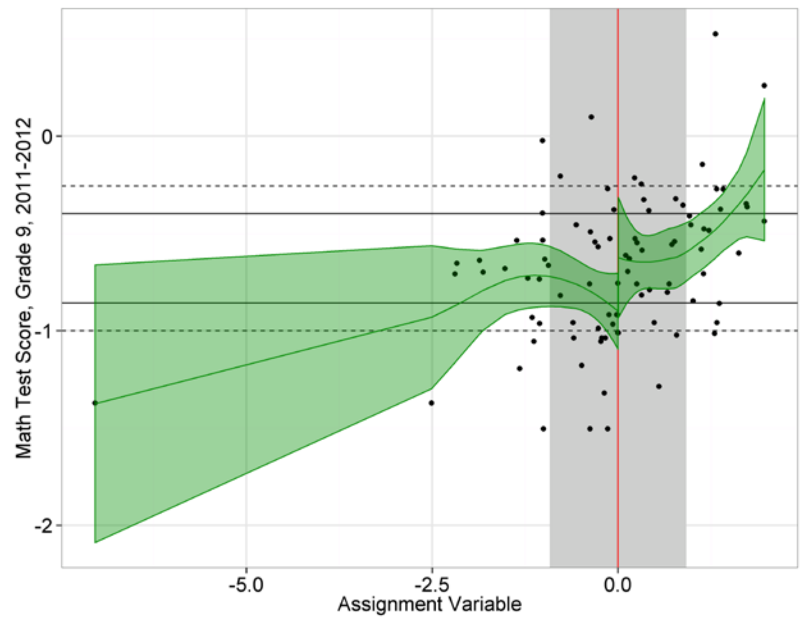


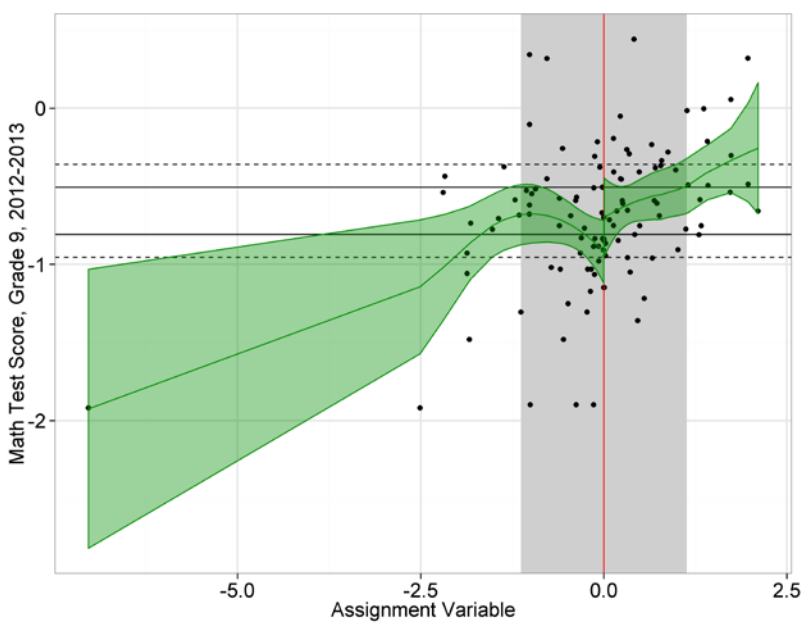
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.76. Math test score in grade 9, accounting for student mobility



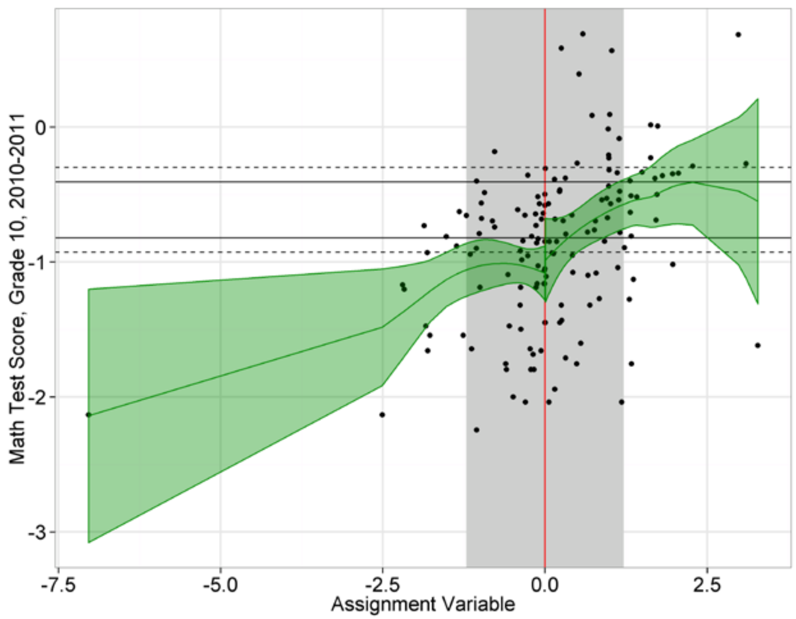


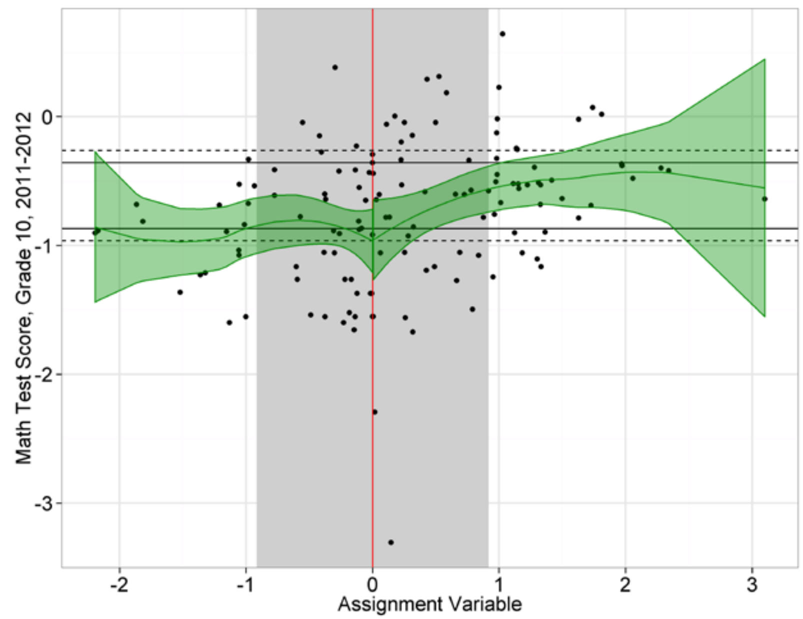


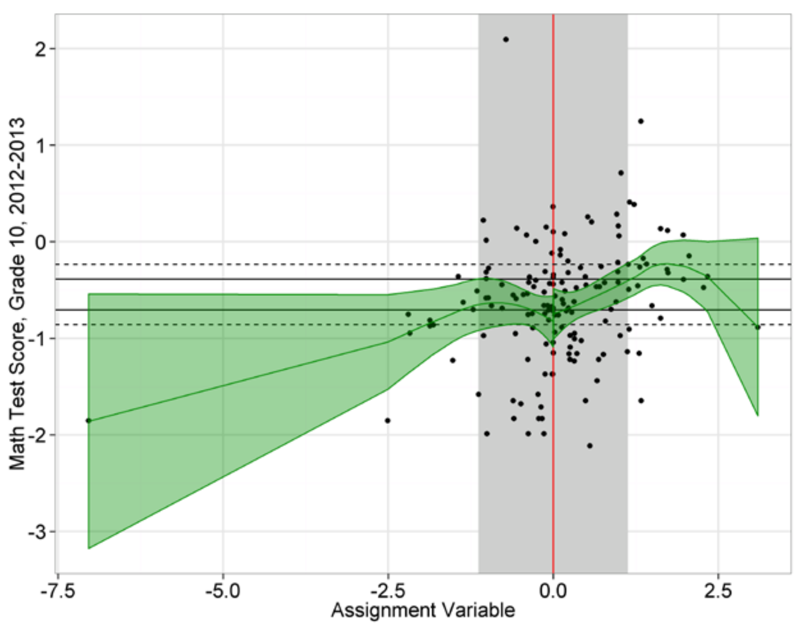
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.77. Math test score in grade 10, accounting for student mobility



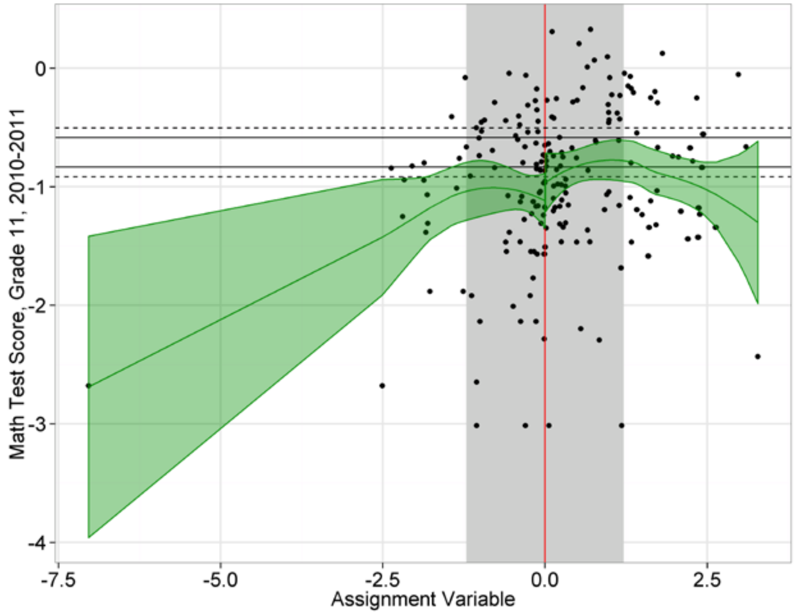


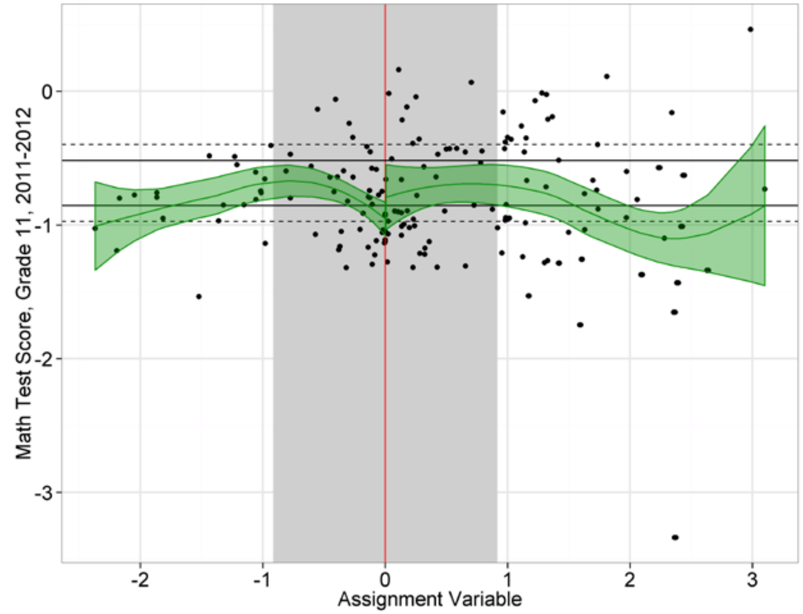


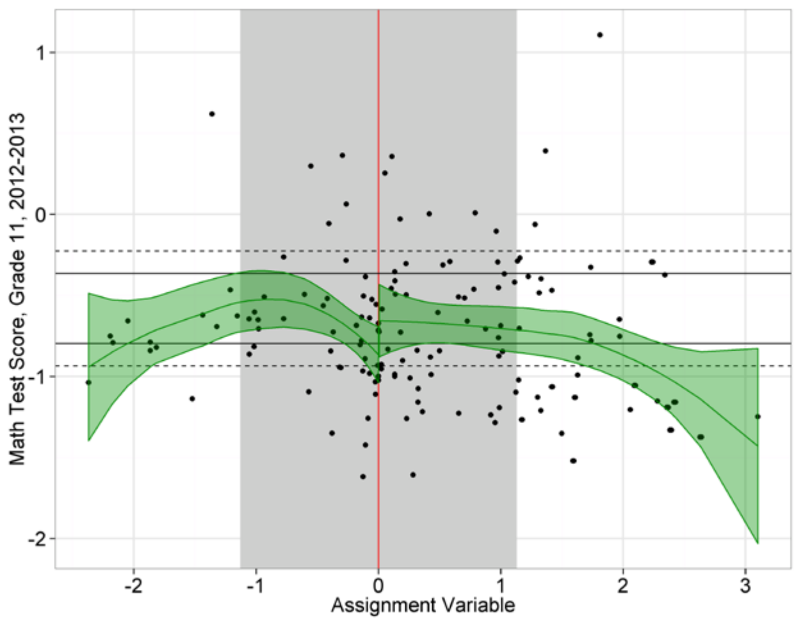
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.78. Math test score in grade 11, accounting for student mobility



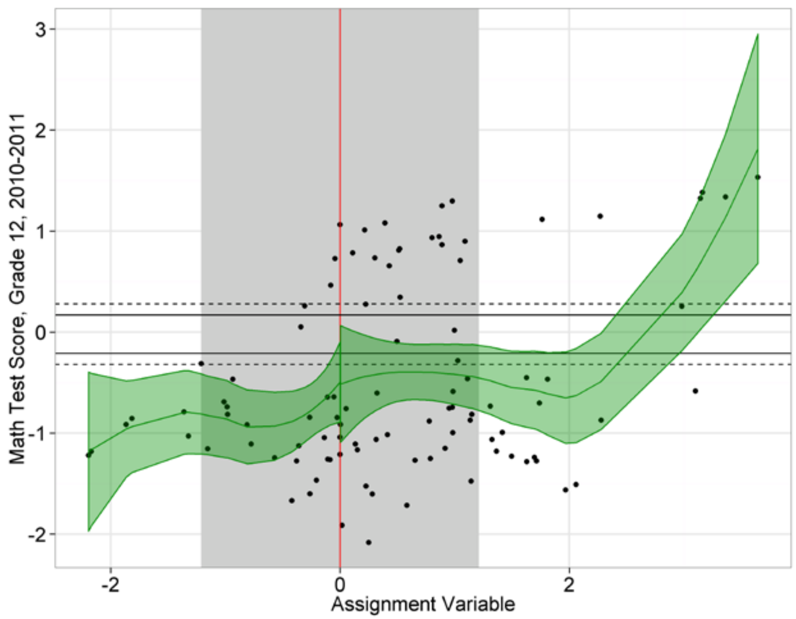


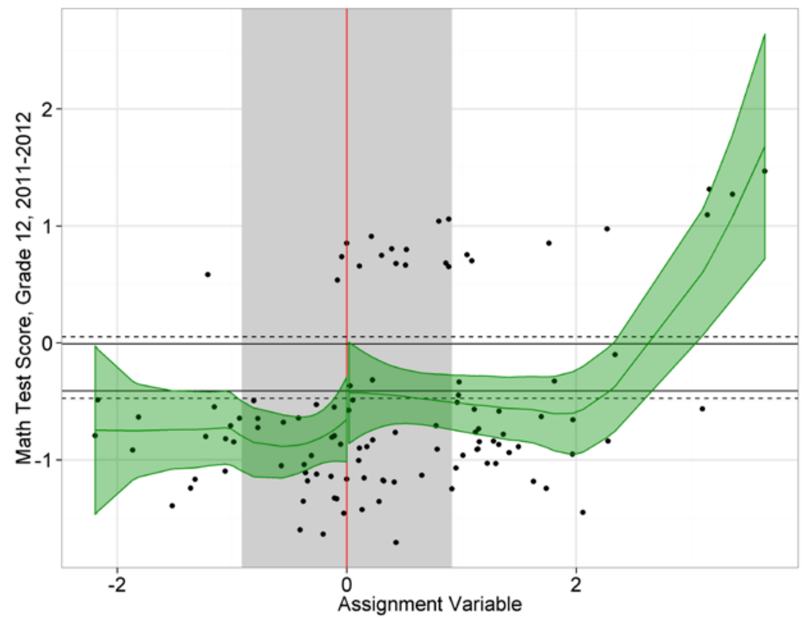


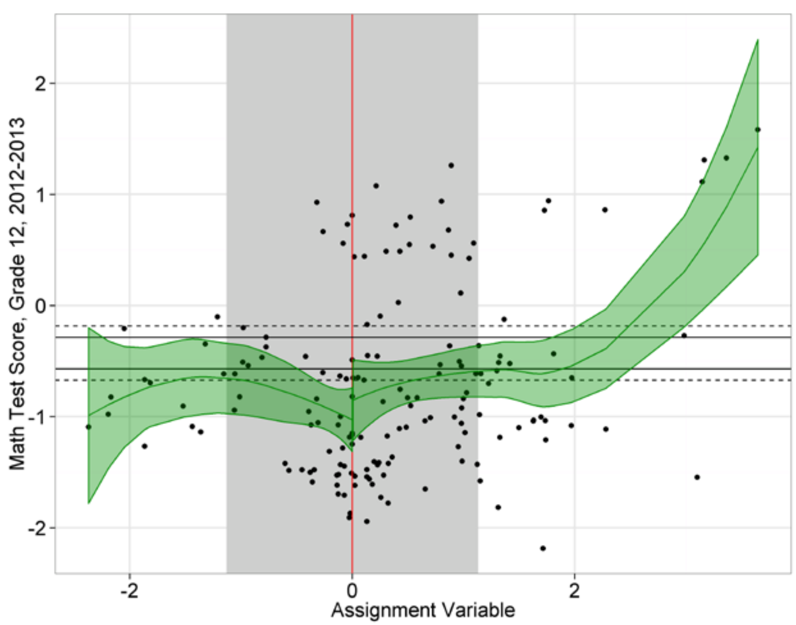
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.79. Math test score in grade 12, accounting for student mobility



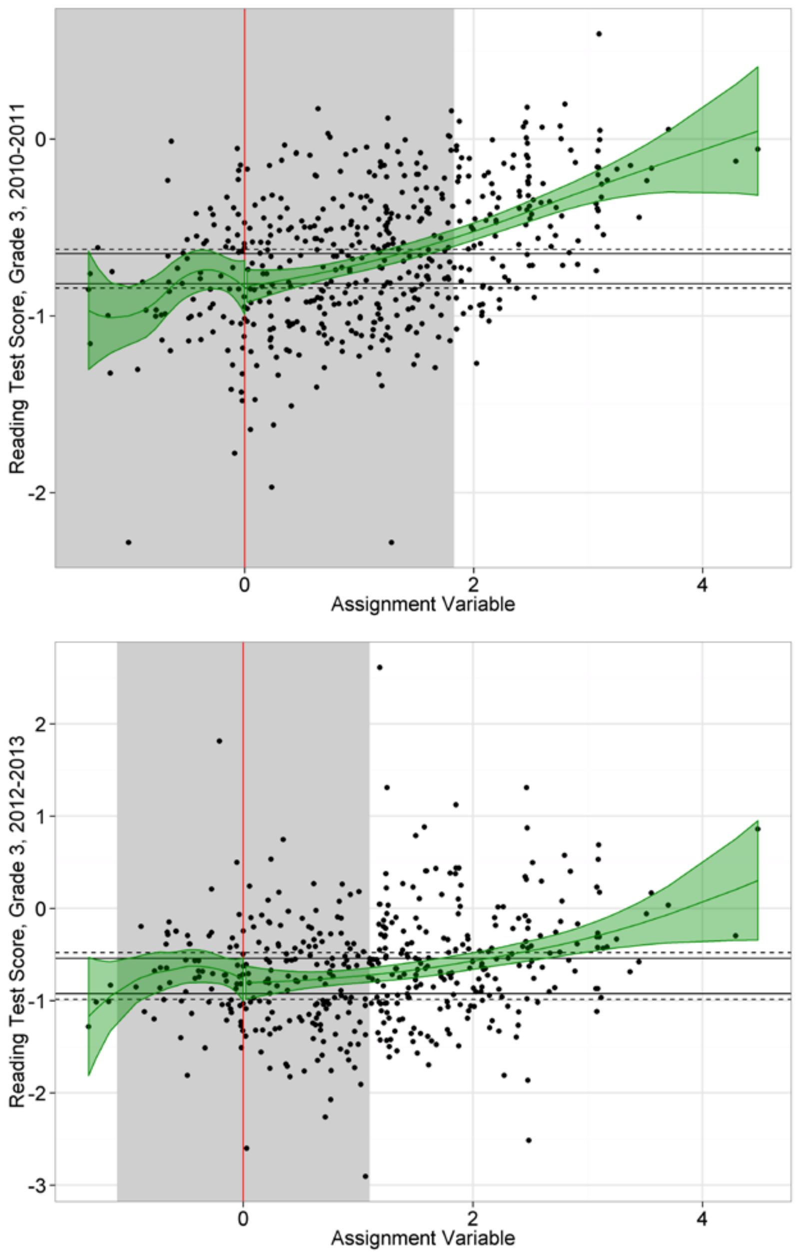




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

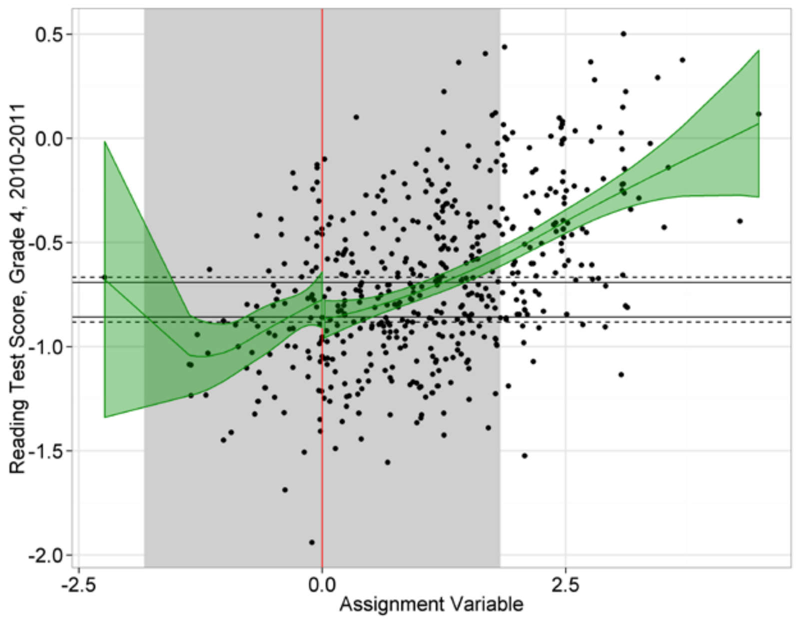
Figure A.80. Reading test score in grade 3, accounting for student mobility

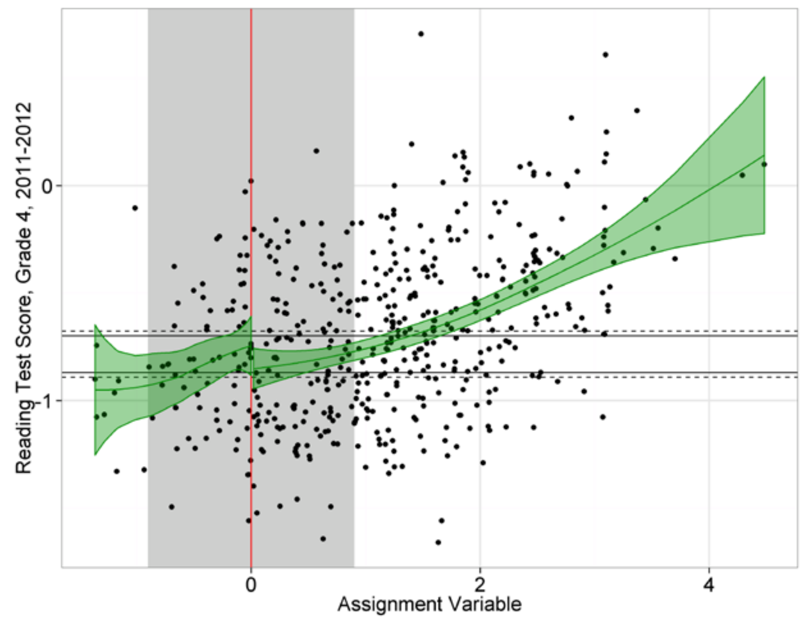


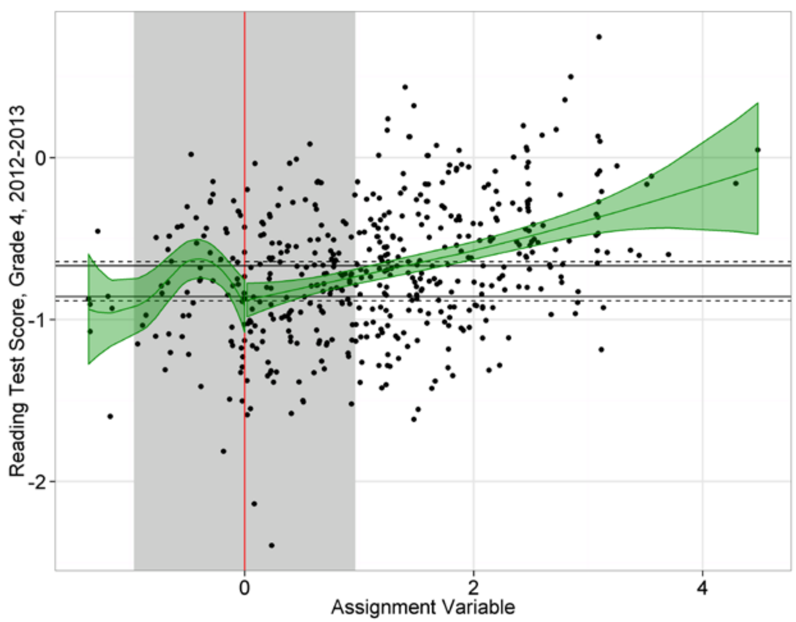
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.81. Reading test score in grade 4, accounting for student mobility



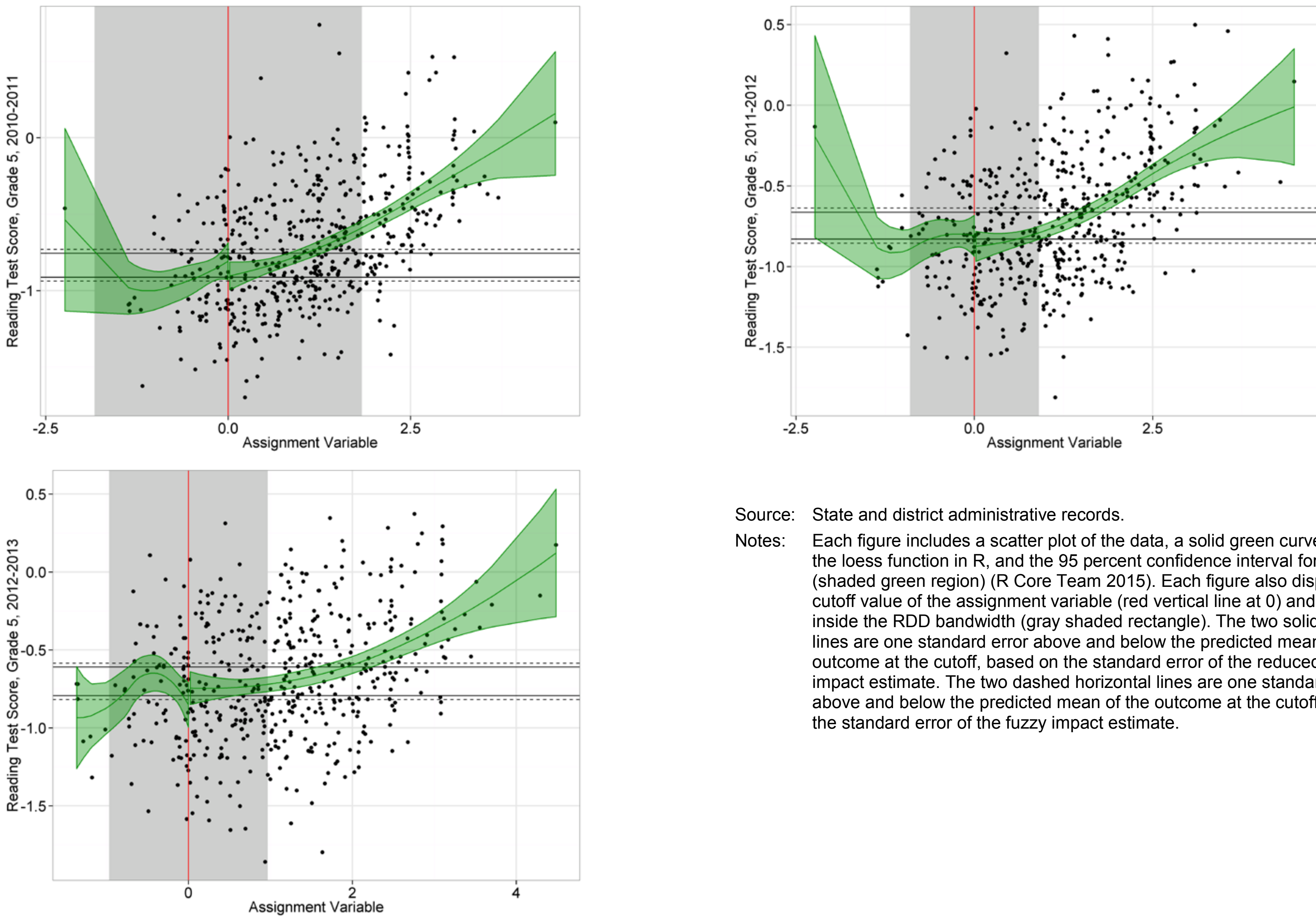




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

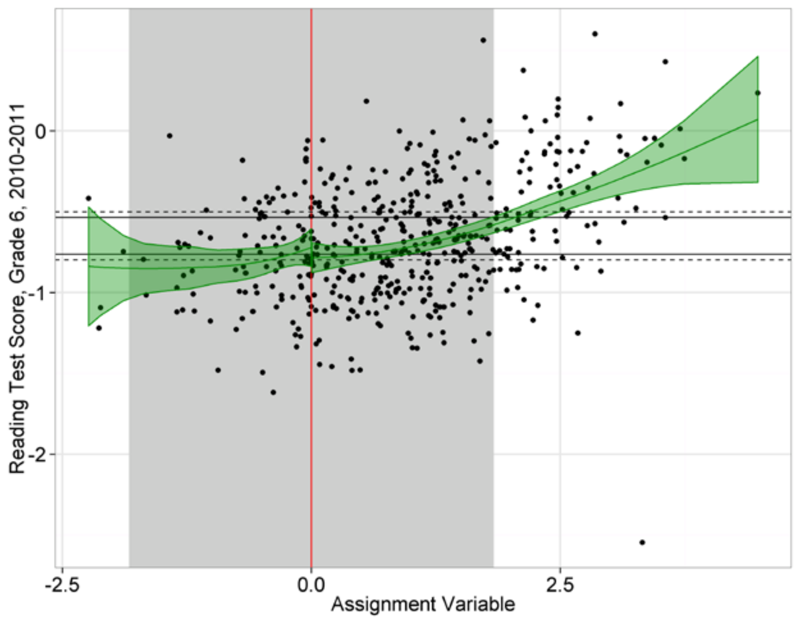
Figure A.82. Reading test score in grade 5, accounting for student mobility

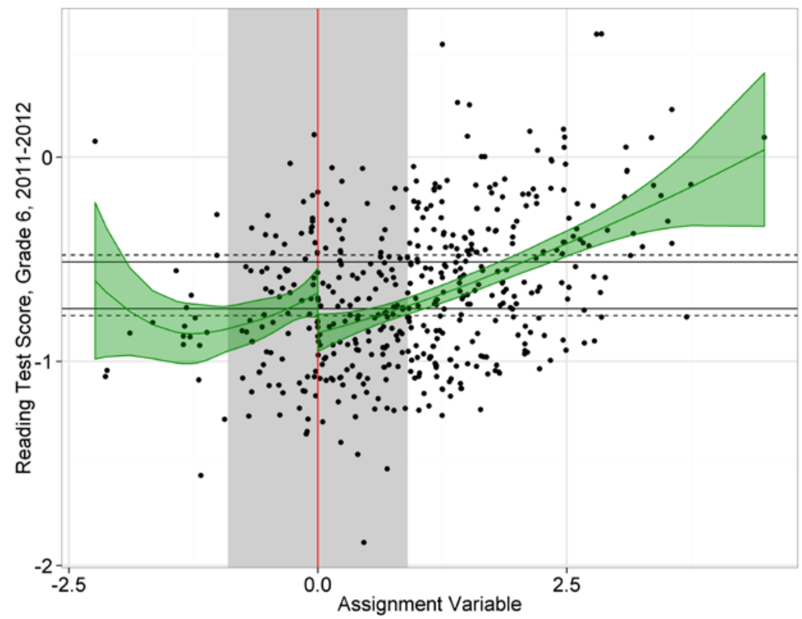


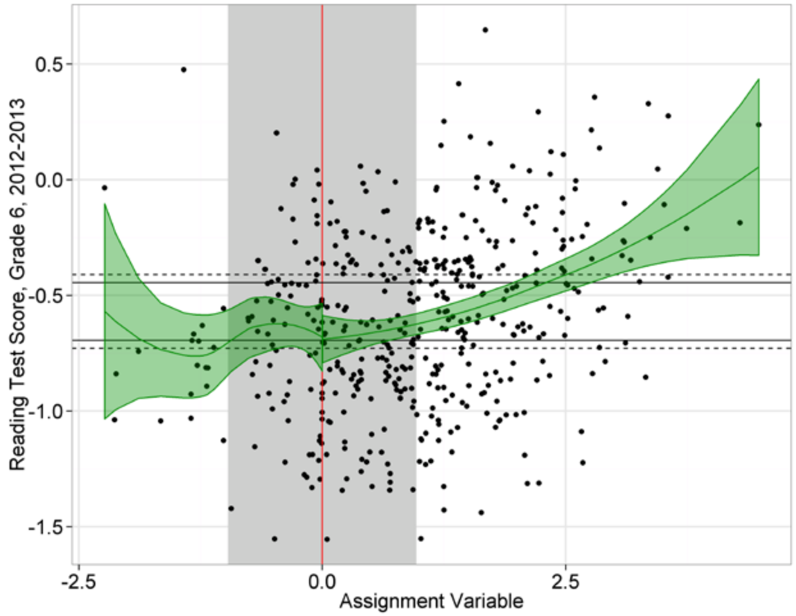
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.83. Reading test score in grade 6, accounting for student mobility



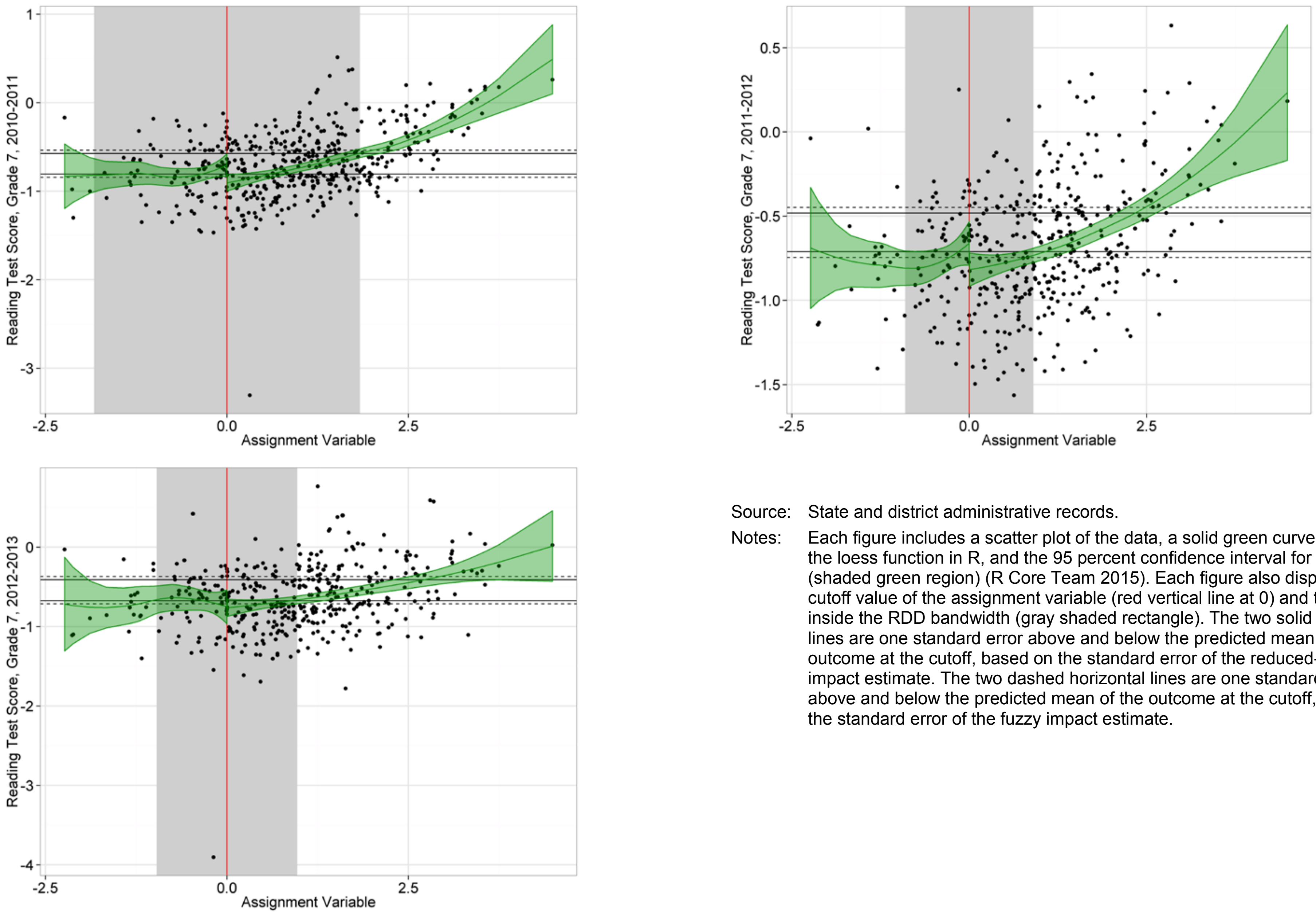




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

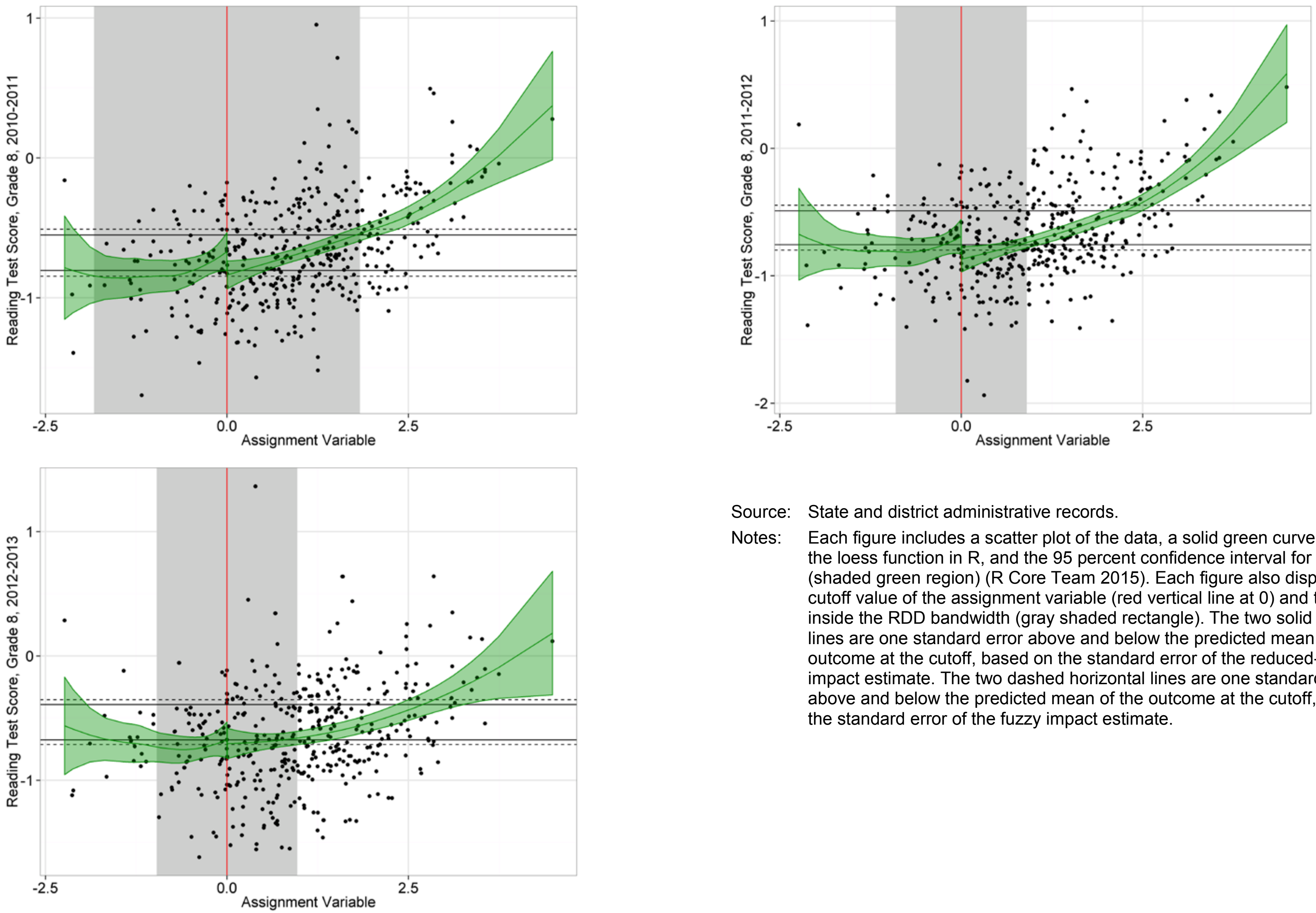
Figure A.84. Reading test score in grade 7, accounting for student mobility



Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

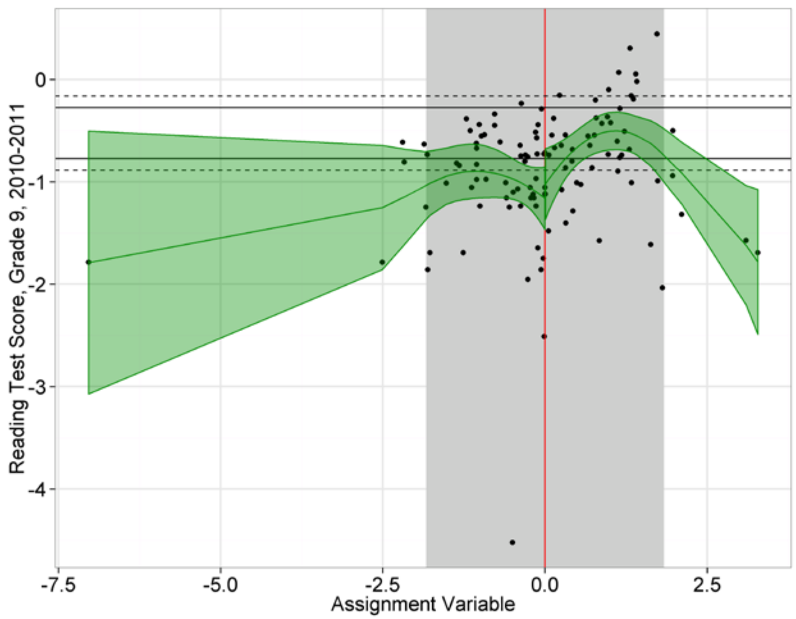
Figure A.85. Reading test score in grade 8, accounting for student mobility

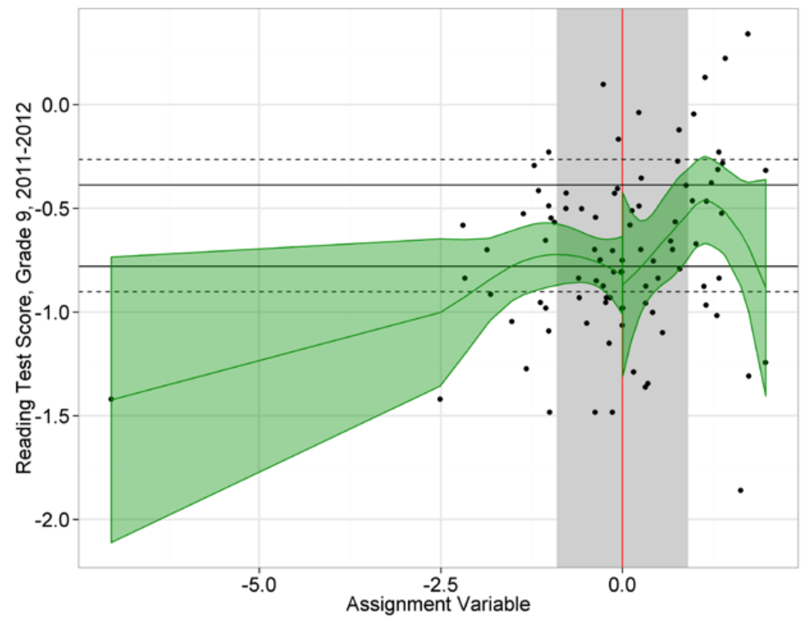


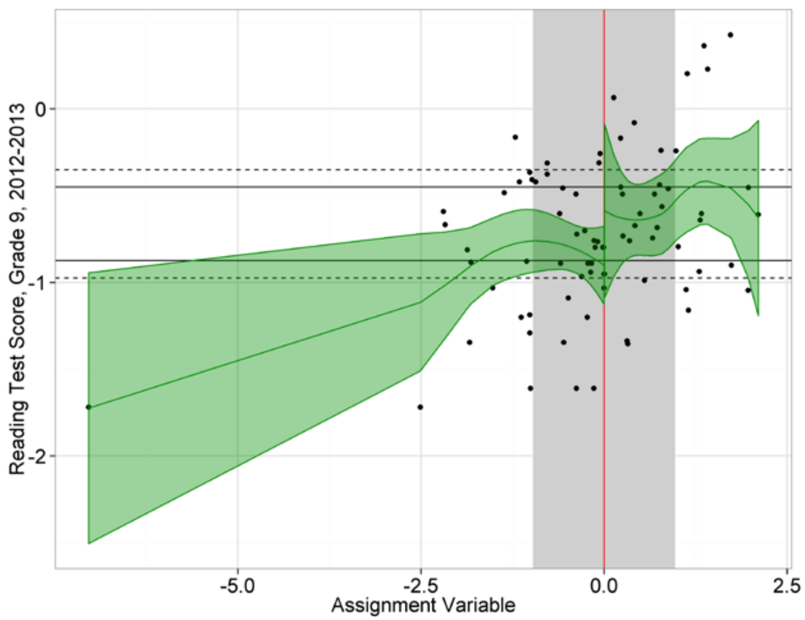
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.86. Reading test score in grade 9, accounting for student mobility



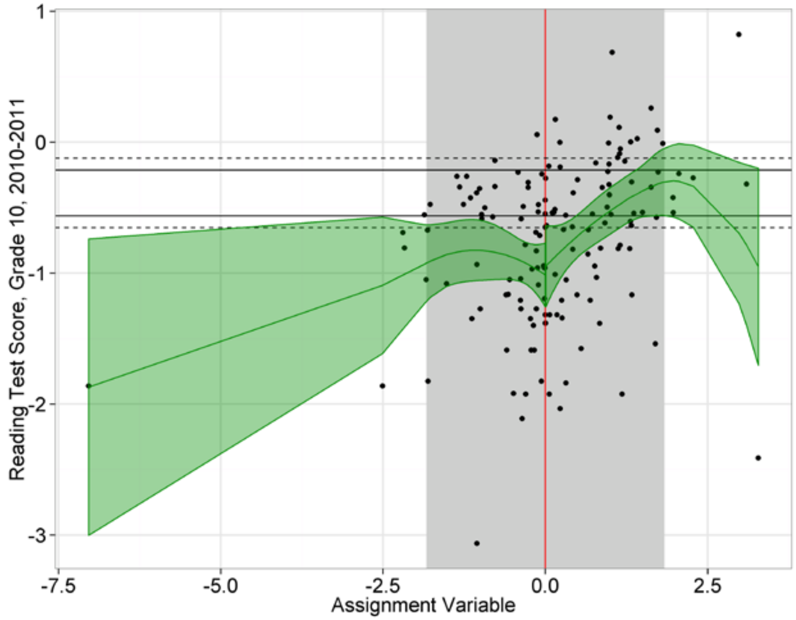


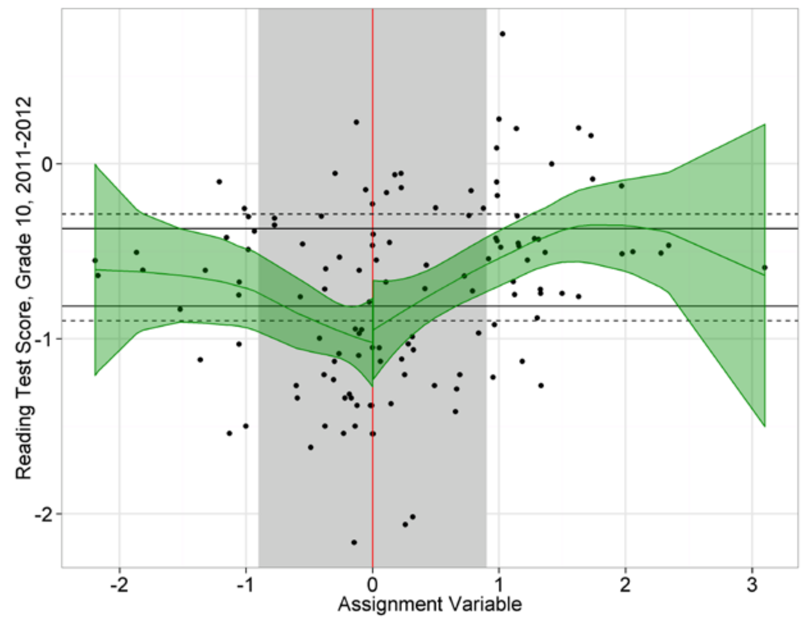


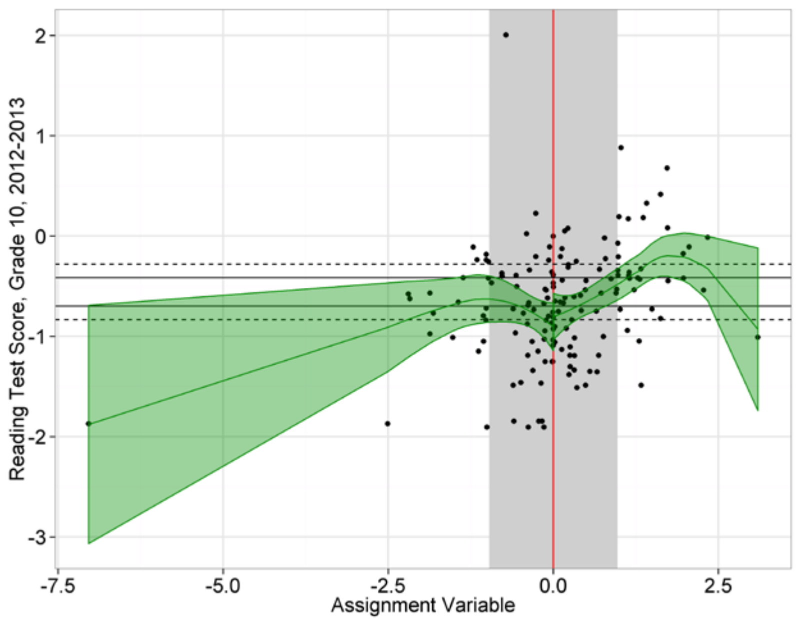
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.87. Reading test score in grade 10, accounting for student mobility



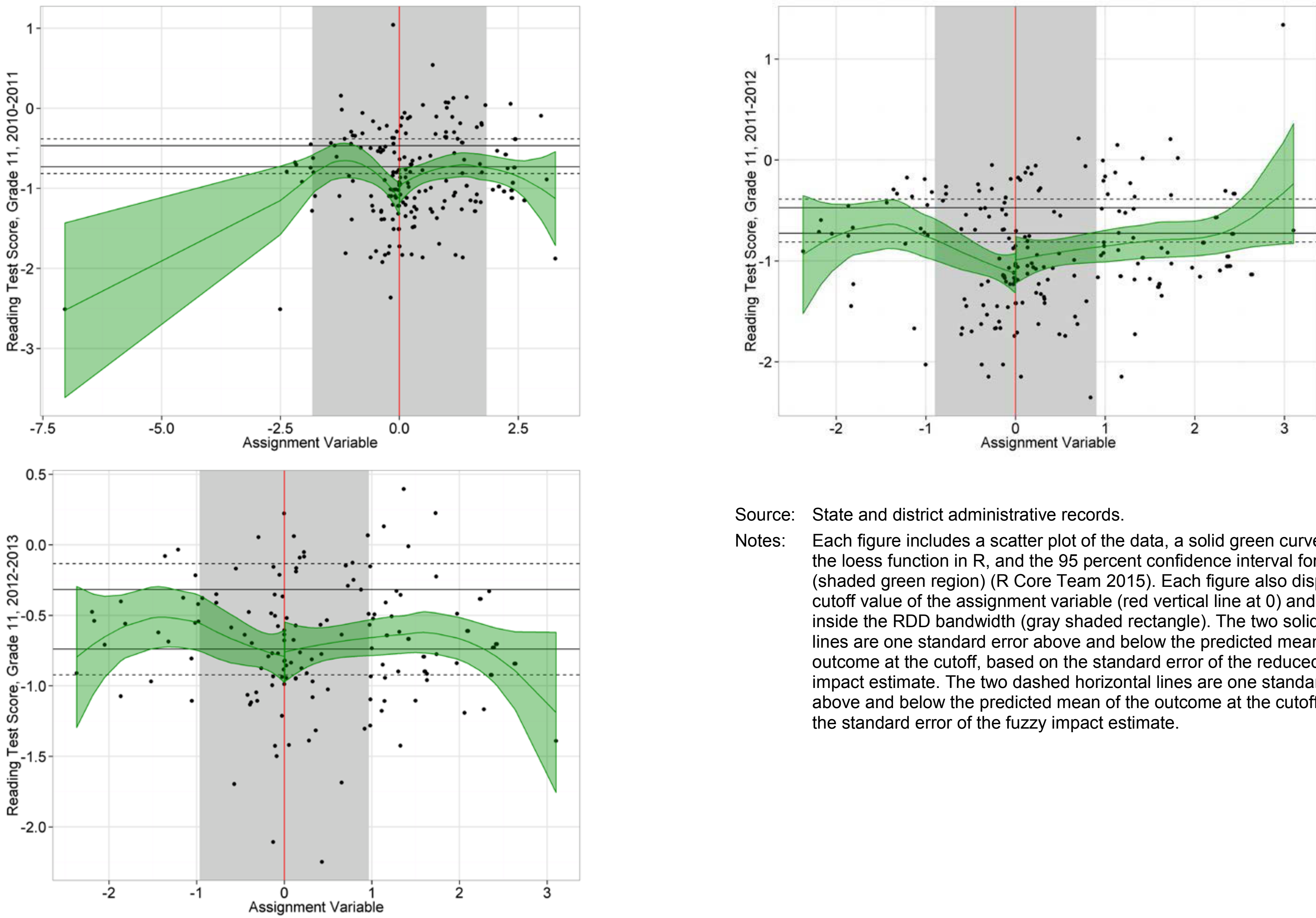




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

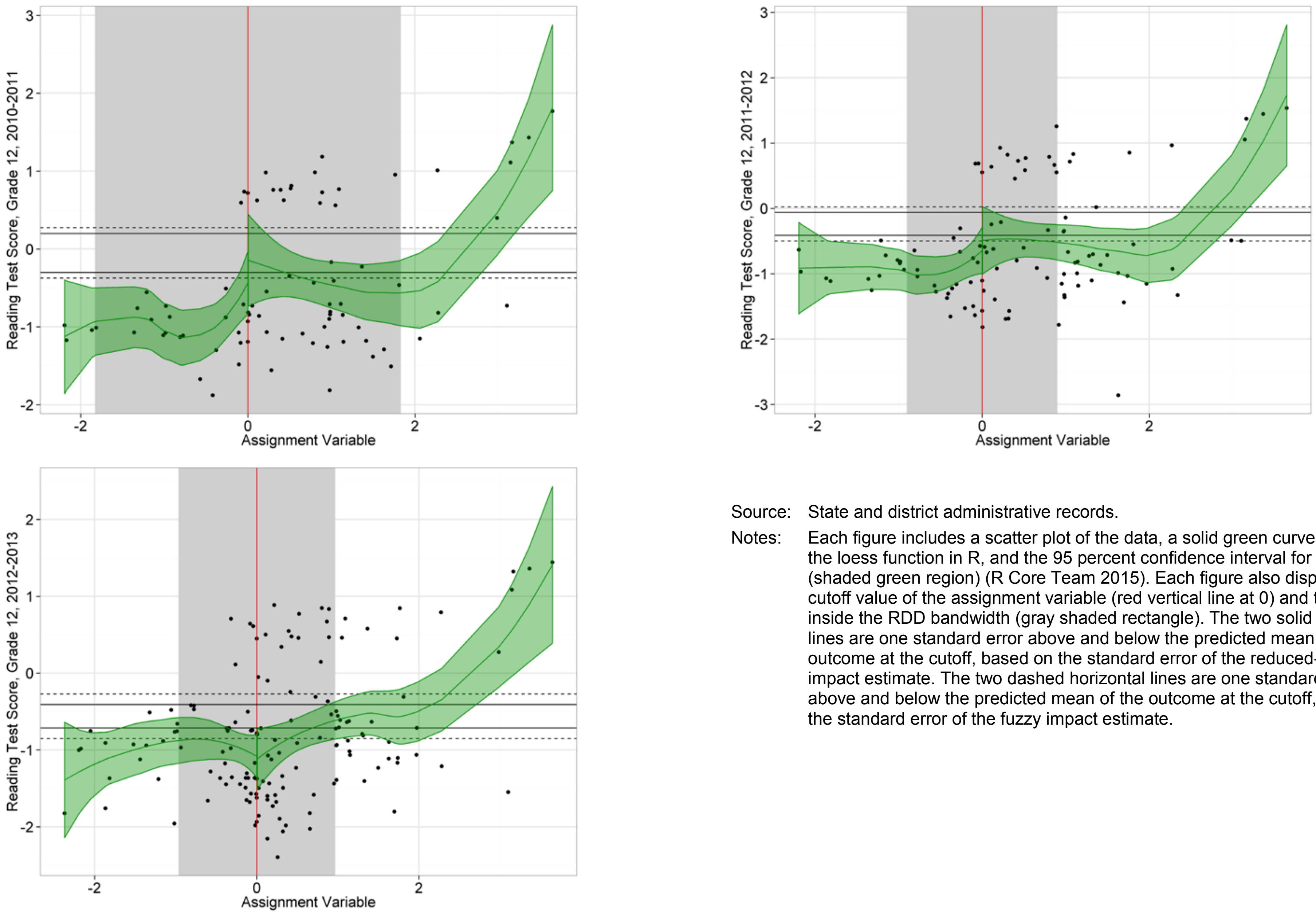
Figure A.88. Reading test score in grade 11, accounting for student mobility



Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

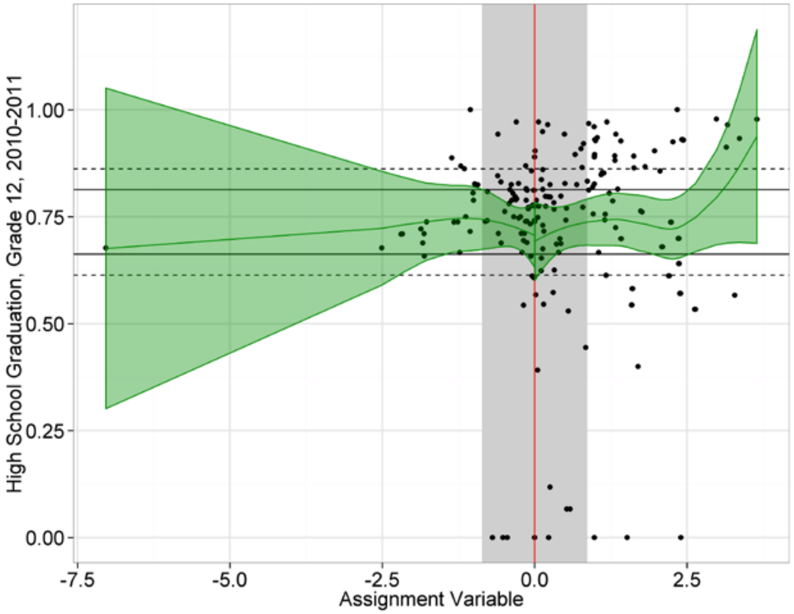
Figure A.89. Reading test score in grade 12, accounting for student mobility

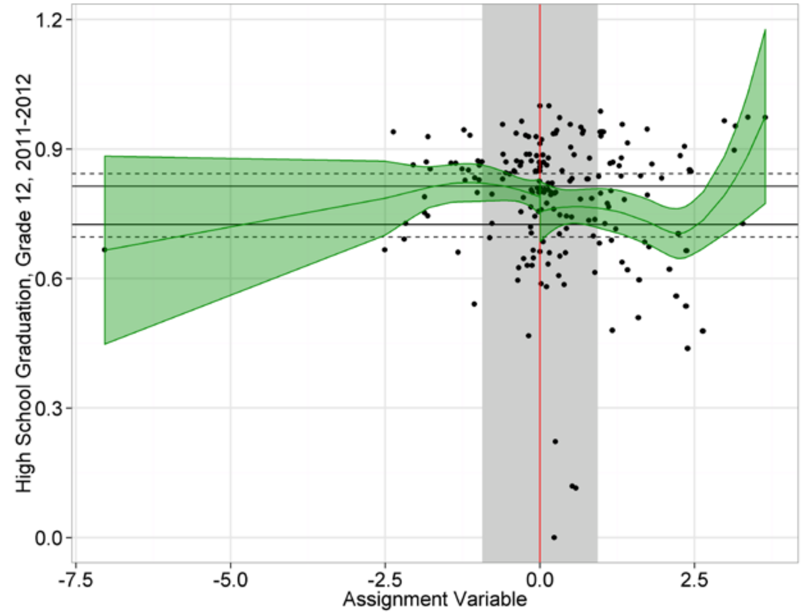


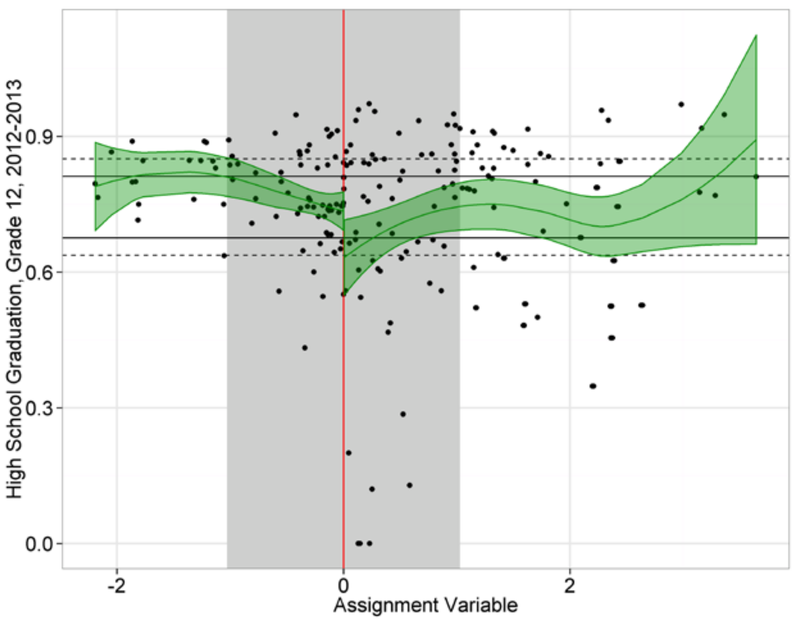
Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.90. High school graduation, accounting for student mobility



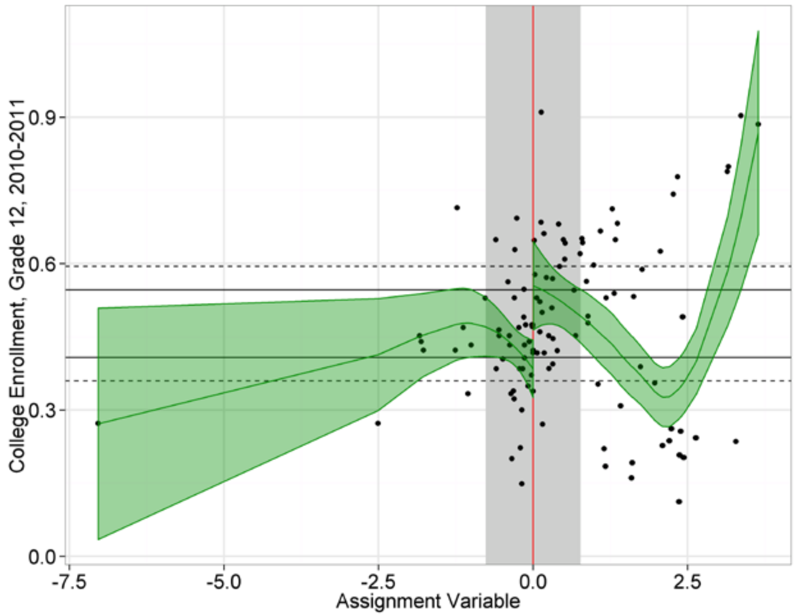


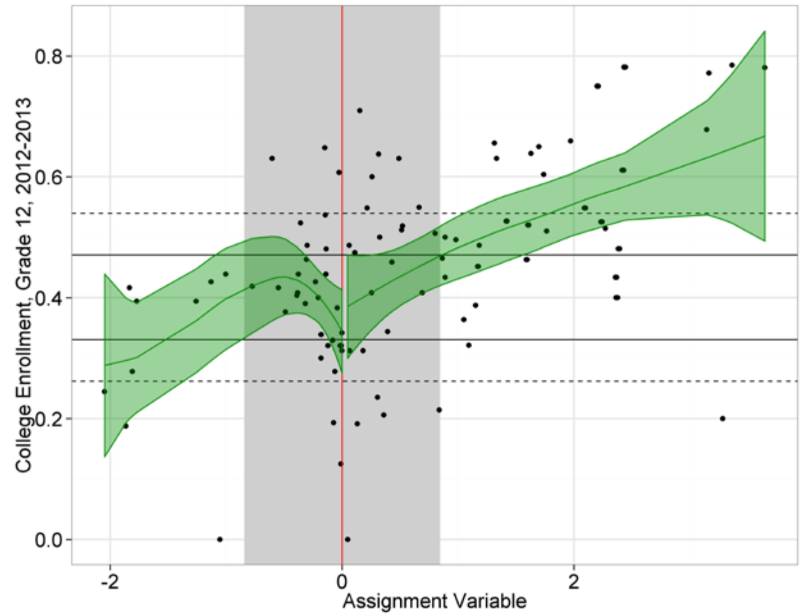


Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate.

Figure A.91. College enrollment, accounting for student mobility

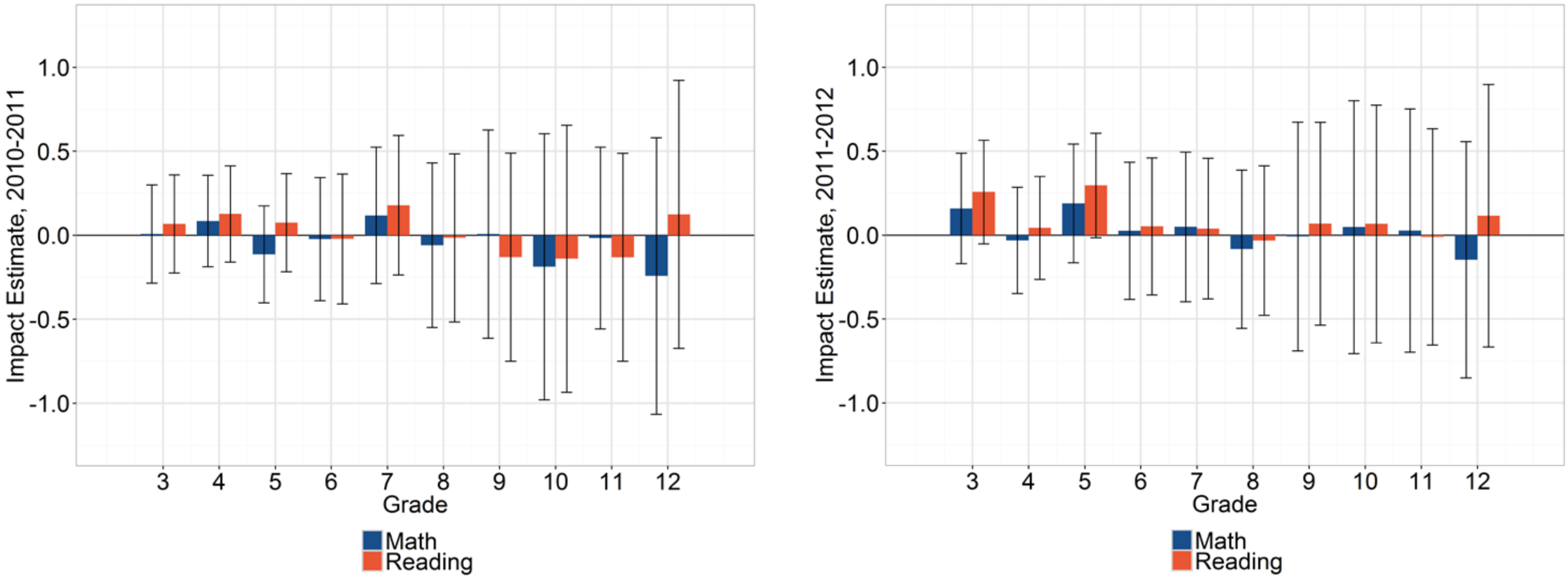




Source: State and district administrative records.

Notes: Each figure includes a scatter plot of the data, a solid green curve fit using the loess function in R, and the 95 percent confidence interval for the curve (shaded green region) (R Core Team 2015). Each figure also displays the cutoff value of the assignment variable (red vertical line at 0) and the region inside the RDD bandwidth (gray shaded rectangle). The two solid horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the reduced-form impact estimate. The two dashed horizontal lines are one standard error above and below the predicted mean of the outcome at the cutoff, based on the standard error of the fuzzy impact estimate. We were not able to estimate an impact for 2011–2012 due to insufficient sample sizes.

Figure A.92. Impacts of SIG-funded models on student test scores, by grade



Source: State and district administrative records.

Notes: We standardized test score outcomes to have a standard deviation of 1, so we report test score impact estimates in effect size units. Black bars show 95 percent confidence intervals. We calculated the results shown in this figure using the regression discontinuity design (RDD) methods described in this appendix.

APPENDIX B

CORRELATIONAL ANALYSIS

We used a correlational analysis to examine the relationship between the SIG intervention model a school implemented and changes in student outcomes. We analyzed changes in math and reading test scores between 2009–2010 (the year prior to SIG) and 2012–2013 (the latest year for which we had data) for schools implementing different models, separately for elementary grades (2 through 5) and higher grades (6 through 12). To investigate whether these changes were due to factors other than the type of model implemented, we analyzed (1) the amount of student mobility experienced by schools implementing different models, (2) changes in the composition of students attending schools implementing different models, and (3) baseline characteristics of schools implementing different models.

In Section A, we describe how we conducted the analysis of changes in math and reading test scores, as well as present results for all three outcome years (2010–2011, 2011–2012, and 2012–2013). We also provide findings related to the amount of student mobility experienced by schools implementing different models (Section B), changes in the composition of students attending schools implementing different models (Section C), sensitivity analyses (Section D), and baseline characteristics of schools implementing different models (Section E).

A. Analysis of changes in math and reading test scores

We faced two challenges when examining the relationship between the type of SIG model implemented and changes in student outcomes: (1) different states used different achievement measures; and (2) states may have changed their achievement measures during the years following SIG implementation (as part of Race to the Top reforms, for example). Our analysis approach addressed these challenges by (1) basing our estimates only on within-state variation (by including state indicator variables in our regressions), not between-state variation, to ensure that differences in state assessments were not confounded with SIG models, and (2) focusing on *differences* in student achievement gains between SIG models rather than on the achievement gains associated with individual models. If we had focused on student achievement gains for individual models, the changes in outcomes could be affected by a change in the state's assessments. For example, an increase in scores that appeared to be associated with the turnaround model might have been due to a change in the state assessment. Instead of looking at the effect of the turnaround model in isolation, we focused on the effect of the turnaround model *relative to* the transformation model, so that any changes in state assessments were differenced out (because those changes in the assessment also affected schools that implemented the transformation model).

To place test scores from different grades and states on a common and interpretable scale, we converted all test score variables to Normal Curve Equivalent (NCE) scores. If test scores are normally distributed, then about 98 percent of NCE scores will fall between the values of 0 and 100. This involved two steps. First we converted test scores to z-scores by subtracting statewide means and dividing by statewide standard deviations. We did this separately for each grade. Second, we converted z-scores to NCE scores using the formula: 50 + 21.06 * z-score (Mertler 2002).

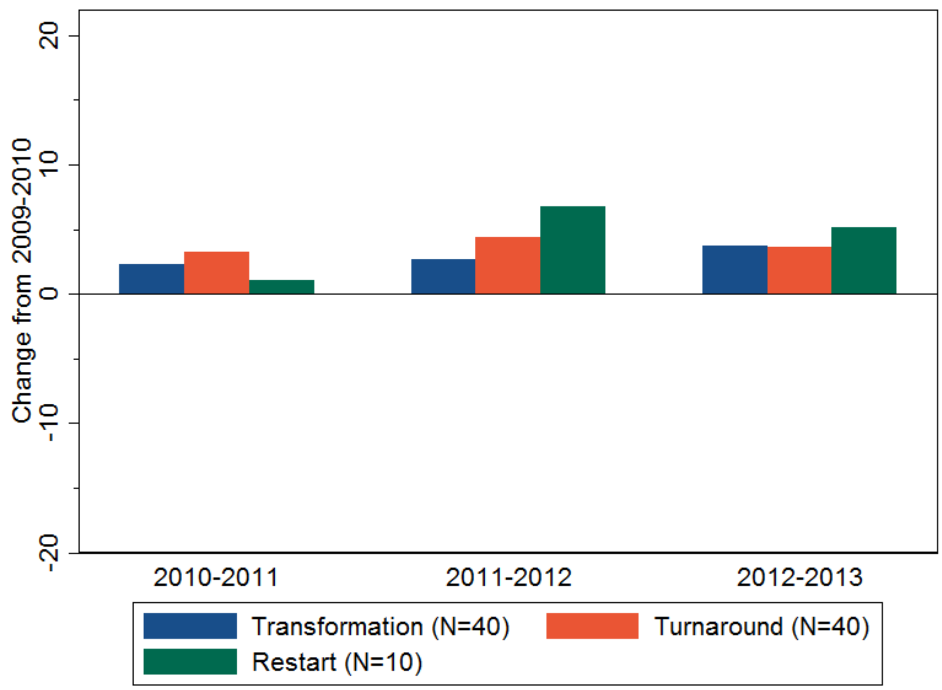
To measure the relationship between the SIG model implemented and the change in math or reading test scores, we used a weighted regression. The dependent variable was the change in math or reading scores. To create the change in math and reading test score variables, we subtracted the average NCE for each grade within each school in the pre-SIG year from the average NCE for the same grade within the same school in each of the post-SIG years. The independent variables were indicators for school intervention model, and indicators for grade and state. We controlled for grade and state so that our estimates were influenced only by within-grade and within-state variation. In addition, we accounted for the fact that each observation was a grade level within a school (thus we included multiple observations from the same school) by clustering the standard errors of the regression at the school level.

This analysis did not follow the same students over time; rather, we calculated gains by subtracting the average score for a grade in 2009–2010 from the average score for that same grade (containing a different set of students) in each outcome year. Because different sets of students were included in the pre-intervention and each post-intervention year, student compositional changes could affect test score gains, apart from effects of school intervention models (as we describe in more detail in section C).

To create the weights, we used the harmonic mean of the number of students in the relevant school and grade in the pre-SIG year and in each post-SIG year. For two numbers x_1 and x_2, the harmonic mean is $2x_1x_2/(x_1+x_2)$. We used the harmonic mean rather than the arithmetic mean because the harmonic mean results in lower weights on observations in which the numbers of students in the pre- and post-SIG years are imbalanced. For example, if the number of students in a school equaled 200 in both the pre-SIG and post-SIG years, the arithmetic mean would equal the harmonic mean (200). If, however, the number of students in the pre-SIG year was 300 and the number in the post-SIG year was 100, the arithmetic mean would equal 200 but the harmonic mean would equal 150. This feature of the harmonic mean is desirable because a large imbalance between the number of students in a grade in the pre- and post-SIG years may indicate a data problem such as a grade being phased out of a school (which would mean the number of students in that grade was large in the pre-SIG year but small in the post-SIG year). In that case, we would want to give this school-grade observation a lower weight than a school-grade observation of a similar size in which the grade was not being phased out.

For elementary grades, between 2009–2010 and each outcome year, there were no significant differences in math or reading gains between schools implementing different models (Figures B.1 and B.2). For higher grades, turnaround schools experienced larger gains in math than transformation schools between 2009–2010 and 2011–2012 and between 2009–2010 and 2012–2013, and restart schools experienced larger gains in math and reading than transformation schools between 2009–2010 and 2012–2013 (Figures B.3 and B.4).

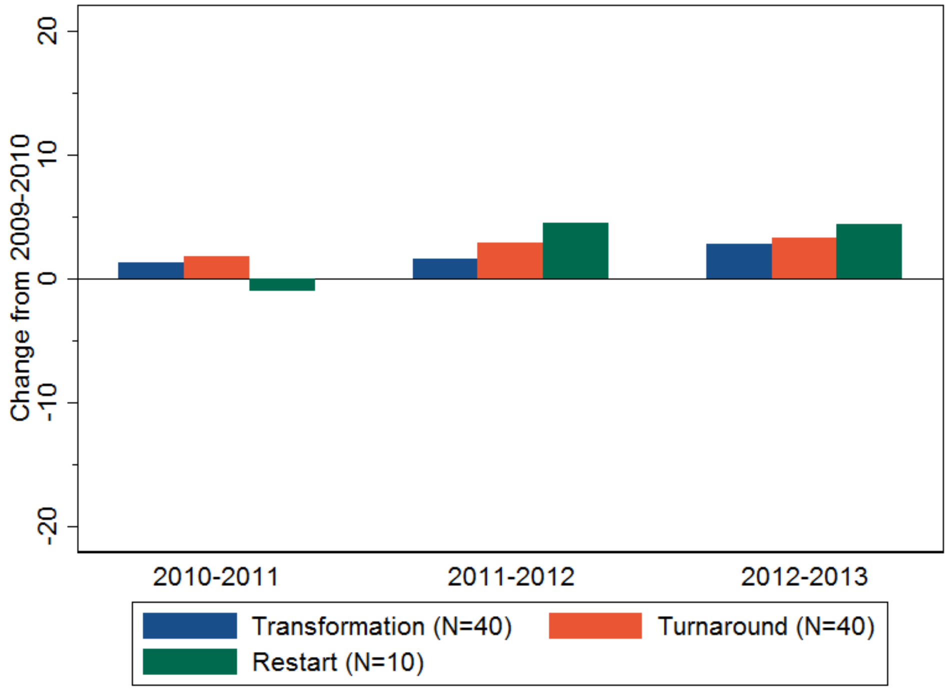
Figure B.1. Changes in math test scores in elementary grades, by model



Source: State administrative data.

Notes: This figure shows regression-adjusted changes in math test scores between the baseline year (2009–2010) and each post-SIG year (2010–2011, 2011–2012, and 2012–2013) in grades 2 through 5. Changes in math test scores were regression-adjusted for state and grade using a linear model. Units are normal curve equivalent (NCE) scores. There were no statistically significant differences between schools implementing different models.

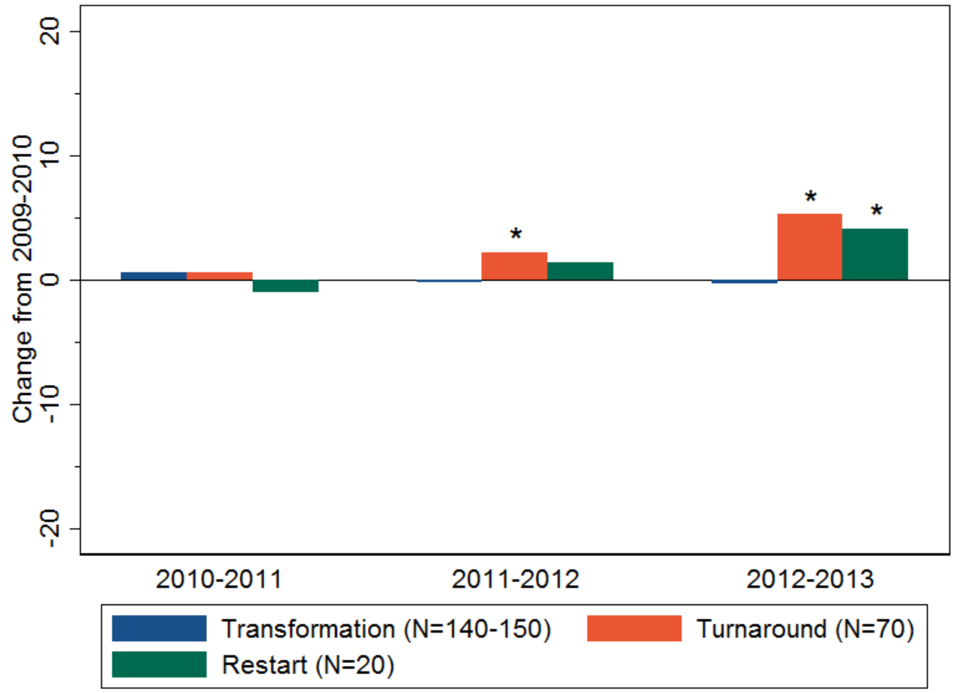
Figure B.2. Changes in reading test scores in elementary grades, by model



Source: State administrative data.

Notes: This figure shows regression-adjusted changes in reading test scores between the baseline year (2009–2010) and each post-SIG year (2010–2011, 2011–2012, and 2012–2013) in grades 2 through 5. Changes in reading test scores were regression-adjusted for state and grade using a linear model. Units are normal curve equivalent (NCE) scores. There were no statistically significant differences between schools implementing different models.

Figure B.3. Changes in math test scores in higher grades, by model

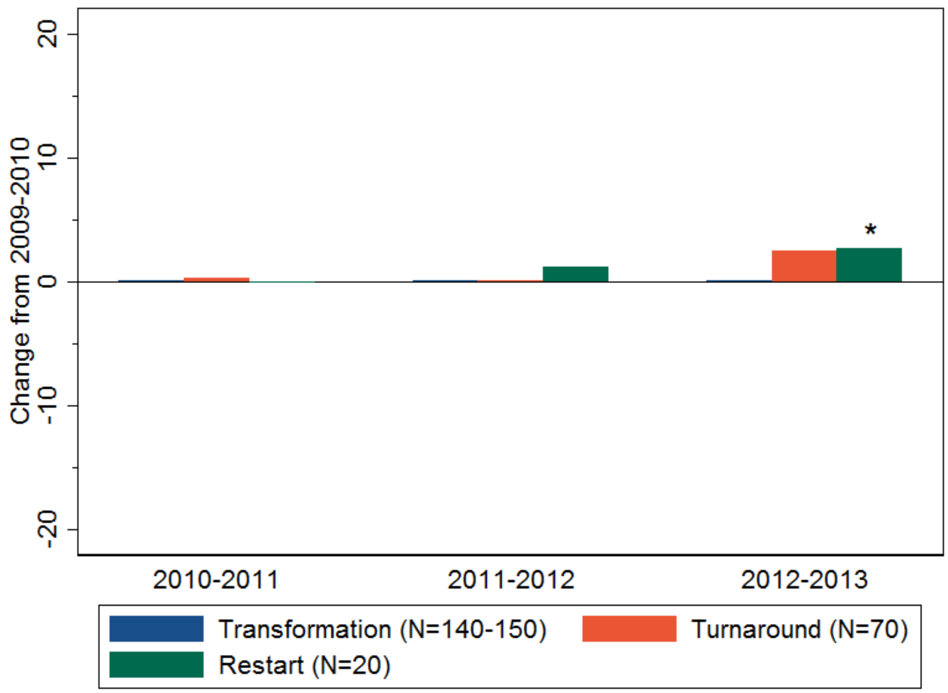


Source: State administrative data.

Notes: This figure shows regression-adjusted changes in math test scores between the baseline year (2009–2010) and each post-SIG year (2010–2011, 2011–2012, and 2012–2013) in grades 6 through 12. Changes in math test scores were regression-adjusted for state and grade using a linear model. Units are normal curve equivalent (NCE) scores.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

Figure B.4. Changes in reading test scores in higher grades, by model



Source: State administrative data.

Notes: This figure shows regression-adjusted changes in reading test scores between the baseline year (2009–2010) and each post-SIG year (2010–2011, 2011–2012, and 2012–2013) in grades 6 through 12. Changes in reading test scores were regression-adjusted for state and grade using a linear model. Units are normal curve equivalent (NCE) scores.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

B. Amount of student mobility for schools implementing different models

Because school intervention models might influence which schools students attend, we examined the extent of student mobility for schools implementing different models (described in this section), as well as changes in the composition of students attending schools implementing different models (described in the next section). Because our main analysis used only the students who were present in the relevant grade within the school in the pre- and post-SIG years, it comingled effects of intervention models on the academic outcomes of individual students with effects on the composition of the students attending these schools. For example, suppose the restart model had no effect on the outcomes of any individual student between the baseline year and 2012–2013, but more disadvantaged students left restart schools in greater proportions than schools implementing the transformation model. The finding that restart schools showed significantly higher gains than transformation schools could result from this shift in the composition of the student body and not be an effect of the models.

If we found a positive relationship between implementation of a particular intervention model and improved student achievement, it could be for either of two reasons. First, implementation of that intervention model might improve student achievement because schools implementing that model are more effective at educating children. Second, implementation of that model might change which students attend the schools implementing that model because, for example, more motivated parents might send their children to these schools even if the schools are not any better at educating children.

To investigate whether the second reason might help explain any observed changes in test scores in the main analysis, we analyzed student mobility patterns. We examined the proportions of students who were (1) new to the school in the years after SIG awards were made (inward mobility), or (2) no longer in the school in the years after SIG awards were made (outward mobility). Mobility for each post-intervention year was calculated relative to the prior school year. We calculated mobility based on which schools students were slated to attend. We identified the school a student was slated to attend each year based on the school they attended in the prior year and typical school feeder patterns in the district. We then ran a regression of the proportions of students who were new to the school or no longer in the school on indicators for school intervention model, state, and the amount of messiness observed in each district's school feeder patterns using a generalized linear model with a logit link function. Schools were weighted by their number of students in 2009–2010.

We created a four-category district-level variable to indicate the level of messiness in school feeder patterns and included this variable in the regressions for three reasons. First, we included them to account for the uncertainty about which school a student was slated to attend. In some districts, the school feeder patterns between elementary and middle school and between middle and high school were straightforward. For example, students from elementary school A nearly always attended middle school B. In other districts, the feeder patterns were messier. For example, students from elementary school A could attend middle school B, C, or D.

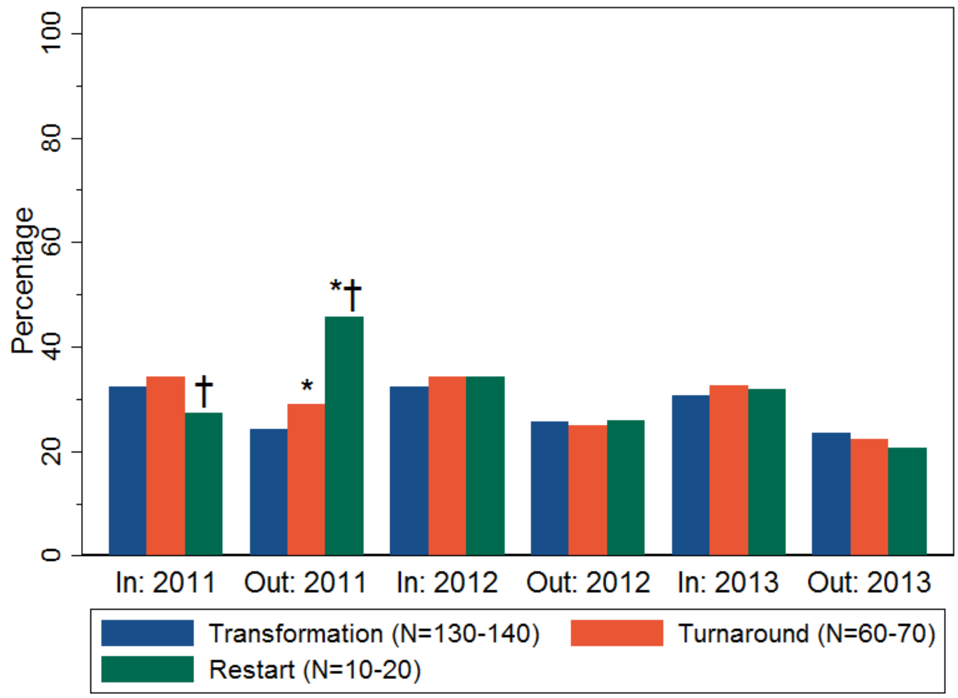
Second, we included the messiness indicators in the regressions to help ensure that our estimates of the relationship between school intervention models and the amount of student mobility were not biased by a possible correlation between model selection and messiness in feeder patterns. This correlation is a potential concern because both messiness and model selection could be the result of district policies. For example, feeder pattern messiness could be caused by a district policy that allows for school choice. Districts that have such a policy might also make different model selections than other districts.

Third, we included these indicators because they might explain variation in the student mobility outcome variables (the proportion of students new to the school or no longer in the school). By explaining this variation, the model can more precisely determine the relationship between school intervention model and student mobility. If we did not include the messiness indicators, the true relationship between school intervention model and student mobility could be swamped by variation caused by the different levels of messiness.

For schools serving students in grades 6 through 12, schools implementing different models experienced different amounts of student mobility (Figure B.5). Turnaround schools experienced more outward mobility than transformation schools between 2009–2010 and 2010–2011. Restart schools experienced less inward mobility than turnaround schools between 2009–2010 and 2010–2011 and more outward mobility than transformation and turnaround schools. There were no significant differences between models in the amount of mobility in later years (2011–2012 or

2012–2013). We did not present results for elementary grades because there were no significant differences in achievement gains between schools implementing different models for elementary grades.

Figure B.5. Inward and outward mobility in schools serving students in higher grades, by model



Source: State administrative data.

Notes: This figure shows regression-adjusted percentages of students new to the school in the post-SIG year (inward mobility) or no longer in the school in the post-SIG year (outward mobility) in schools serving students in grades 6 through 12. Mobility for 2010–2011 is relative to 2009–2010, mobility for 2011–2012 is relative to 2010–2011, and mobility for 2012–2013 is relative to 2011–2012. Percentages were regression-adjusted for state and messiness in district feeder patterns using a generalized linear model with a logit link function and robust standard errors.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

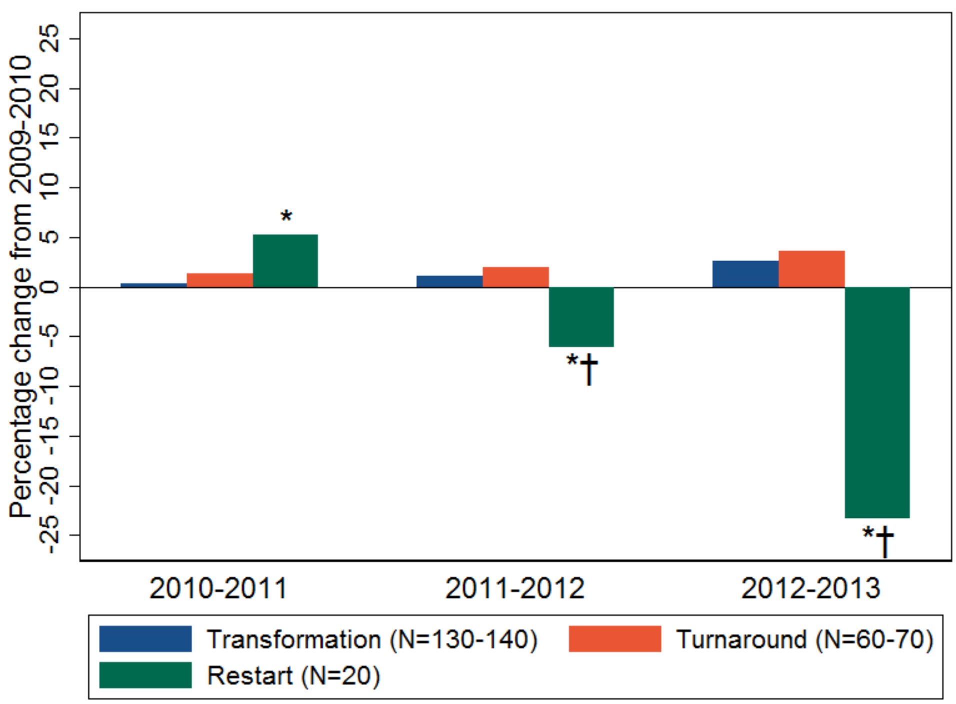
†Significantly different from turnaround model at the 0.05 level, two-tailed test.

C. Changes in composition of students attending schools implementing different models

To examine whether SIG-funded models changed which students attended the schools implementing those models, we analyzed changes over time in the student body composition of schools. Specifically, we examined changes over time with respect to several variables, including the percentage of students eligible to receive free or reduced-price lunch and the average math test scores from 2009–2010 (the year before SIG funding was received). For example, for schools implementing the transformation model, we calculated the school-level average baseline test score for all students in a school in 2009–2010, and subtracted this average from the school-level average baseline test score for all students in the school in 2012–2013. We did the same for schools implementing the turnaround model. We then tested whether the change in average achievement levels between the two years was statistically significantly different for transformation and turnaround schools. We ran a regression of changes in achievement levels on indicators for state and the amount of messiness in district feeder patterns using a linear model in which observations were weighted using the harmonic mean number of students described previously.

In Chapter VI, we described some of the results from this analysis. In this appendix, we present Figures B.6 through B.13 showing results for all years for the following characteristics: the percentage of students who were eligible for free or reduced-price lunch, ELL, Hispanic, black, white, and other race, and baseline math and reading test scores.

Figure B.6. Percentage change in students eligible for free or reduced-price lunch in schools serving students in higher grades, by model



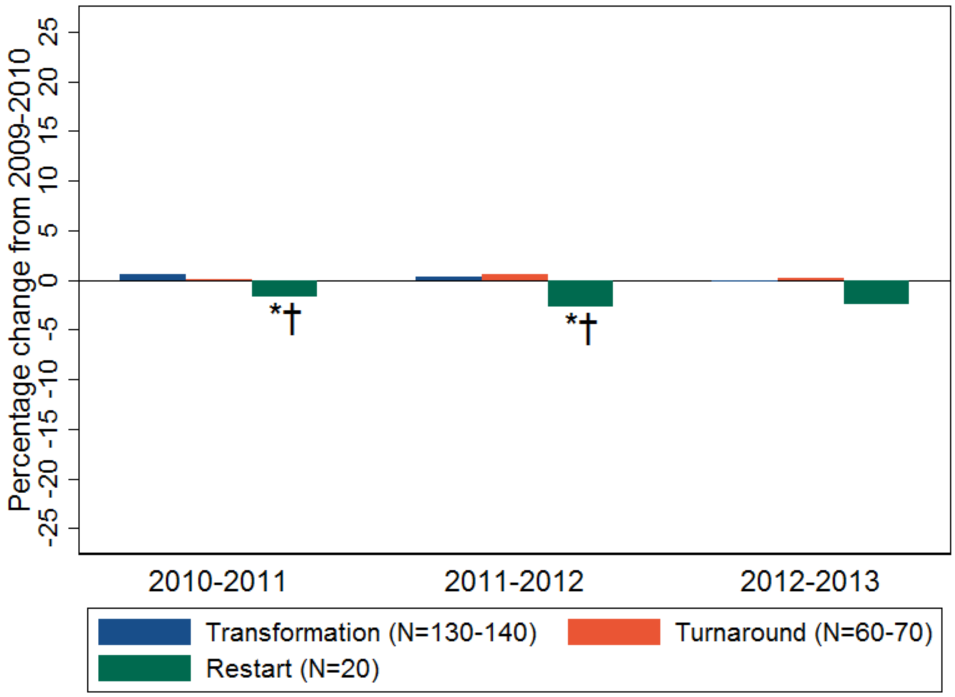
Source: State administrative data.

Notes: This figure shows regression-adjusted changes in the percentage of students eligible for free or reduced-price lunch relative to the baseline year (2009–2010) for schools serving students in grades 6 through 12. Percentage changes were regression-adjusted for state and messiness in district feeder patterns using a linear model.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

†Significantly different from turnaround model at the 0.05 level, two-tailed test.

Figure B.7. Percentage change in English language learner (ELL) students in schools serving students in higher grades, by model



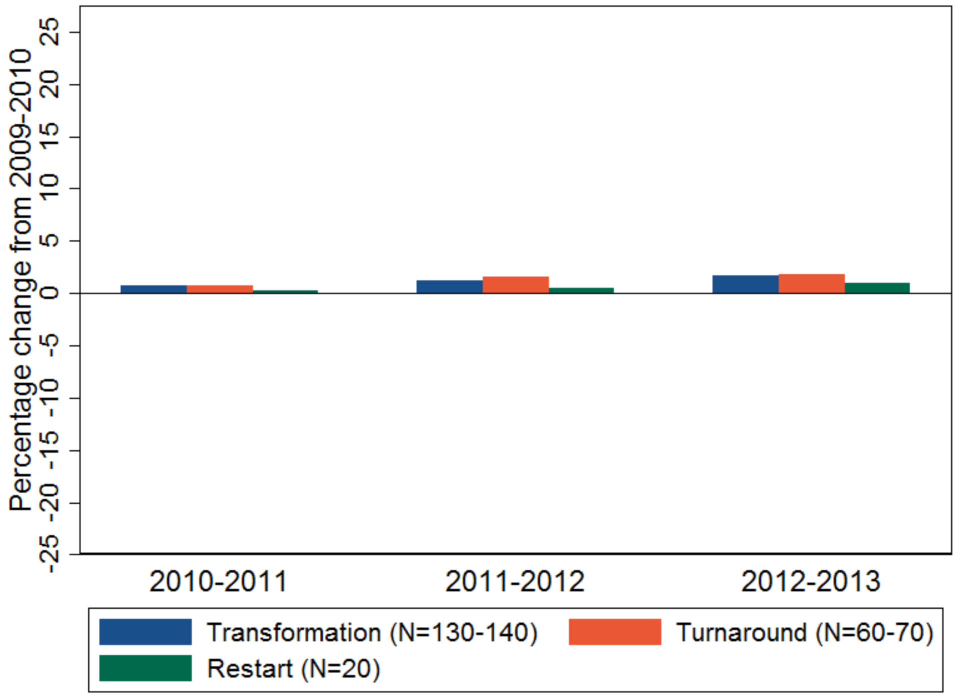
Source: State administrative data.

Notes: This figure shows regression-adjusted changes in the percentage of ELL students relative to the baseline year (2009–2010) for schools serving students in grades 6 through 12. Percentage changes were regression-adjusted for state and messiness in district feeder patterns using a linear model.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

†Significantly different from turnaround model at the 0.05 level, two-tailed test.

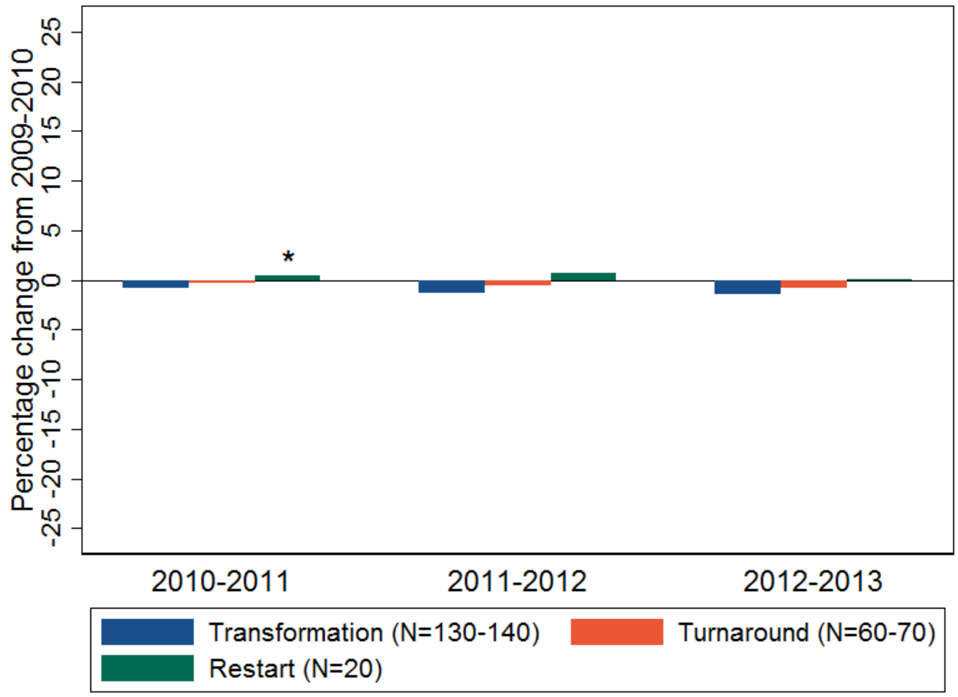
Figure B.8. Percentage change in Hispanic students in schools serving students in higher grades, by model



Source: State administrative data.

Notes: This figure shows regression-adjusted changes in the percentage of Hispanic students relative to the baseline year (2009–2010) for schools serving students in grades 6 through 12. Percentage changes were regression adjusted for state and messiness in district feeder patterns using a linear model. There were no statistically significant differences between schools implementing different models.

Figure B.9. Percentage change in Black students in schools serving students in higher grades, by model

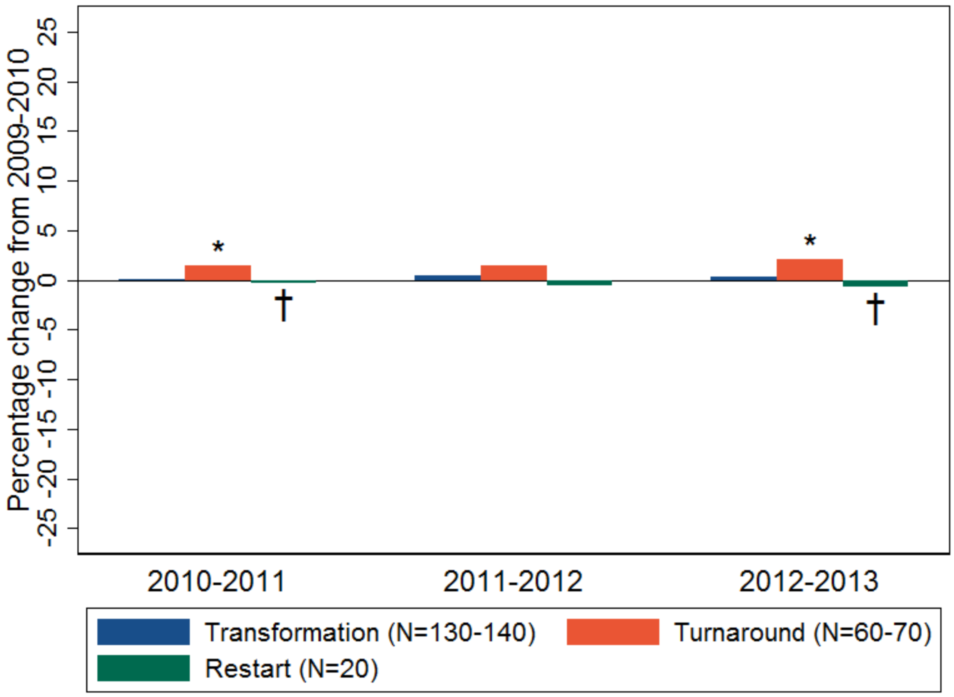


Source: State administrative data.

Notes: This figure shows regression-adjusted changes in the percentage of black students relative to the baseline year (2009–2010) for schools serving students in grades 6 through 12. Percentage changes were regression-adjusted for state and messiness in district feeder patterns using a linear model.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

Figure B.10. Percentage change in White students in schools serving students in higher grades, by model



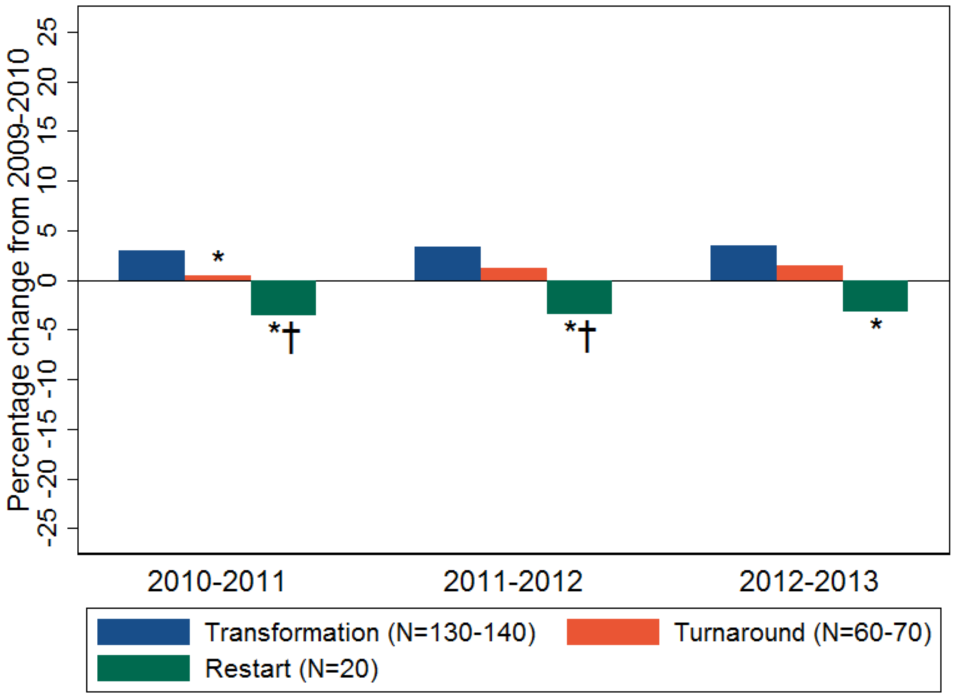
Source: State administrative data.

Notes: This figure shows regression-adjusted changes in the percentage of white students relative to the baseline year (2009–2010) for schools serving students in grades 6 through 12. Percentage changes were regression-adjusted for state and messiness in district feeder patterns using a linear model.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

†Significantly different from turnaround model at the 0.05 level, two-tailed test.

Figure B.11. Percentage change in students of other race in schools serving students in higher grades, by model



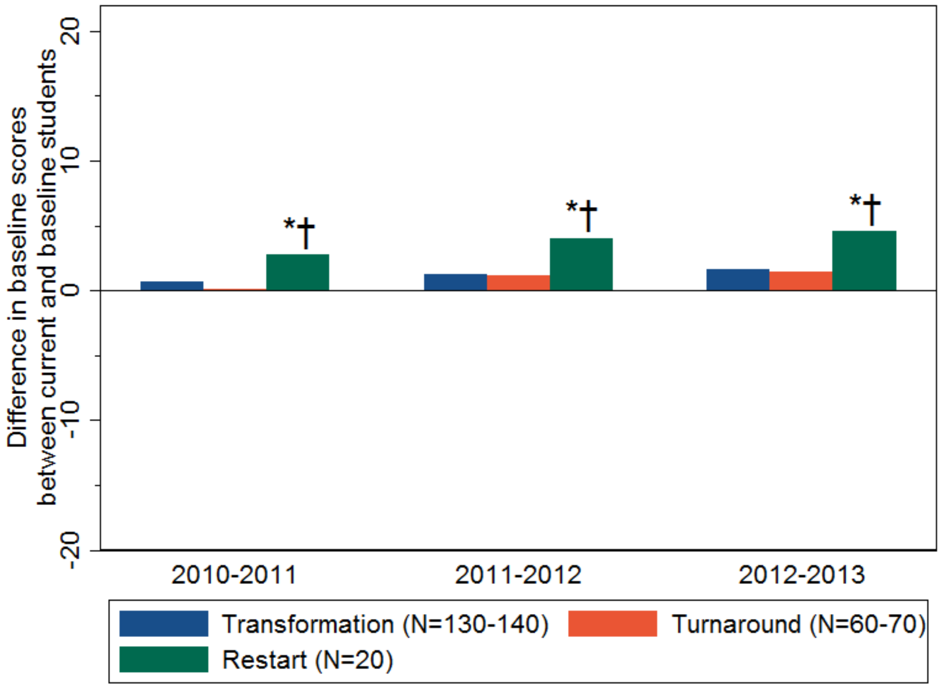
Source: State administrative data.

Notes: This figure shows regression-adjusted changes in the percentage of students of "other" race (that is, Asian, Pacific Islander, Native American, or more than one race) relative to the baseline year (2009–2010) for schools serving students in grades 6 through 12. Percentage changes were regression-adjusted for state and messiness in district feeder patterns using a linear model.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

†Significantly different from turnaround model at the 0.05 level, two-tailed test.

Figure B.12. Difference in baseline math scores between current and baseline students in schools serving students in higher grades, by model



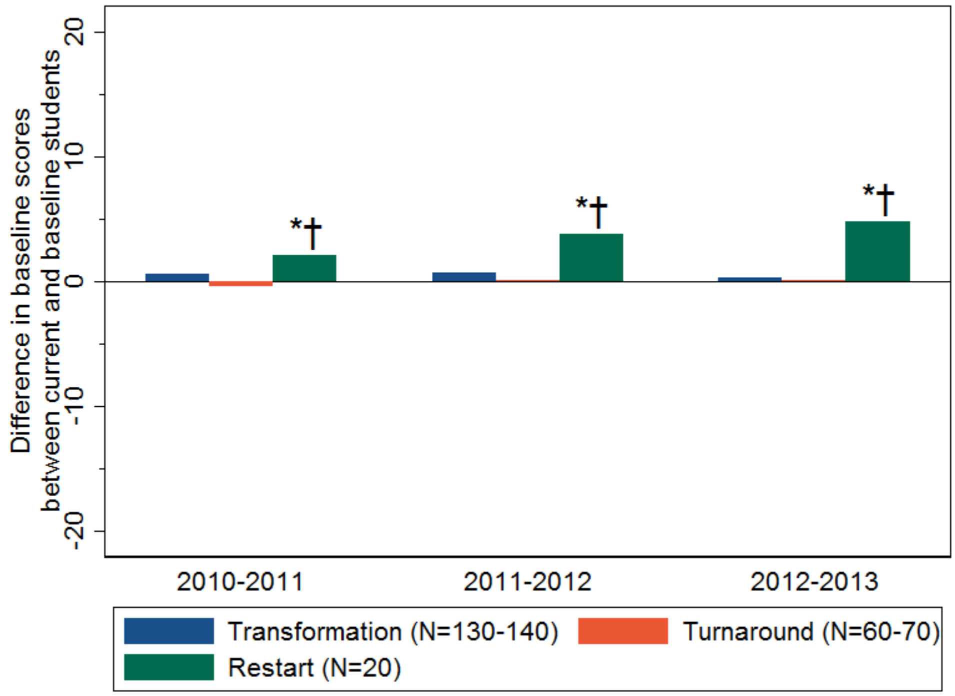
Source: State administrative data.

Notes: This figure shows regression-adjusted average differences in baseline (2009–2010) math scores between students attending schools serving students in grades 6 through 12 in each post-SIG year (2010–2011, 2011–2012, or 2012–2013) and students attending those same schools in the baseline year. Differences in baseline test scores were regression-adjusted for state and messiness in district feeder patterns using a linear model. Units are normal curve equivalent (NCE) scores.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

†Significantly different from turnaround model at the 0.05 level, two-tailed test.

Figure B.13. Difference in baseline reading scores between current and baseline students in schools serving students in higher grades, by model



Source: State administrative data.

Notes: This figure shows regression-adjusted average differences in baseline (2009–2010) reading scores between students attending schools serving students in grades 6 through 12 in each post-SIG year (2010–2011, 2011–2012, or 2012–2013) and students attending those same schools in the baseline year. Differences in baseline test scores were regression-adjusted for state and messiness in district feeder patterns using a linear model. Units are normal curve equivalent (NCE) scores.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

†Significantly different from turnaround model at the 0.05 level, two-tailed test.

D. Sensitivity analyses

1. Grade-specific analyses

When we estimated the relationship between SIG model and test score changes for each grade separately, we found that in higher grades, the turnaround model and, to a lesser extent, the restart model were associated with higher achievement gains than the transformation model (Table B.1). In lower grades, there were consistently no significant differences between the models.

Table B.1. Results from grade-specific analyses of changes in test scores

Grade	Regression-adjusted changes in test scores between 2009–2010 and 2012–2013		
	Transformation	Turnaround	Restart
Math			
3	5.3	3.4	6.5
4	3.2	3.9	4.3
5	4.9	3.7	6.4
6	2.6	4.2	4.9
7	2.7	4.8	5.7*
8	0.3	4.9**	10.1**,†
9	0.7	4.6	0.0
10	0.0	8.6**	9.0
Reading			
3	4.7	3.3	4.9
4	3.2	2.8	5.7
5	3.4	3.3	5.5
6	1.9	4.4*	4.8*
7	2.2	3.0	6.7**,†
8	2.4	3.0	5.5*
9	1.8	3.3	1.8
10	1.8	5.7*	9.9

Source: State administrative data.

Notes: This figure shows regression-adjusted changes in test scores between the baseline year (2009–2010) and 2012–2013 for grades 3–10. Units are normal curve equivalent (NCE) scores. Changes in test scores were regression-adjusted for state using a linear model.

*/**Significantly different from transformation at the 0.10 level/0.05 level, two-tailed test.

†Significantly different from turnaround model at the .05 level, two-tailed test.

2. Analysis that accounted for student mobility

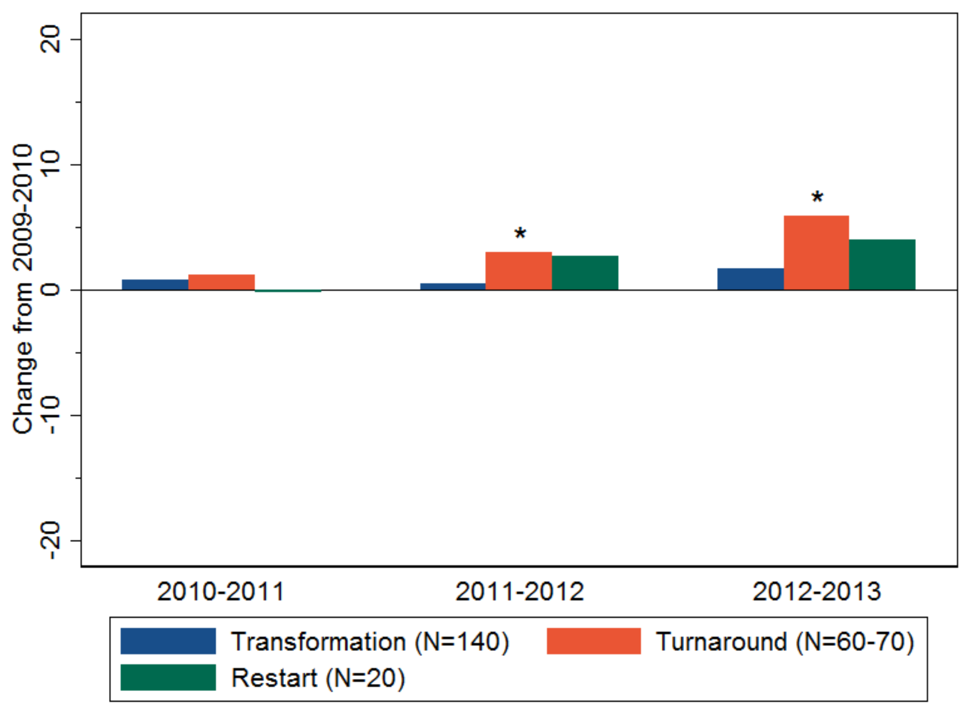
The analyses described in Section C showed that the student body composition of schools implementing different models changed over time. To account for this, we conducted a sensitivity analysis to determine whether and how much of the observed differences in outcome changes between models could be due to student mobility. The sensitivity analysis involved re-estimating our main model using outcome changes calculated in a way that accounted for student mobility. Specifically, we calculated outcome changes using test scores of students who were *slated to attend* a particular school, as opposed to students who *actually attended* the school (which was the approach used in the benchmark analysis). Unlike the benchmark analysis, this sensitivity test included schools that implemented the closure model, because we analyzed outcome changes using test scores of students who were slated to attend the closure schools had they not closed.

We identified the school a student was slated to attend based on (1) the school they attended in the baseline year, and (2) typical school feeder patterns in the district. We created the dependent variables, changes in math and reading test scores, in the same way as in the benchmark analysis except that, in each post-SIG year, students were associated with the school they were slated to attend rather than the school they actually attended. That is, we subtracted the average NCE for each grade within each school in the pre-SIG year from the average NCE for students slated to attend that school in that grade in each of the post-SIG years. These outcome changes were regressed on indicators for grade, state, and the amount of messiness in district feeder patterns.

The sensitivity analysis that accounted for student mobility changed the findings from the benchmark analysis for restart schools. Specifically, we found that after accounting for student mobility, there were no longer significant gains in math or reading test scores for restart schools relative to transformation schools, for any outcome year (Figures B.14 and B.15).

The sensitivity analysis that accounted for student mobility did not change the finding from the benchmark analysis of larger math gains in turnaround schools than in transformation schools. Specifically, we found that even after accounting for student mobility, turnaround schools experienced larger gains in math than transformation schools between 2009–2010 and 2011–2012 and between 2009–2010 and 2012–2013 (Figure B.14). As in the benchmark analysis, there were no significant differences between these two models with respect to gains in reading for any year (Figure B.15).

Figure B.14. Changes in math test scores in higher grades, accounting for student mobility, by model

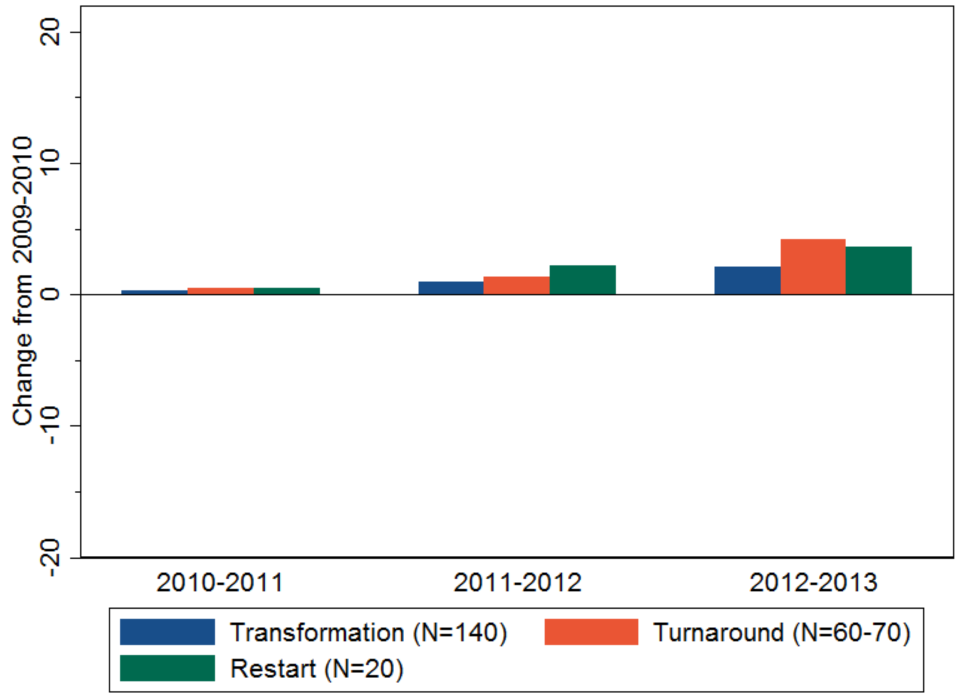


Source: State administrative data.

Notes: This figure shows regression-adjusted changes in math test scores between the baseline year (2009–2010) and each post-SIG year (2010–2011, 2011–2012, or 2012–2013) in grades 6 through 12, using changes calculated in a way that accounted for student mobility. Changes in math test scores were regression-adjusted for state, grade, and messiness in district feeder patterns using a linear model. Units are normal curve equivalent (NCE) scores.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

Figure B.15. Changes in reading test scores in higher grades, accounting for student mobility, by model



Source: State administrative data.

Notes: This figure shows regression-adjusted changes in reading test scores between the baseline year (2009–2010) and each post-SIG year (2010–2011, 2011–2012, or 2012–2013) in grades 6 through 12, using changes calculated in a way that accounted for student mobility. Changes in reading test scores were regression-adjusted for state, grade, and messiness in district feeder patterns using a linear model. Units are normal curve equivalent (NCE) scores. There were no statistically significant differences between schools implementing different models.

3. Other sensitivity analyses

We conducted several additional sensitivity analyses. First, we included additional school characteristics (measured in 2009–2010, prior to SIG) as control variables in our benchmark analysis regressions. These additional characteristics were math and reading test scores; the percentage of students who were male, white, Black, Hispanic, other race, eligible for free or reduced-price lunch, and ELLs; school level (that is, whether the school included grade 3, grade 7, and/or grade 10 to differentiate among elementary, middle, high, and K-8 schools); and whether the school was located in a large city, a small or mid-sized city, a suburb, or a rural area. Second, we ran our benchmark analysis regressions without weights (that is, we ran regressions in which each school received an equal weight). Third, we included additional school characteristics in our regressions that accounted for student mobility (or "mobility-robust" regressions, for short). We used the same characteristics listed above.

The findings for elementary grades were robust across all sensitivity analyses. Between 2009–2010 and each outcome year, there were no significant differences in math or reading gains between schools implementing different models (Figures B.16 and B.17).

Among higher grades, the findings were generally robust across sensitivity analyses, though there were a few exceptions. The finding that turnaround schools experienced larger math gains than transformation schools was robust across most sensitivity analyses (it remained significant in all analyses except the regression that used an equal weight for each school) (Figure B.18). The finding of no significant differences in reading gains between turnaround and transformation schools was robust across some, but not all, sensitivity analyses (a few analyses showed that turnaround schools experienced larger reading gains than transformation schools) (Figure B.19). The finding that restart schools experienced larger math and reading gains than transformation schools was not robust to accounting for student mobility and additional school characteristics (Figures B.18 and B.19).

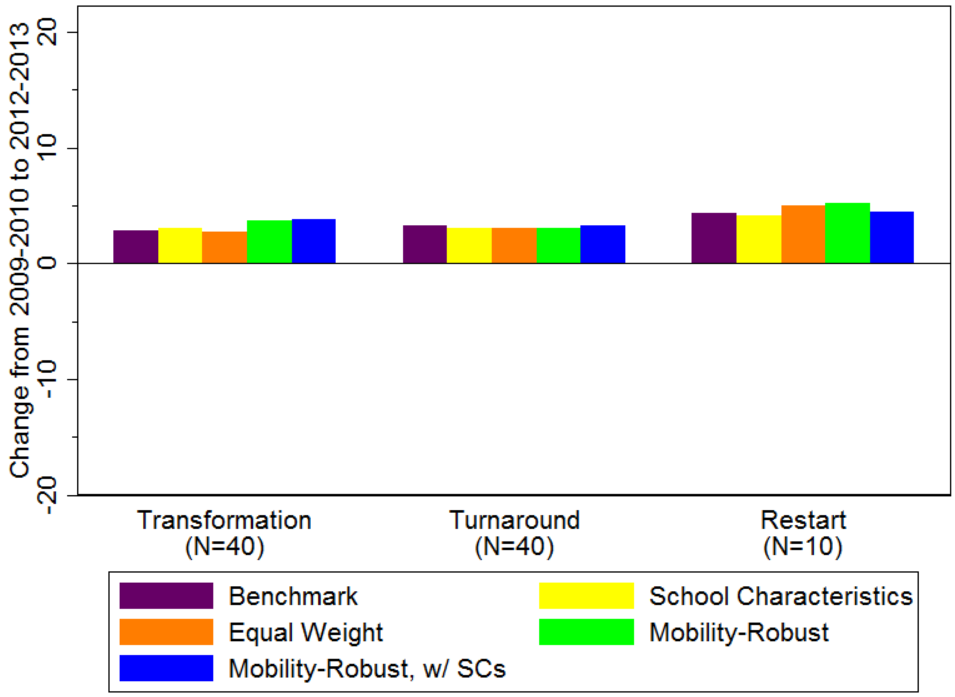
Figure B.16. Changes in math test scores in elementary grades, sensitivity analyses, by model

Source: State administrative data.

Notes: Units are normal curve equivalent (NCE) scores. There were no statistically significant differences between schools implementing different models.

SCs = school characteristics.

Figure B.17. Changes in reading test scores in elementary grades, sensitivity analyses, by model

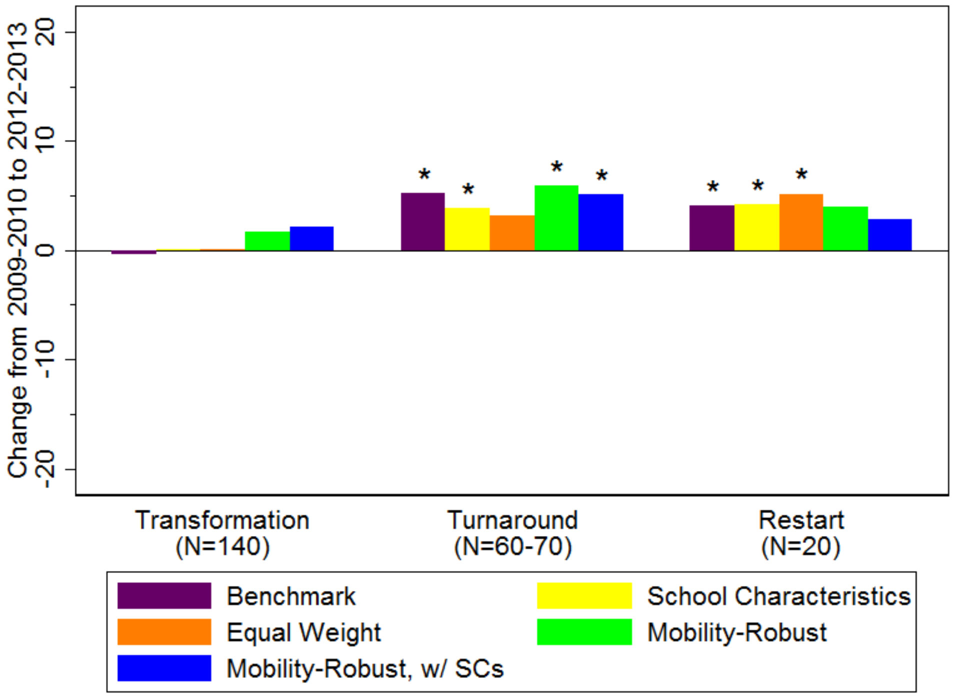


Source: State administrative data.

Notes: Units are normal curve equivalent (NCE) scores. There were no statistically significant differences between schools implementing different models.

SCs = school characteristics.

Figure B.18. Changes in math test scores in higher grades, sensitivity analyses, by model



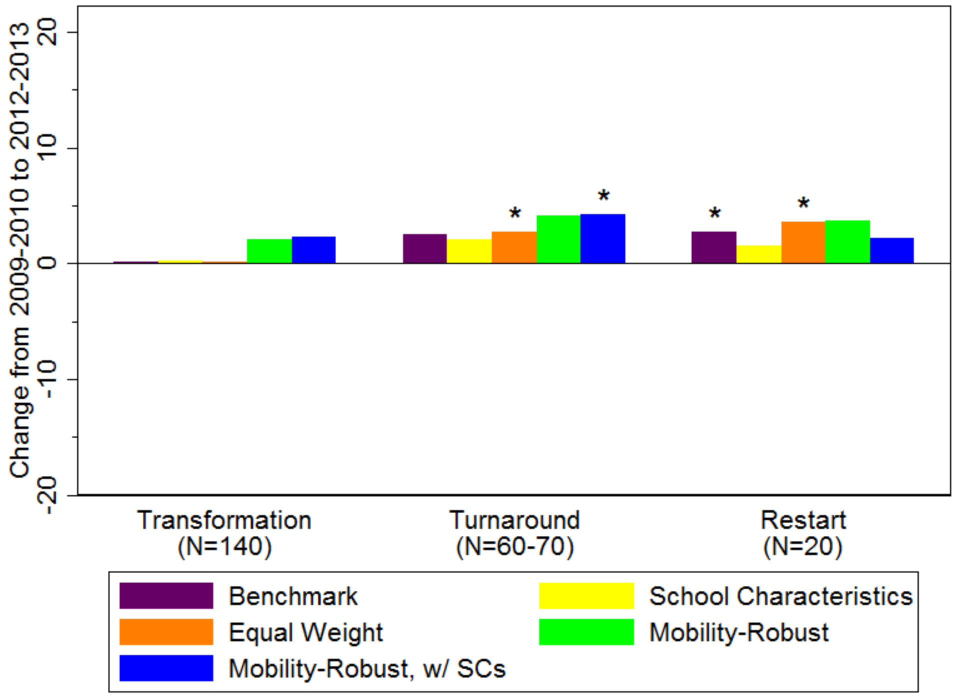
Source: State administrative data.

Notes: Units are normal curve equivalent (NCE) scores.

SCs = school characteristics.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

Figure B.19. Changes in reading test scores in higher grades, sensitivity analyses, by model



Source: State administrative data.

Notes: Units are normal curve equivalent (NCE) scores.

SCs = school characteristics.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

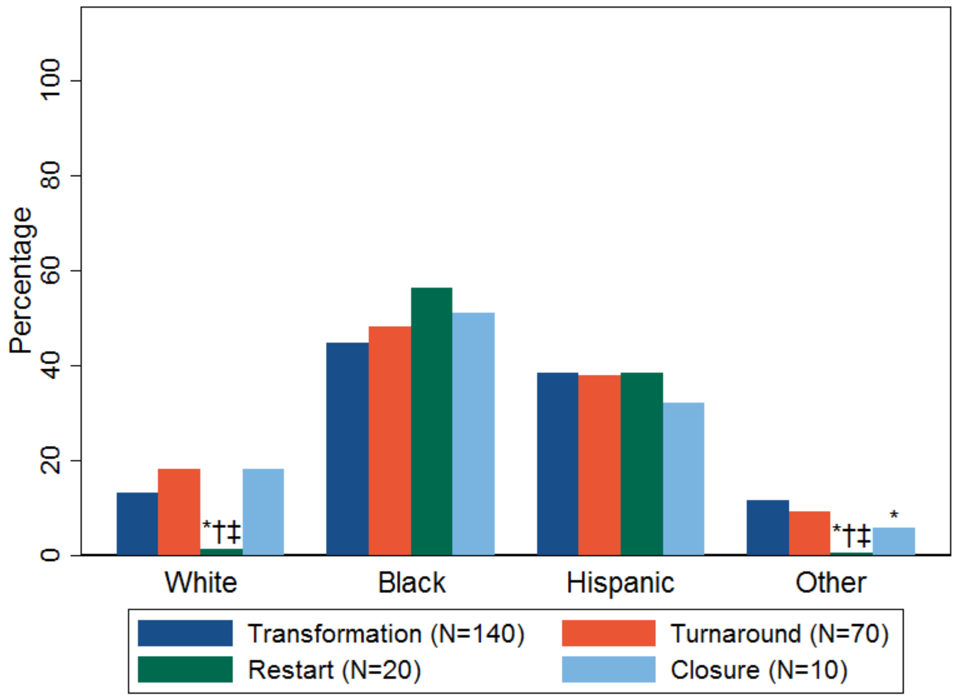
E. Baseline characteristics of schools implementing different models

In Chapter VI we reported differences between schools implementing different models with respect to two baseline characteristics—percentage of students eligible for free or reduced-price lunch and average baseline math scores. In this appendix, we report results for additional baseline characteristics: percentages of students who were ELLs, white, black, Hispanic, or other race; urbanicity; and reading test scores.

For characteristics measured as percentages—percentage of students who were ELLs, white, black, Hispanic, or other race—we used a generalized linear model with a logit link function, and we included state indicators in the model. For urbanicity categories, we used a logit model, reported results as percentages, and included state indicators in the model. For baseline test scores (in NCE units), we used a linear model and included state indicators in the model. For all characteristics, each school was weighted by its number of students at baseline.

For schools serving students in grades 6 through 12, we found several baseline differences between schools implementing different models. Restart schools had lower percentages of white students than other schools (Figure B.20). Transformation schools were more likely than other schools to be located in a mid-sized or small city, and, relative to turnaround schools, less likely to be located in a large city or suburban or rural area (Figure B.21). Turnaround schools served lower-achieving students than transformation schools, and restart schools served lower-achieving students than transformation and turnaround schools (Figure B.22). We found no differences between models with respect to the percentage of ELL students at baseline (Figure B.23). We did not present results for elementary grades because there were no significant differences in achievement gains between schools implementing different models for elementary grades.

Figure B.20. Baseline race/ethnicity percentages in schools serving students in higher grades, by model



Source: State administrative data.

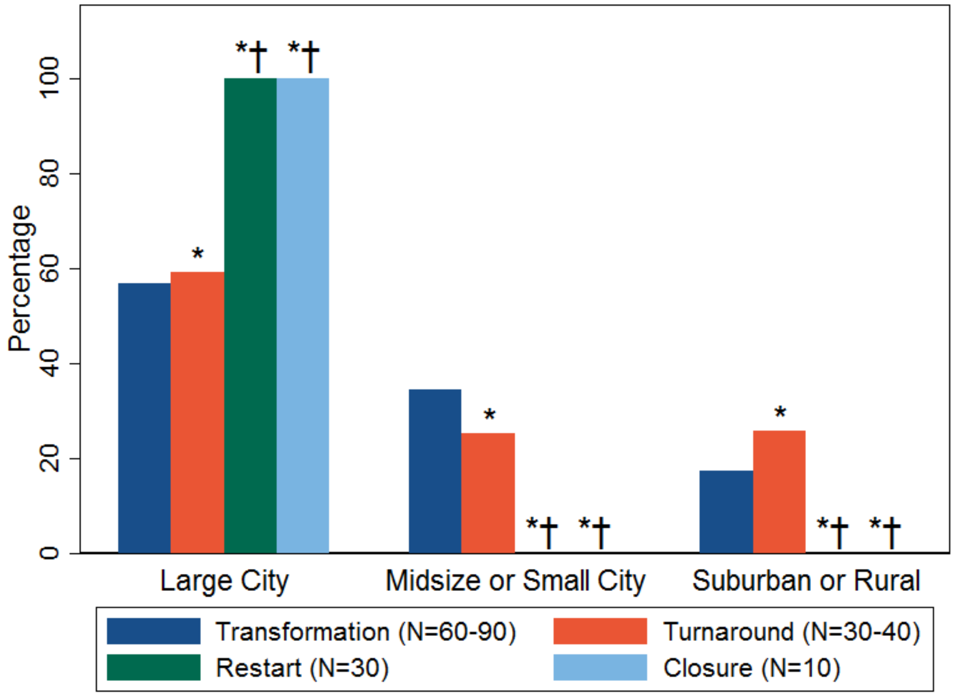
Notes: This figure shows regression-adjusted percentages of students in various race/ethnicity categories at baseline (the 2009–2010 school year) in schools serving students in grades 6 through 12. Percentages were regression-adjusted for state using a generalized linear model with a logit link function.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

†Significantly different from turnaround model at the 0.05 level, two-tailed test.

‡Significantly different from closure model at the 0.05 level, two-tailed test.

Figure B.21. Baseline urbanicity in schools serving students in higher grades, by model



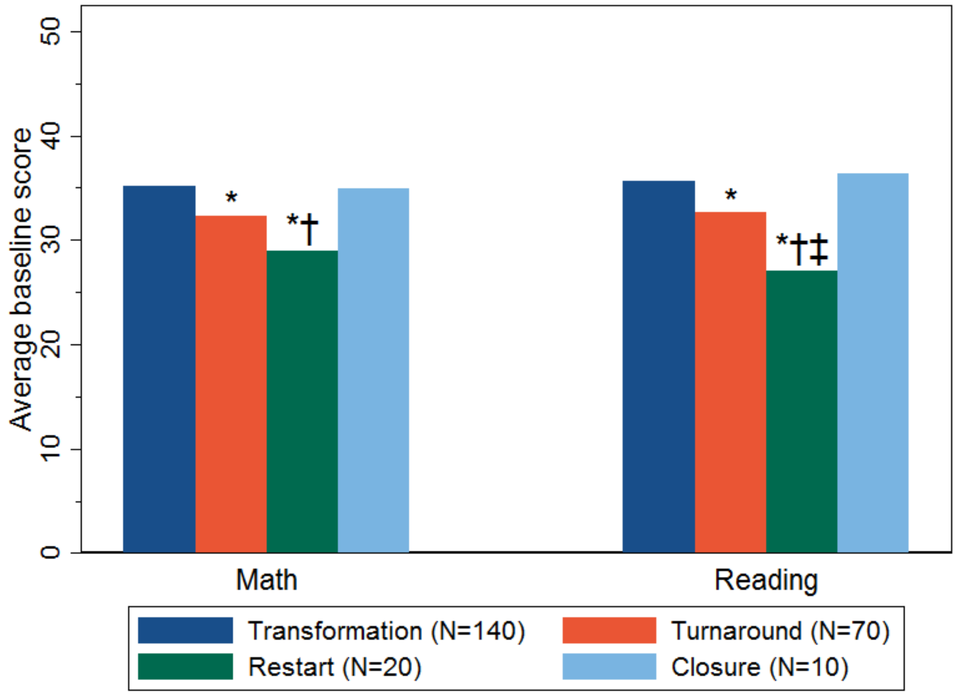
Source: State administrative data.

Notes: This figure shows regression-adjusted percentages of schools in different urbanicity categories at baseline (the 2009–2010 school year) in schools serving students in grades 6 through 12. Percentages were regression-adjusted for state using a logit model.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

†Significantly different from turnaround model at the 0.05 level, two-tailed test.

Figure B.22. Average baseline test scores in schools serving students in higher grades, by model



Source: State administrative data.

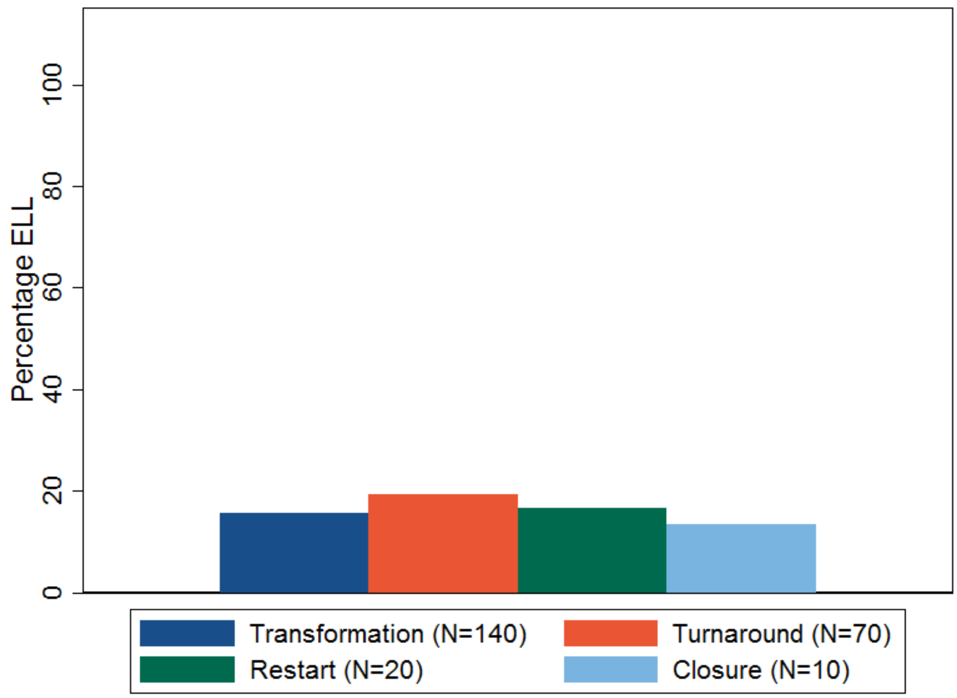
Notes: This figure shows regression-adjusted average math and reading test scores at baseline (the 2009-2010 school year) in schools serving students in grades 6 through 12. Scores were regression-adjusted for state using a linear model. Units are normal curve equivalent (NCE) scores.

*Significantly different from transformation model at the 0.05 level, two-tailed test.

†Significantly different from turnaround model at the 0.05 level, two-tailed test.

‡Significantly different from closure model at the 0.05 level, two-tailed test.

Figure B.23. Baseline percentages of English language learner (ELL) students in schools serving students in higher grades, by model



Source: State administrative data.

Notes: This figure shows regression-adjusted percentages of ELL students at baseline (the 2009–2010 school year) in schools serving students in grades 6 through 12. Percentages were regression-adjusted for state using a generalized linear model with a logit link function. There were no statistically significant differences between schools implementing different models.

ELL = English language learner.

APPENDIX C

ADDITIONAL FIGURES BASED ON SCHOOL SURVEYS

In Chapter IV, we summarized the extent to which schools reported using the practices promoted by SIG based on data from school surveys. In Section A of this appendix, we describe how we analyzed the school survey data. In Sections B through E, we present additional figures that are directly related to the analyses presented in Chapter IV. We focus on the same four topic areas addressed in Chapter IV. For each area, we present a series of figures, one for each subtopic, showing schools' use of the practices aligned with the SIG application criteria for that subtopic, similar to the figures shown in Chapter IV for each topic area. In Section F, we present additional figures that show variation across states and districts in the number of practices schools used.

A. Analysis methods

Given the large number of questions in the surveys, it is difficult to discern broad patterns or form overall conclusions by separately examining responses to individual questions. Therefore, we analyzed data from the surveys using methods designed to provide information about broad patterns in the data. Readers interested in the responses to specific survey questions can refer to Appendix E.

The process of summarizing findings involved several steps:

1. **Selecting subtopics.** For each of the four topic areas in the SIG application, we selected subtopics of interest using the SIG application criteria as a guide. For example, the section of the application criteria focusing on comprehensive instructional reform strategies identified activities in seven subsections (such as using data to identify and implement an instructional program and promoting the continuous use of data to identify and address the needs of individual students). We used each of the seven subsections as subtopics in our analysis.
2. **Selecting questions aligned with the SIG application criteria.** For each subtopic, we used a systematic approach to select survey questions that aligned with the practices that SIG sought to affect in that area (by either requiring or permitting them for specific school intervention models). First, two Mathematica researchers independently selected questions corresponding to each practice based on whether he or she determined them to be aligned with the SIG application (the agreement rate was 100 percent). It was important to have two researchers independently select questions to ensure that the questions selected for each subtopic were aligned with the SIG application criteria. Measurement of the extent to which the first and second researchers agreed on which questions were aligned with the SIG application criteria is called "inter-rater reliability" in statistics. Inter-rater reliability is traditionally measured using the percent agreement rate, calculated as the number of questions for which the first and second researchers agreed on whether or not the question was aligned with the SIG application criteria, divided by the total number of survey questions (Gwet 2014).

   We determined the topic area and subtopic into which each survey question fell based on the section of the SIG application with which it aligned. We did not use a survey question for more than one subtopic because doing so would have resulted in the question being overweighted in the overall topic area. When a question could be used for more than one subtopic, we assigned it to the subtopic (and corresponding section of the application) with which it was best aligned. The survey questions addressed all four topic areas and all but 2 of 15 subtopics from the application criteria.

3. **Constructing practice variables from survey questions.** For each practice in the SIG application for which we identified one or more relevant questions, we constructed a variable ranging from 0 to 100 percent using those questions. A value of 100 indicates that the school responded "yes" to all the questions aligned with that practice in the application, a value of 0 indicates that the school responded "yes" to none of the questions aligned with that practice, and a value between those two limits indicates that the school responded "yes" to some of the questions aligned with that practice.

   Many questions were originally structured with two response options, with a response of "yes" (recoded to a value of 100) indicating that the school reported using the practice and a response of "no" (recoded to a value of 0) indicating that the school did not report using the practice. In some cases, however, it was necessary to combine multiple survey questions to determine whether or not a school reported fully adopting a particular practice. For example, one practice in the application was that schools use teacher evaluation systems that took into account several factors. The survey asked nine separate questions about whether each of nine different measures of teacher performance (such as classroom observations and student surveys) was used for teacher evaluations. In this case, a school received less than 100 percent (in this example, a value of 11.1 percent, or one-ninth of 100 percent) for each "yes" response. This approach helped to ensure that we did not overweight some survey questions relative to how they were represented in the application.

4. **Summing the practices for each school.** To determine each school's progress in using the practices aligned with the SIG application, we summed the variables created in step 3. This sum was calculated separately for each subtopic. We then summed across subtopics to create a sum for each topic area. Thus, one or more survey *questions* were used to create a variable for each *practice*, one or more practices formed a *subtopic*, and one or more subtopics formed a *topic area* (or *area*, for short). If a particular school was missing values for a particular practice, we took the mean of the non-missing practices and multiplied it by the total number of practices for the overall area. For example, for the comprehensive instructional reform strategies area, which has eight practices, if a school had data available for five practices, and reported using two of them, the number of the school's reported used practices would be equal to (2/5)*8.

   Across all schools and all subtopics, the average percentage of practices that were missing was 4.1 percent. To assess how our coding of missing data might have affected our results, we conducted a bounding exercise in which we re-calculated the results twice: once setting all missing responses to "no" (that is, assuming all missing responses indicated that the practice was not used) and once setting all missing responses to "yes" (that is, assuming all missing responses indicated that the practice was used). The results were largely unchanged. The magnitude of differences between schools implementing a SIG-funded model in 2012–2013 and schools not implementing one with respect to the number of SIG-promoted practices used were very similar to the magnitudes reported in Chapter IV. In addition, across the eight statistical significance tests conducted as part of this bounding exercise (two for each of the four topic areas), only one result differed from what is shown in Chapter IV: when setting all missing responses to "no," the difference between schools implementing a SIG-funded model in 2012–2013 and schools not implementing one with respect to the number of SIG-promoted practices used in the area of comprehensive instructional reform strategies was no longer statistically significant.

5. **Averaging the number of practices across schools.** For each group of schools (that is, schools implementing a SIG-funded model and schools not implementing one), we averaged the numbers calculated in step 4. We calculated this average number of practices reported for the two groups of schools separately for each topic area and subtopic.

6. **Testing differences between groups of schools.** We conducted statistical tests to assess whether the average number of practices reported differed between the schools implementing a SIG-funded model in 2012–2013 and the schools not implementing one. For this analysis, we used a permutation test, which is the nonparametric counterpart to a t-test. The statistical power of this test differed by topic area and subtopic because it depended on several factors, including the number of survey questions aligned with the SIG application, the number of variables constructed from those questions, and the degree to which the variables were correlated with each other. We did not adjust the standard errors in this analysis for any type of clustering (for example, at the district or state level) because there was no random sampling or random assignment mechanism through which districts or states contributed to random variation in the results. If we had randomly sampled or randomly assigned districts or states, then it would have been appropriate to account for such variation.

   Because the goal of this analysis was to provide descriptive information about the actual levels of practices used by schools implementing a SIG-funded model and schools not implementing one, the results (that is, the mean number of practices reported by each group of schools) were not regression-adjusted to account for pre-existing differences between these two groups of schools. However, we did adjust the means to account for the district in which schools were located. This method ensured that any observed differences in reported practices between the two groups of schools reflected true differences between the two groups, rather than simply reflecting differences between the districts that included these schools. For example, certain districts had higher concentrations of SIG schools than other districts, and certain districts might have had environments that were more conducive to schools' use of SIG practices than other districts were. To account for these types of differences between districts, rather than analyzing raw (that is, unadjusted) means, we calculated regression-adjusted means using regressions that included an indicator for the district in which the school was located. This method ensured that SIG status was not confounded with school district.

   When reporting the findings from this analysis, we focused on the statistical significance of differences between schools implementing a SIG-funded model and schools not implementing one (rather than the magnitude of differences) to ensure that consistent, objective, and transparent criteria were used for reporting findings. One caveat with this approach is that some statistically significant differences might not be substantively important; we indicated places in the report where this might be the case.

We used this same method to summarize findings from the district interview. Those results are presented in Appendix D.

This method of summarizing findings is one way to analyze broad patterns observed in the data, and compare levels of usage of practices across different groups of schools. If variables had been constructed differently (for example, if multiple questions that addressed the same practice had not been combined into a single variable, but had each been included in the analysis as separate variables), the results might change. Therefore, it is important to keep these methods in mind when interpreting the results.

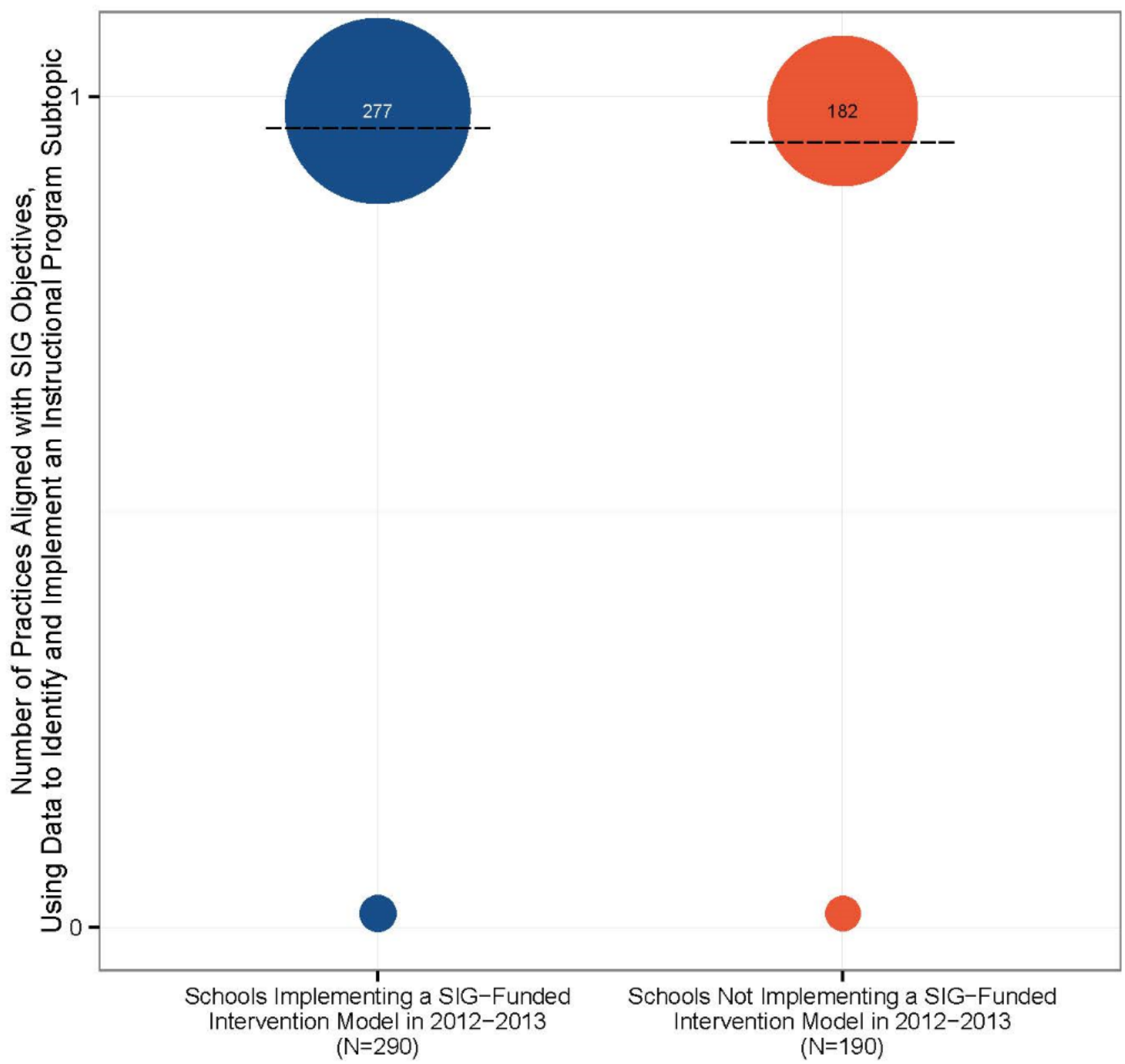
Additional caveats to keep in mind are: the findings are based on self-reported use of these practices, 2 of the 15 SIG objectives were not addressed by our study instruments, we did not collect information about the quality, fidelity, scope, or intensity with which the practices were implemented, and the sample of schools was not randomly selected. For these reasons, the findings from this analysis should be interpreted with caution.

Here we provide several potential explanations for how these limitations might affect the data. Our data do not allow us to determine whether any of these possible explanations are correct, but we offer them as starting points for thinking about how the results might be affected by them. Self-reported levels of practice use might be overestimated (relative to actual use) if schools provided socially desirable responses. This would likely lead to overestimated levels for all schools. The study team took several steps to ensure that schools provided accurate responses, including telling schools that the survey was not an audit and that we would report aggregated responses across schools rather than singling out any individual school. That self-reported levels of practice use (as presented in Chapter IV) are not all 100 percent, and in many cases are much lower than 100 percent, suggests that many respondents did not feel compelled to provide socially desirable responses. However, it is possible that schools implementing a SIG-funded model might have been more likely than schools not implementing one to provide socially desirable responses, given that they received SIG funds to implement the practices we examined. Therefore, the results for schools implementing a SIG-funded model might be more inflated than the results for schools not implementing one, so readers should use caution when interpreting the results. In addition, our study instruments did not address 2 out of 15 SIG objectives, but this factor is unlikely to have a large effect on the overall results because very few objectives were not addressed. Regarding the quality, fidelity, scope, or intensity with which the practices were implemented, our data might overestimate use of practices if schools tended to report that they used a practice if they had at least begun to use it but had not necessarily implemented it fully. However, this would lead to overestimates of use for all schools rather than affecting the differences between the two groups of schools.

Schools implementing the closure model were handled as follows. Schools that had already implemented the closure model as of spring 2013 were not surveyed and were not included in the analysis. Schools that were planning to implement the closure model but had not yet closed as of spring 2013 were surveyed and included in the analysis, for three reasons: (1) Dropping these schools from the analysis would have been inconsistent with how we treated schools that were planning to implement other models, but had not yet implemented a particular practice required by that model. For example, if a school was planning to implement the transformation model but had not yet replaced their principal, we still treated them as schools implementing a SIG-funded model in 2012–2013. Similarly, we treated schools that were planning to close but had not yet closed as of spring 2013 as schools implementing a SIG-funded model in 2012–2013; (2) The SIG application guidance indicated that closure schools may use SIG funds to cover the activities (such as community outreach) that were recommended before closing the school. Because these schools were receiving SIG funds to implement the closure model and the associated activities that preceded the closure, we included them in the analysis for the years before they closed; and (3) The process of closure was not always immediate: some schools closed by allowing current students to finish, but ending enrollment of additional students (that is, the lowest grade closed first, then the next lowest, and so on, until the school was shut down).

B. Comprehensive instruction reform strategies

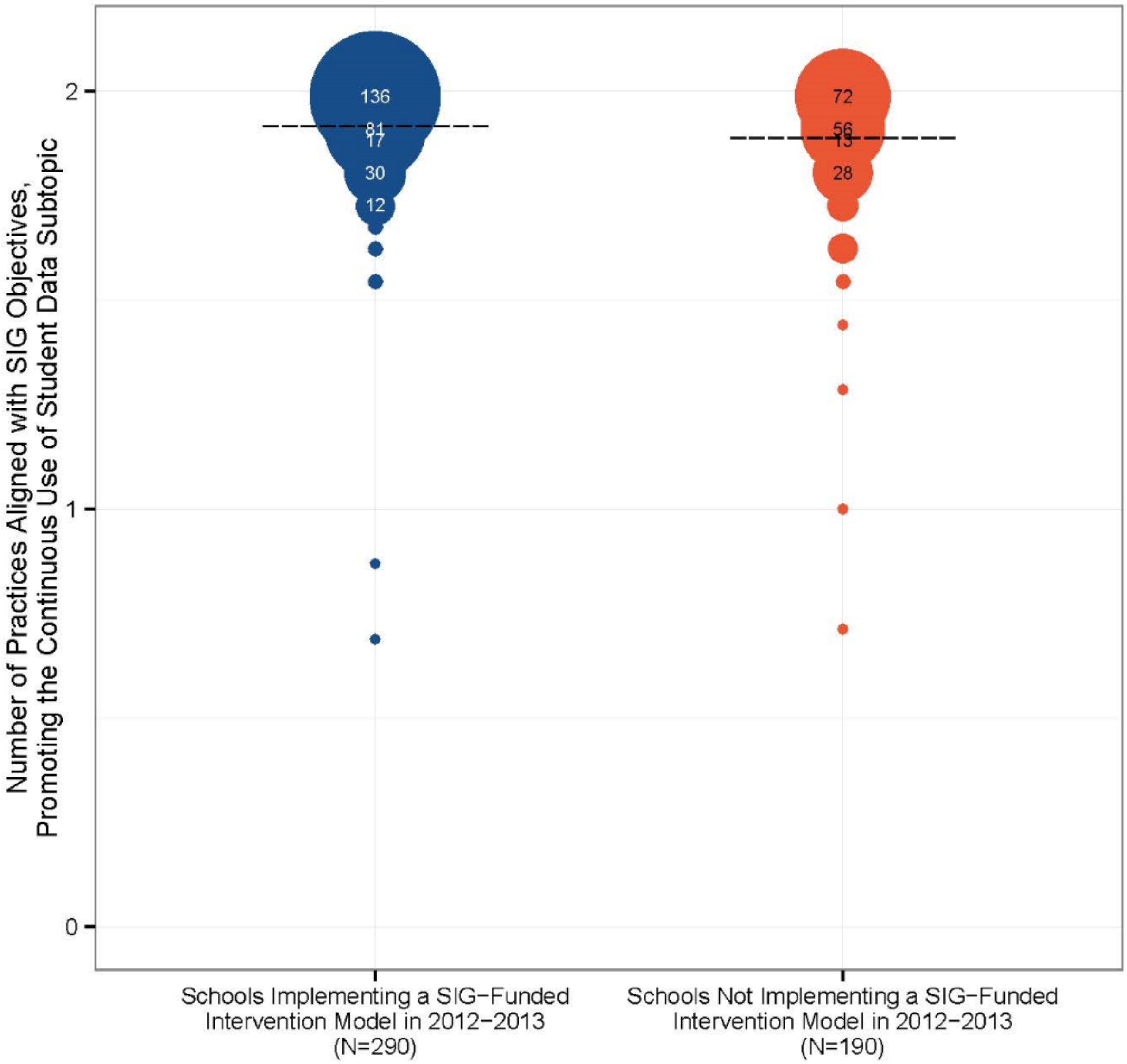
Figure C.1. Use of practices aligned with SIG, using data to identify and implement an instructional program subtopic



Source: Surveys of school administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table IV.1. Each dot in this figure represents the schools that reported using the one practice that was aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot. To protect respondent confidentiality, the number inside the smallest dot for each group of schools has been removed. For this practice, a “yes” response received one point. See Section A of this appendix for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools. There were no statistically significant differences between schools implementing a SIG-funded intervention model in 2012–2013 and schools not implementing one at the 0.05 level using a two-tailed test.

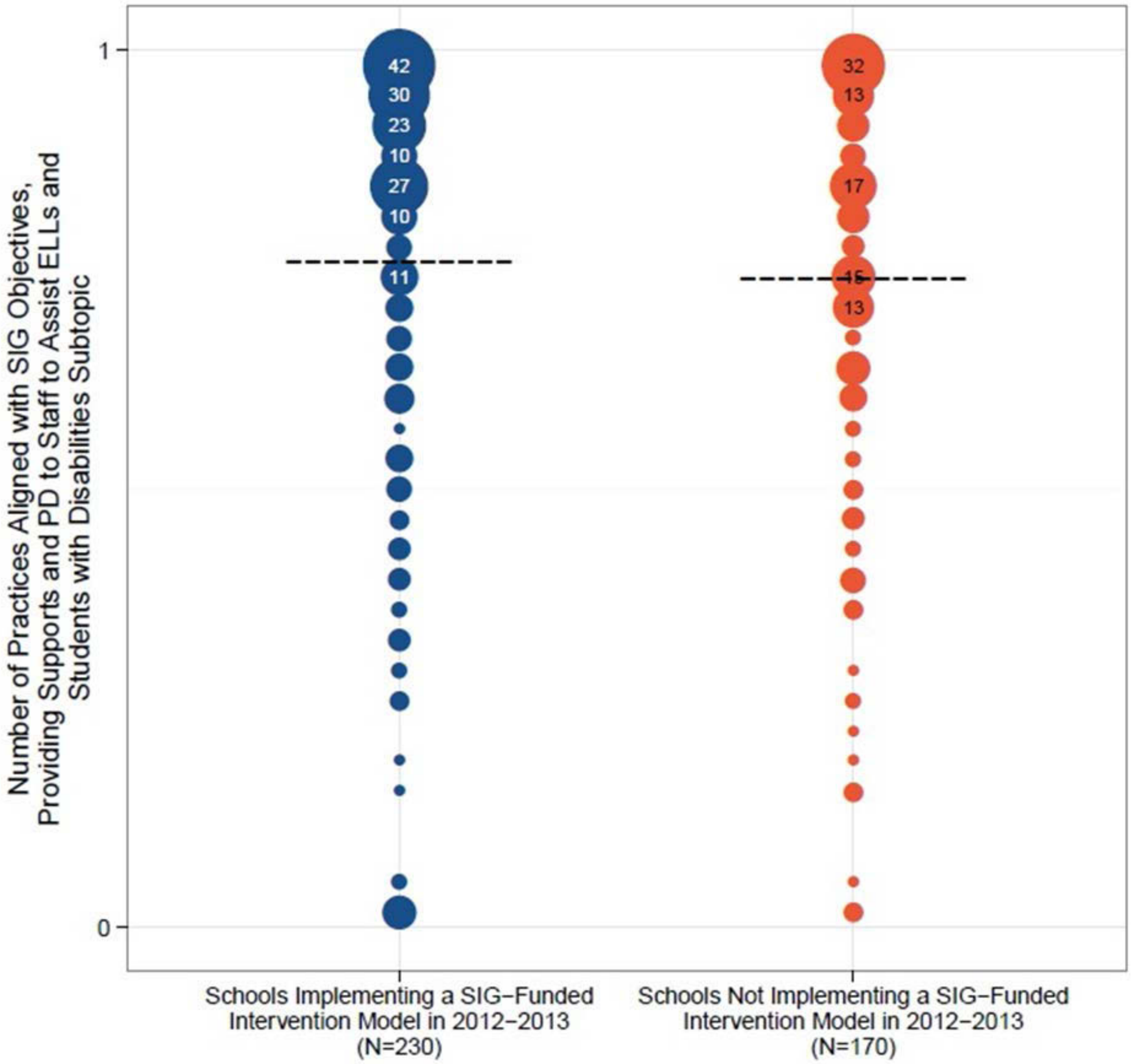
Figure C.2. Use of practices aligned with SIG, promoting the continuous use of student data subtopic



Source: Surveys of school administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table IV.1. Each dot in this figure represents the schools that reported using a particular number of practices (out of two examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For both practices, it was possible for a school to receive a fraction of one point. See Section A of this appendix for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools. There were no statistically significant differences between schools implementing a SIG-funded intervention model in 2012–2013 and schools not implementing one at the 0.05 level using a two-tailed test.

Figure C.3. Use of practices aligned with SIG, providing supports and professional development to staff to assist English language learners and students with disabilities subtopic

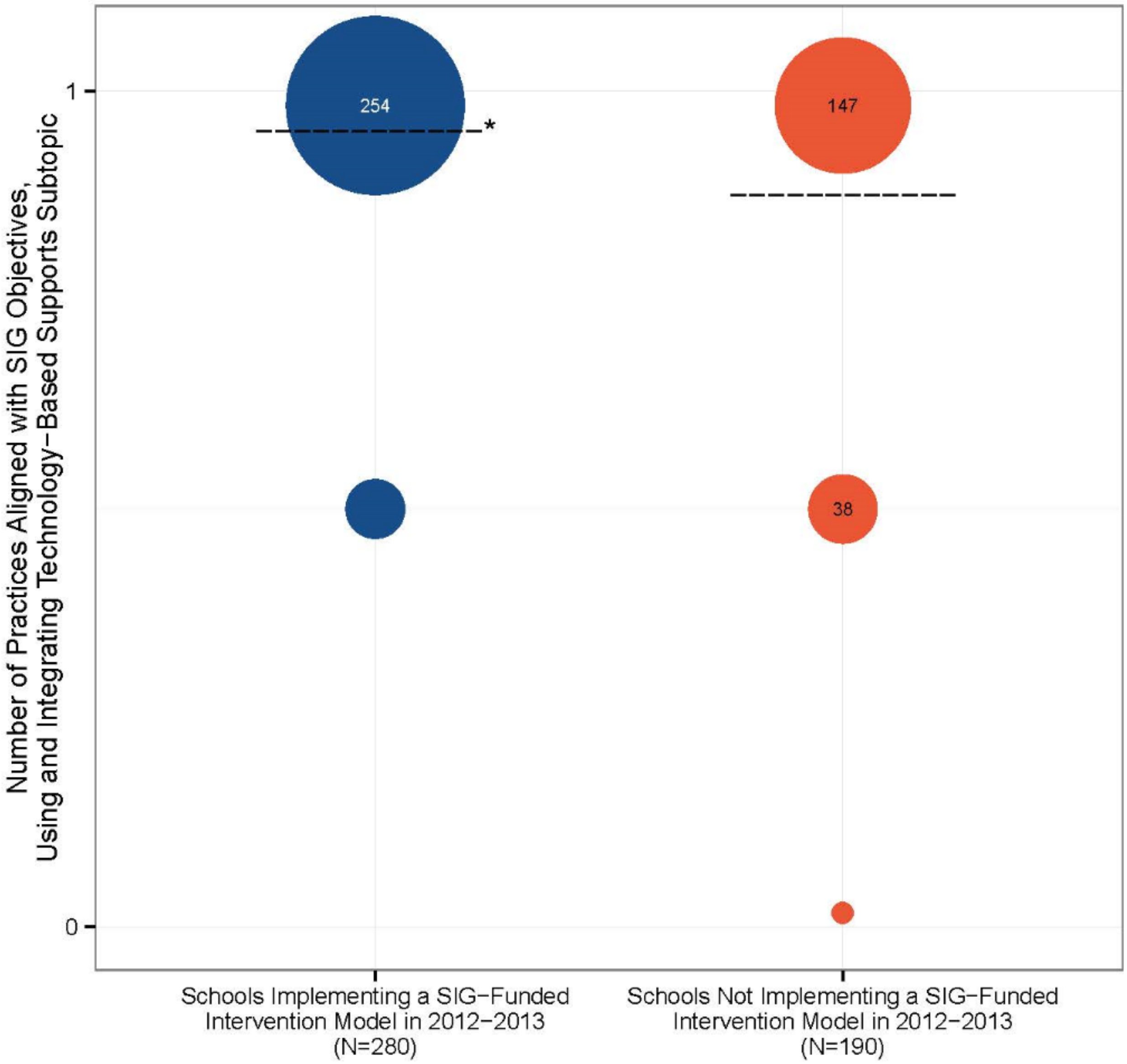


Source: Surveys of school administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table IV.1. This figure presents one practice described in the SIG application criteria to which multiple survey questions aligned. As described in Section A of this appendix, whenever multiple survey questions aligned with a single practice from the application criteria, we used those questions to construct a variable ranging from zero to one, with schools receiving a fraction of a point for each question to which they responded "yes." Each dot in this figure represents the schools that reported using a particular proportion of the survey questions aligned to the practice described in the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. The dashed line denotes the average value for each group of schools. There were no statistically significant differences between schools implementing a SIG-funded intervention model in 2012–2013 and schools not implementing one at the 0.05 level using a two-tailed test.

ELLs = English language learners; PD = professional development.

Figure C.4. Use of practices aligned with SIG, using and integrating technology-based supports subtopic

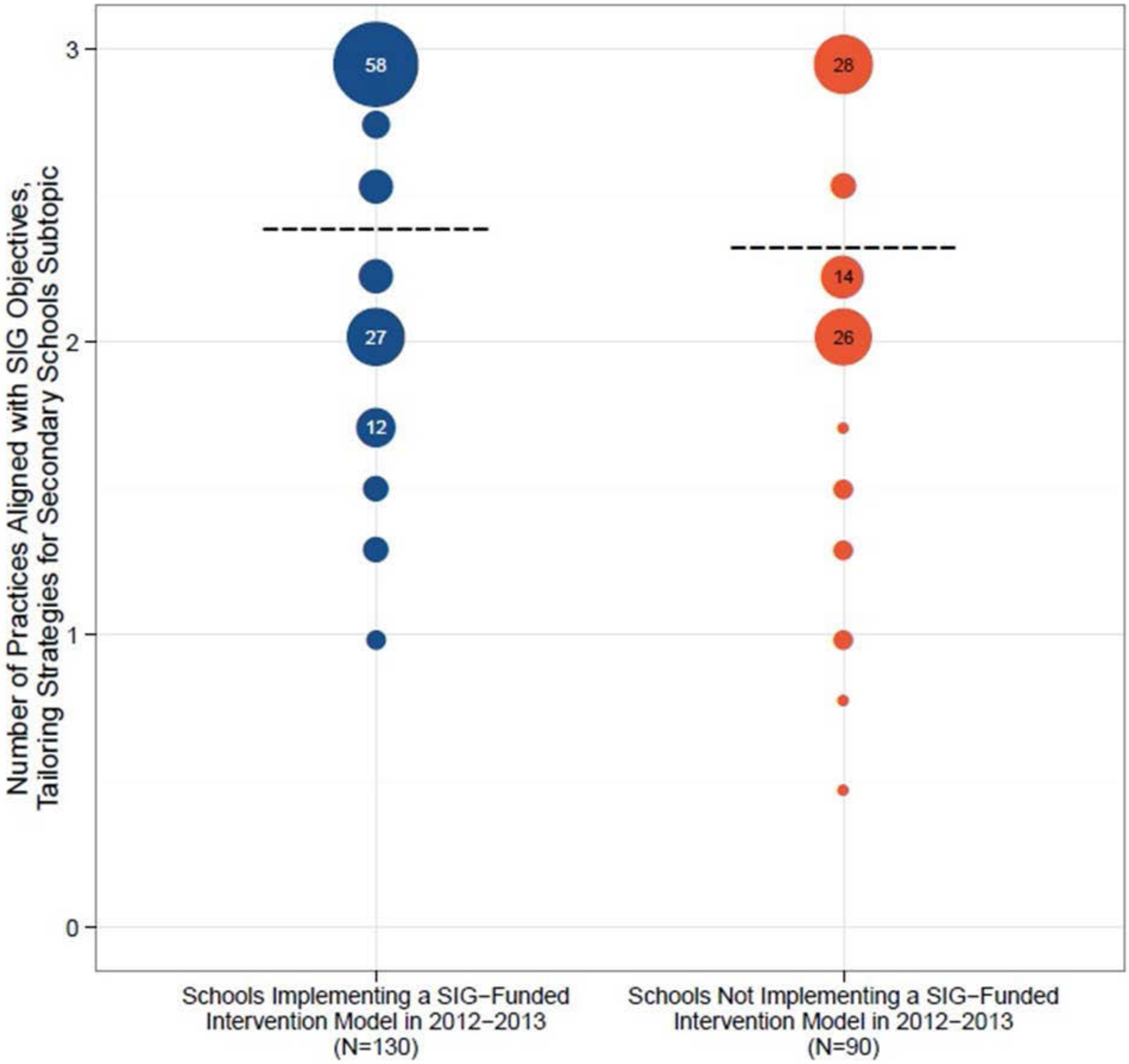


Source: Surveys of school administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table IV.1. This figure presents one practice described in the SIG application criteria to which multiple survey questions aligned. As described in Section A of this appendix, whenever multiple survey questions aligned with a single practice from the application criteria, we used those questions to construct a variable ranging from zero to one, with schools receiving a fraction of a point for each question to which they responded "yes." Each dot in this figure represents the schools that reported using a particular proportion of the survey questions aligned to the practice described in the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. The dashed line denotes the average value for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

Figure C.5. Use of practices aligned with SIG, tailoring strategies for secondary schools subtopic

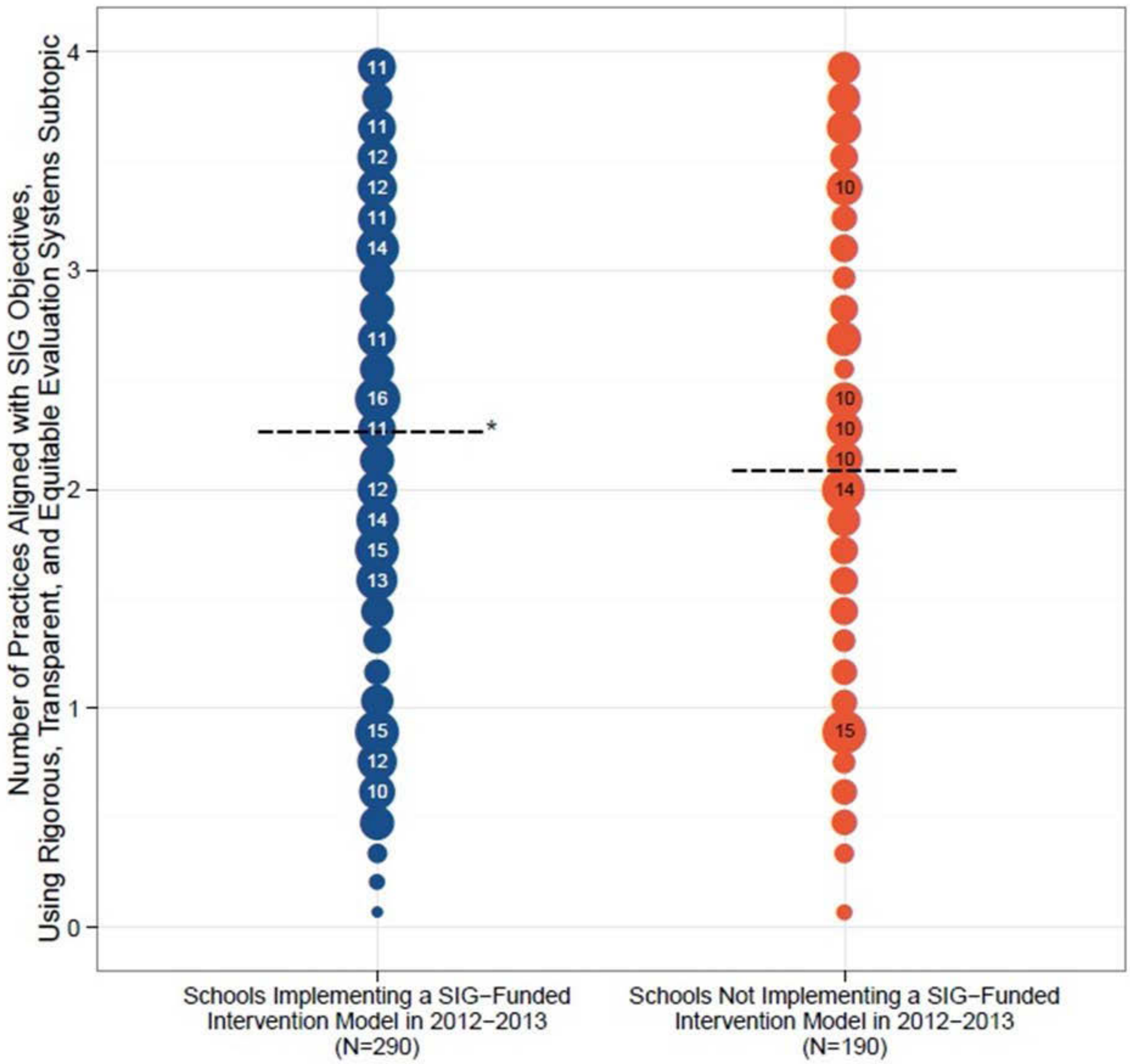


Source: Surveys of school administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table IV.1. Each dot in this figure represents the schools that reported using a particular number of practices (out of three examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For one practice, a “yes” response received one point. In the other two cases, it was possible for a school to receive a fraction of one point. See Section A of this appendix for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools. There were no statistically significant differences between schools implementing a SIG-funded intervention model in 2012–2013 and schools not implementing one at the 0.05 level using a two-tailed test.

C. Teacher and principal effectiveness

Figure C.6. Use of practices aligned with SIG, using rigorous, transparent, and equitable evaluation systems subtopic

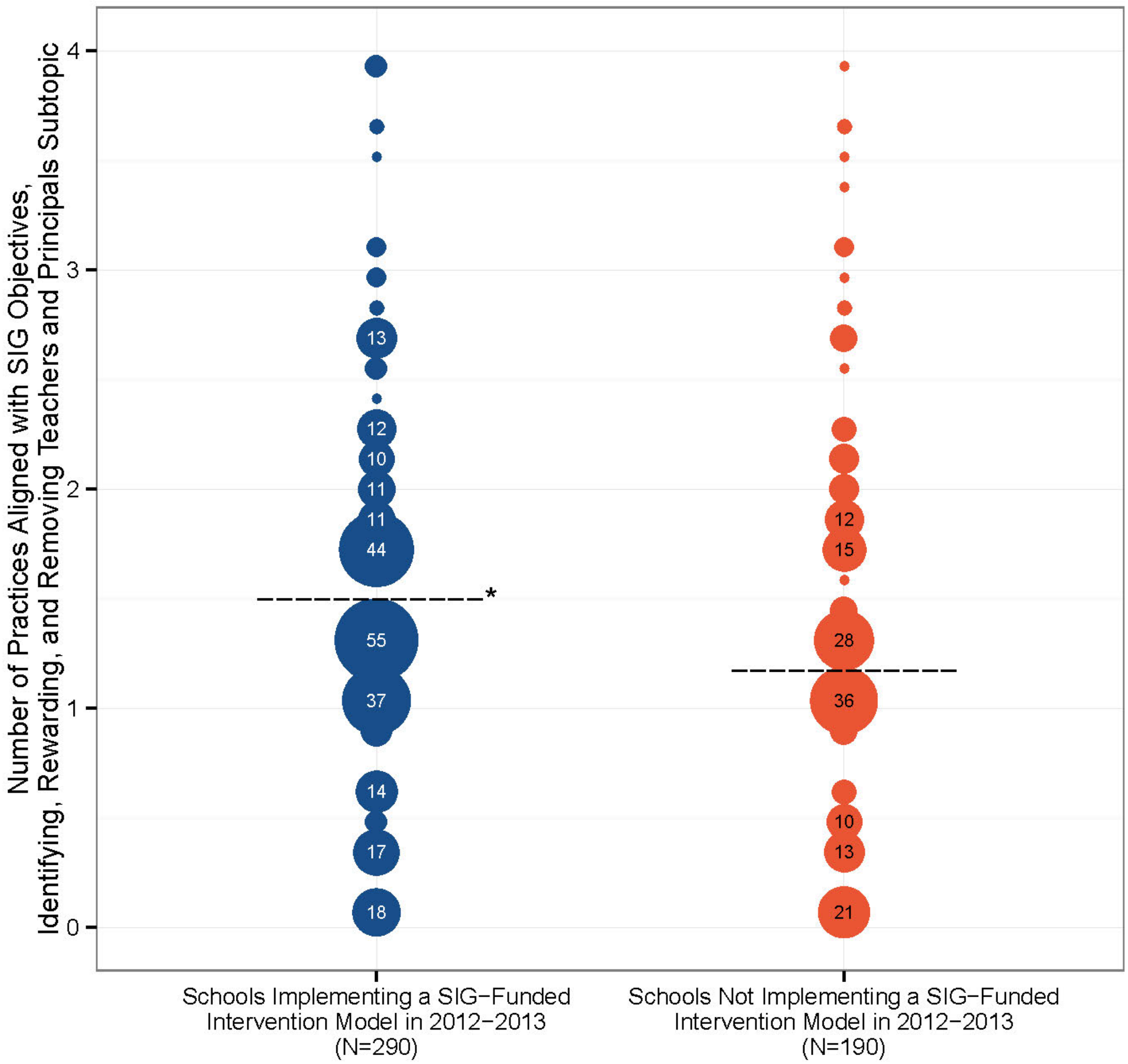


Source: Surveys of school administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table IV.2. Each dot in this figure represents the schools that reported using a particular number of practices (out of four examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For one practice, a “yes” response received one point. In the other three cases, it was possible for a school to receive a fraction of one point. See Section A of this appendix for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

Figure C.7. Use of practices aligned with SIG, identifying and rewarding effective teachers and principals and removing ineffective ones subtopic

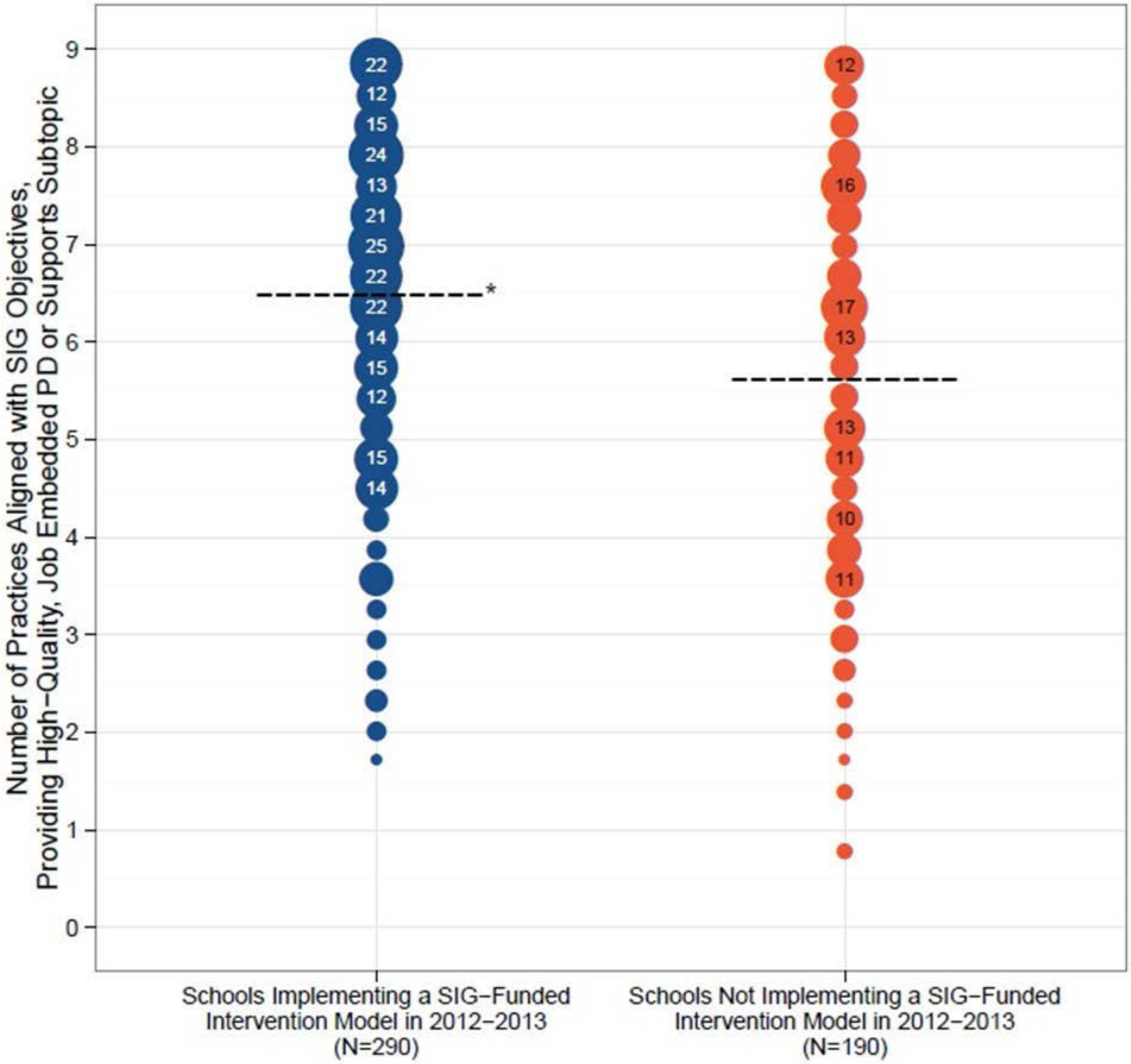


Source: Surveys of school administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table IV.2. Each dot in this figure represents the schools that reported using a particular number of practices (out of four examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For all four practices, it was possible for a school to receive a fraction of one point. See Section A of this appendix for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

Figure C.8. Use of practices aligned with SIG, job-embedded professional development or supports subtopic



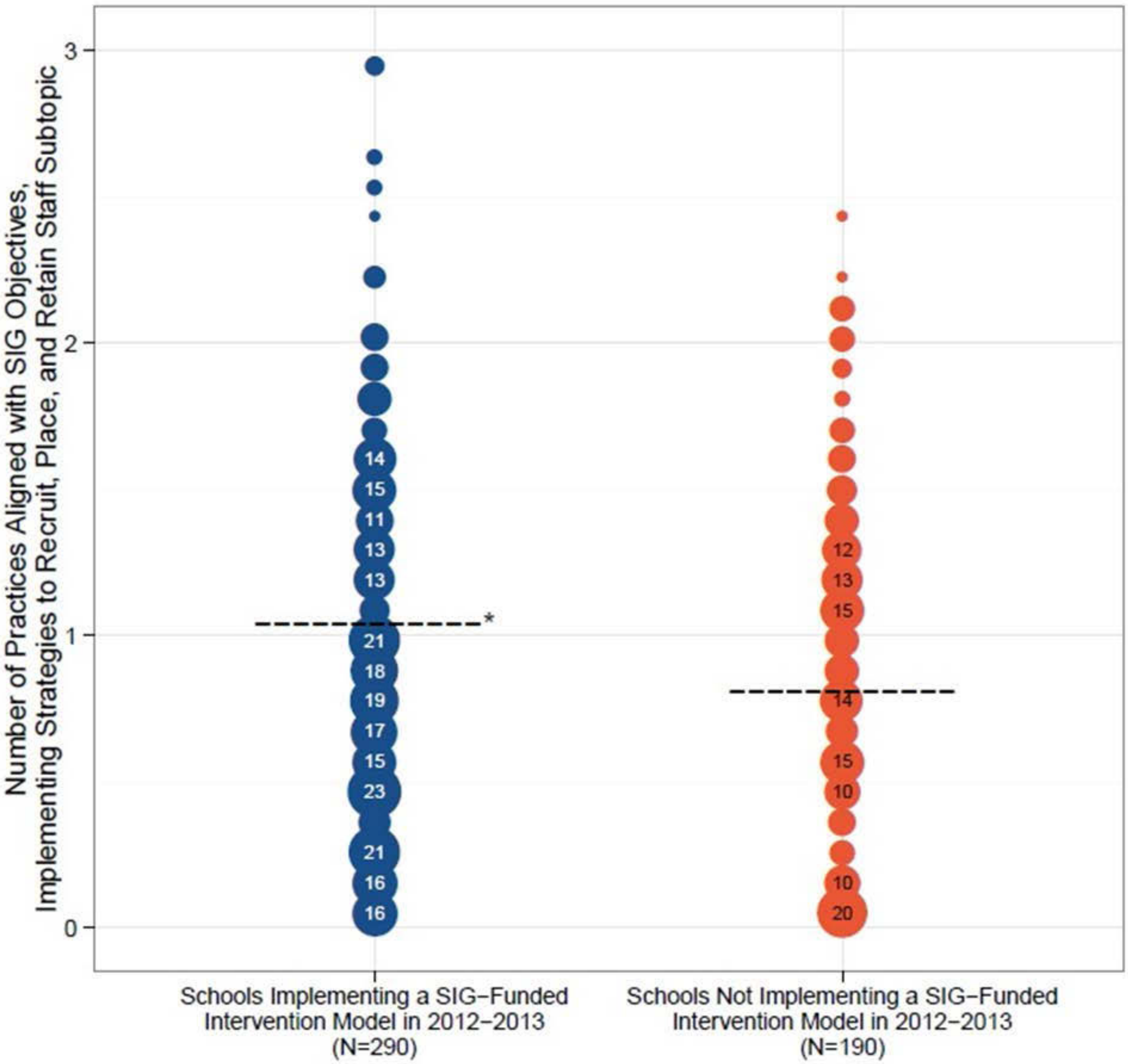
Source: Surveys of school administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table IV.2. Each dot in this figure represents the schools that reported using a particular number of practices (out of nine examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For four of the practices, a "yes" response received one point. In the other five cases, it was possible for a school to receive a fraction of one point. See Section A of this appendix for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools.

PD = professional development.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

Figure C.9. Use of practices aligned with SIG, implementing strategies to recruit, place, and retain staff subtopic



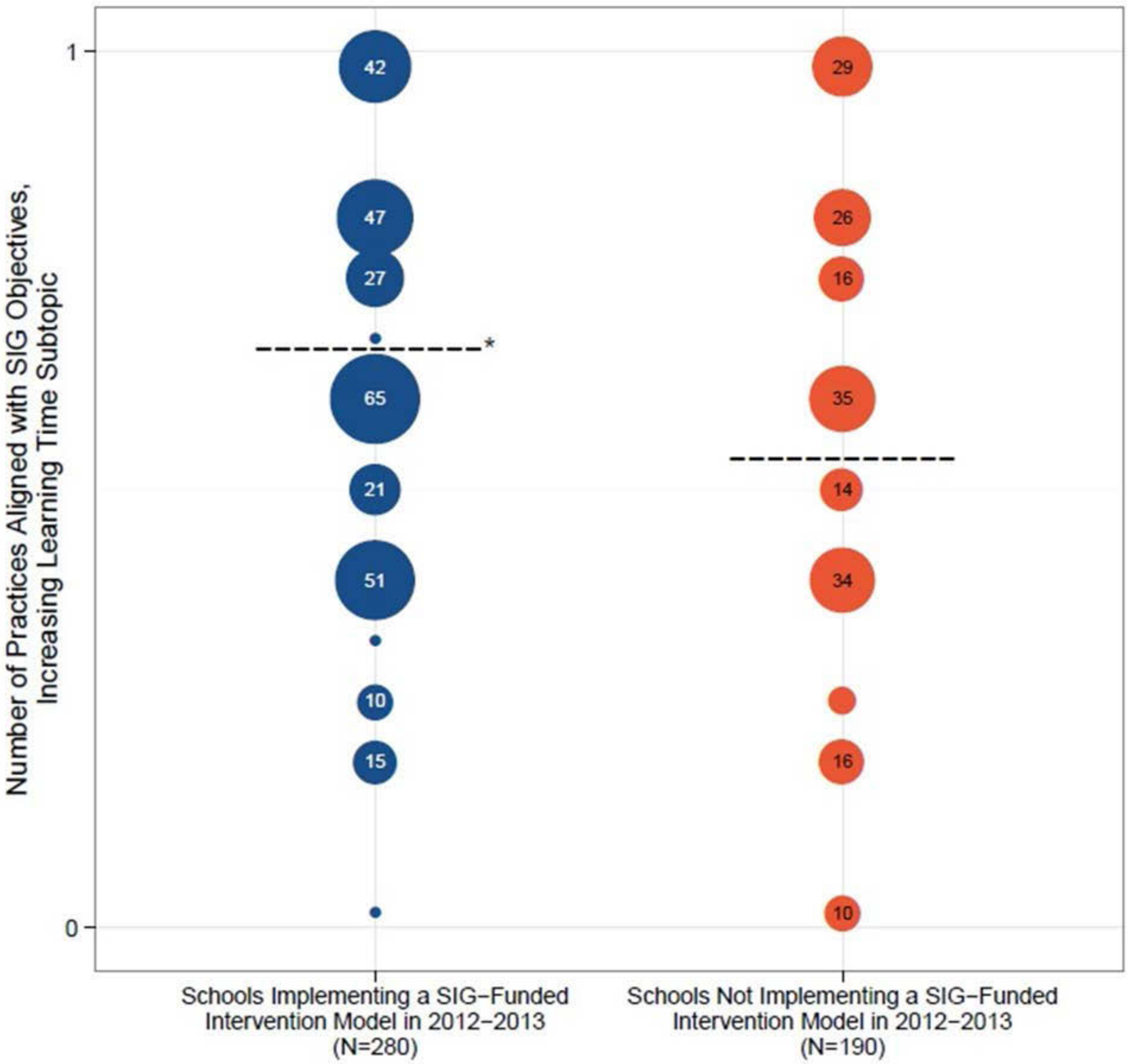
Source: Surveys of school administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table IV.2. Each dot in this figure represents the schools that reported using a particular number of practices (out of three examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For all three practices, it was possible for a school to receive a fraction of one point. See Section A of this appendix for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

D. Learning time and community-oriented schools

Figure C.10. Use of practices aligned with SIG, increasing learning time subtopic

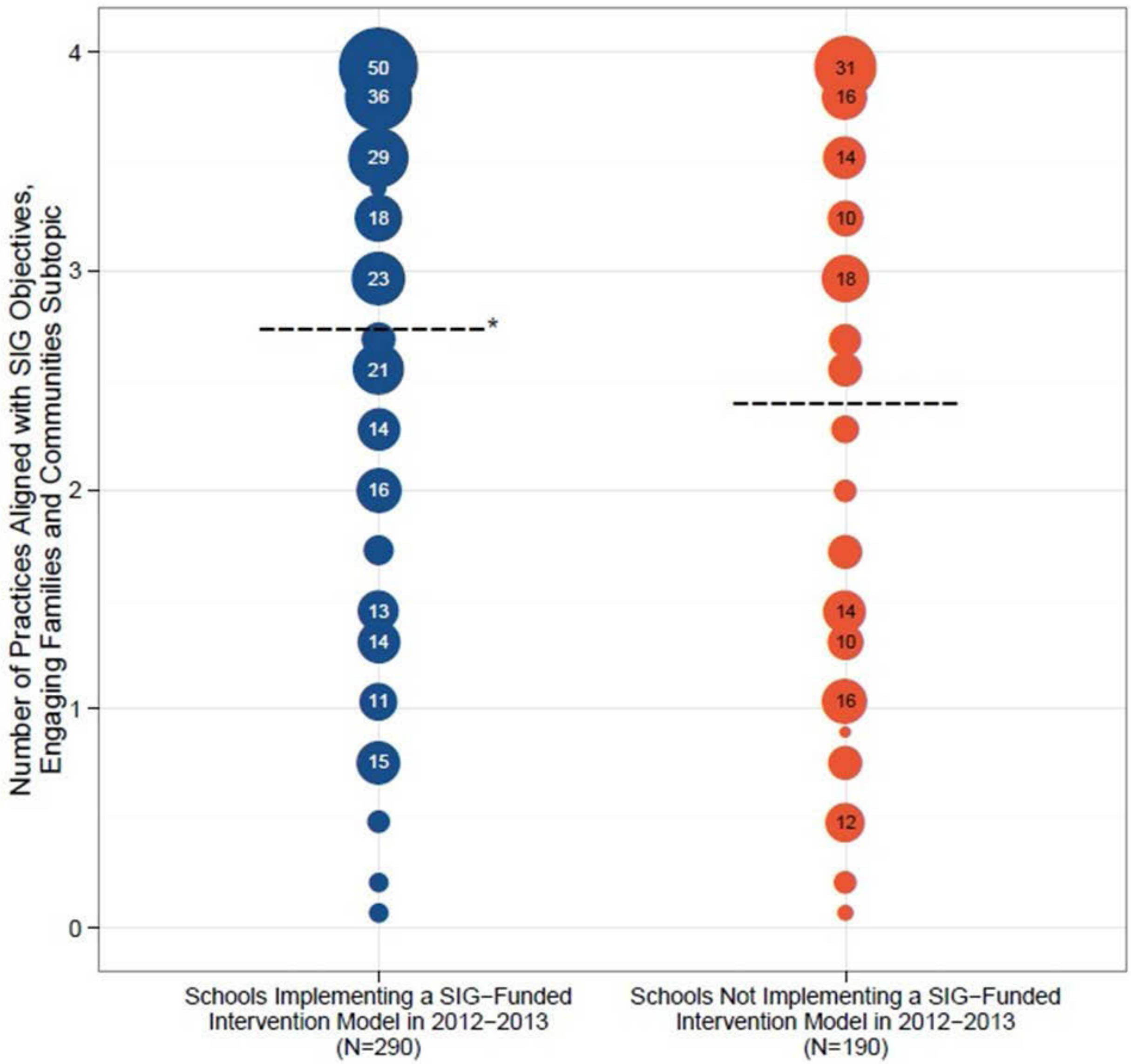


Source: Surveys of school administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table IV.3. This figure presents one practice described in the SIG application criteria to which multiple survey questions aligned. As described in Section A of this appendix, whenever multiple survey questions aligned with a single practice from the application criteria, we used those questions to construct a variable ranging from zero to one, with schools receiving a fraction of a point for each question to which they responded "yes." Each dot in this figure represents the schools that reported using a particular proportion of the survey questions aligned to the practice described in the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. The dashed line denotes the average value for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

Figure C.11. Use of practices aligned with SIG, engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs subtopic



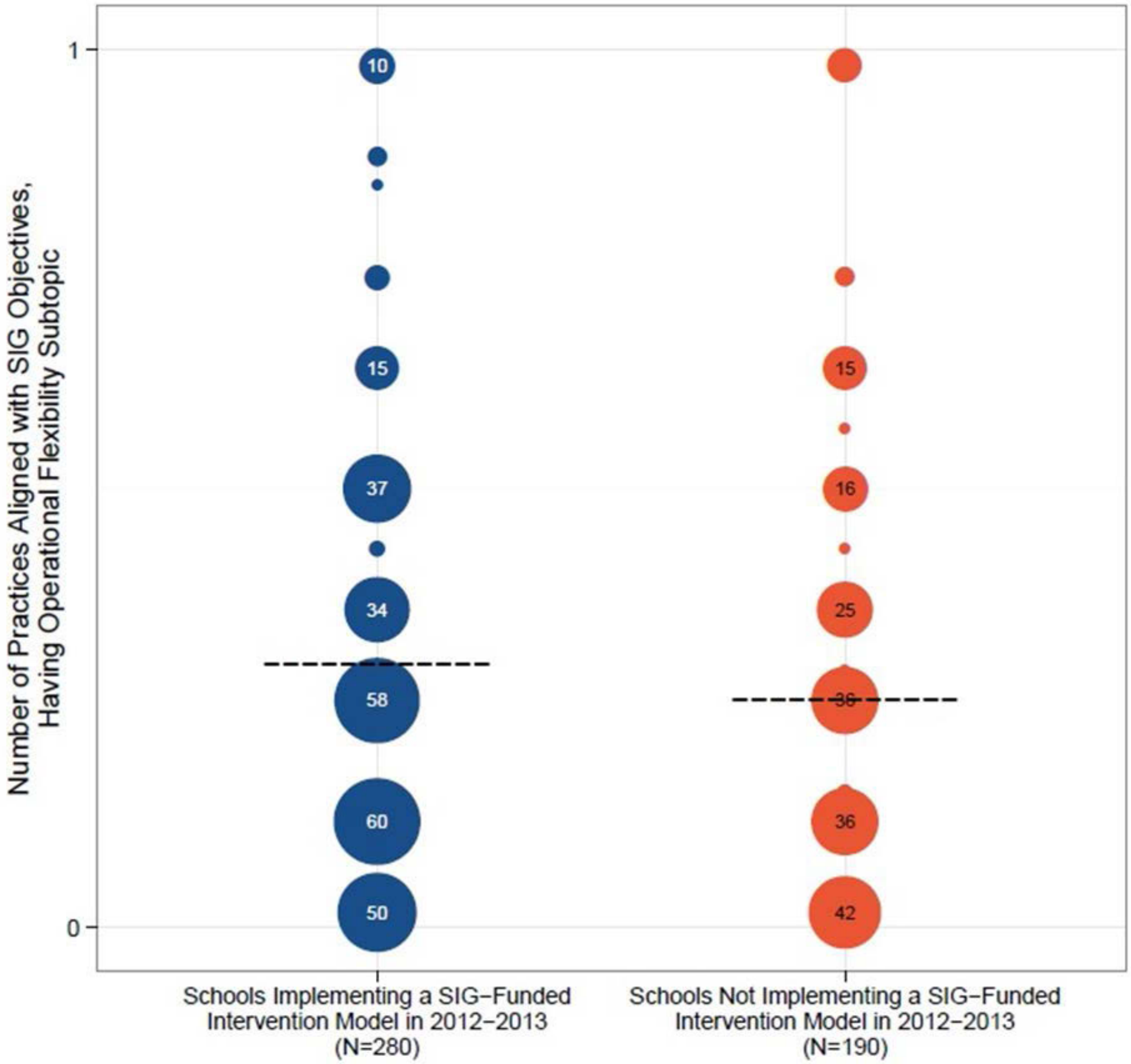
Source: Surveys of school administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table IV.3. Each dot in this figure represents the schools that reported using a particular number of practices (out of four examined) that were aligned with the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. For two of the practices, a "yes" response received one point. In the other cases, it was possible for a school to receive a fraction of one point. See Section A of this appendix for details on how we determined the number of practices for each school. The dashed line denotes the average number of practices for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

E. Operational flexibility and support

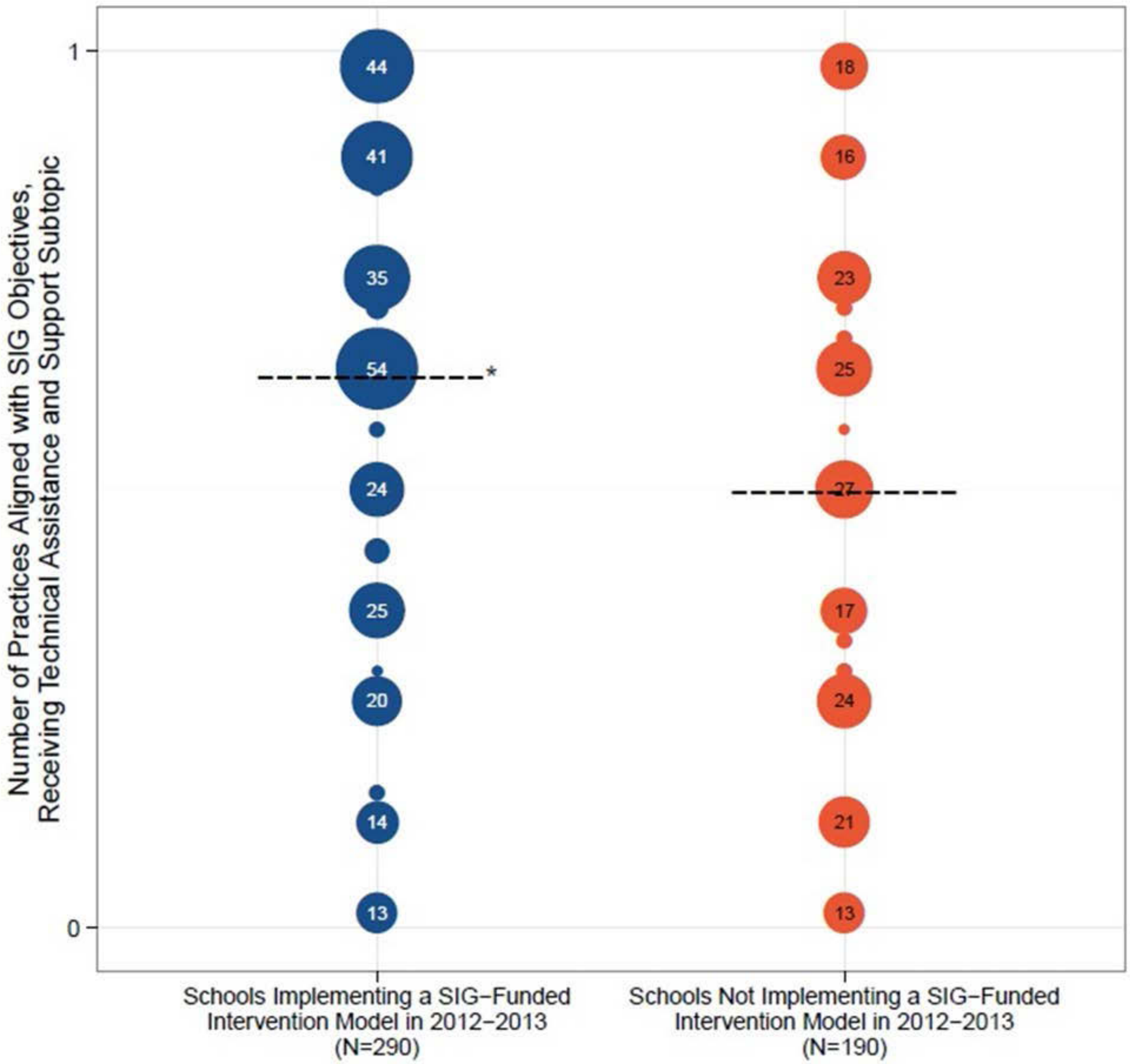
Figure C.12. Use of practices aligned with SIG, having operational flexibility subtopic



Source: Surveys of school administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table IV.4. This figure presents one practice described in the SIG application criteria to which multiple survey questions aligned. As described in Section A of this appendix, whenever multiple survey questions aligned with a single practice from the application criteria, we used those questions to construct a variable ranging from zero to one, with schools receiving a fraction of a point for each question to which they responded “yes.” Each dot in this figure represents the schools that reported using a particular proportion of the survey questions aligned to the practice described in the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. The dashed line denotes the average value for each group of schools. There were no statistically significant differences between schools implementing a SIG-funded intervention model in 2012–2013 and schools not implementing one at the 0.05 level using a two-tailed test.

Figure C.13. Use of practices aligned with SIG, receiving technical assistance and support subtopic



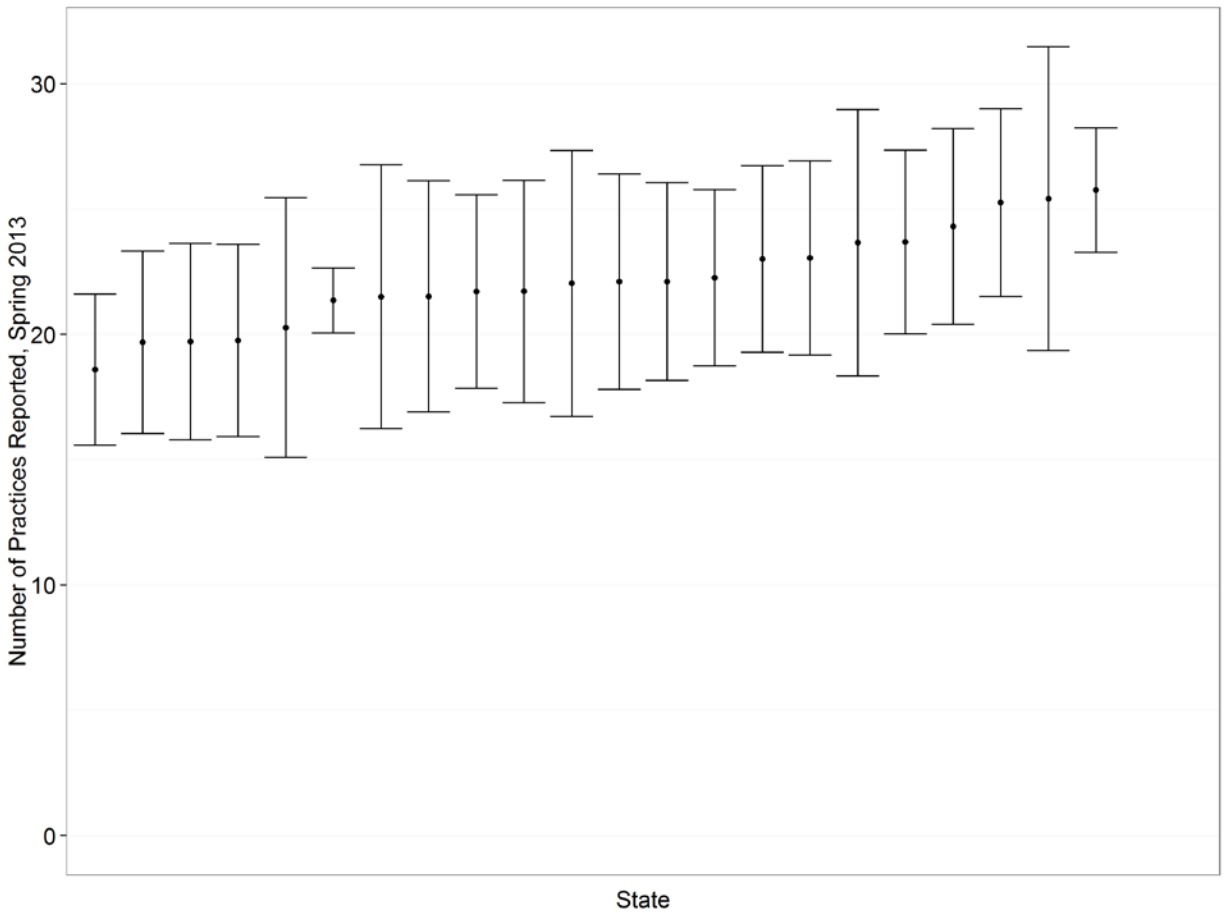
Source: Surveys of school administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table IV.4. This figure presents one practice described in the SIG application criteria to which multiple survey questions aligned. As described in Section A of this appendix, whenever multiple survey questions aligned with a single practice from the application criteria, we used those questions to construct a variable ranging from zero to one, with schools receiving a fraction of a point for each question to which they responded “yes.” Each dot in this figure represents the schools that reported using a particular proportion of the survey questions aligned to the practice described in the SIG application criteria. The number inside each dot is the number of schools represented by the dot; dots that represent fewer than 10 schools have no number inside. The dashed line denotes the average value for each group of schools.

*Significantly different from schools not implementing a SIG-funded intervention model at the 0.05 level, two-tailed test.

E. Variation across states and districts in the number of practices schools used

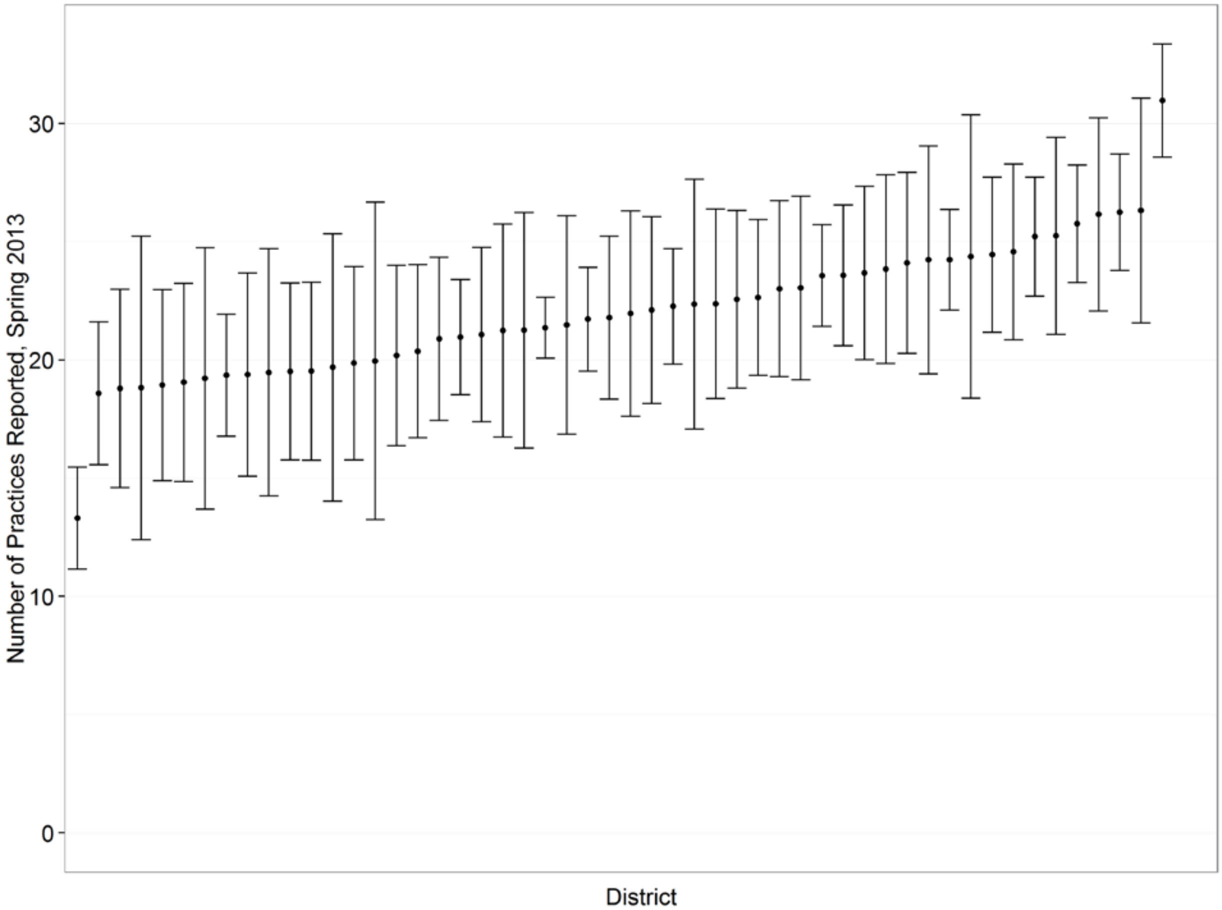
Figure C.14. Average number of practices that schools reported using, by state



Source: Surveys of school administrators in spring 2013.

Note: For each state, we calculated the average number of practices aligned with the SIG application criteria that schools in that state reported using. Each line in this figure represents one state and presents its average (indicated by a dot) and +/- 1 standard deviation from the average (indicated by the black vertical lines). The analysis includes 290 schools implementing a SIG-funded intervention model in 2012–2013 and 190 schools not implementing a SIG-funded intervention model in 2012–2013 in 22 states.

Figure C.15. Average number of practices that schools reported using, by district



Source: Surveys of school administrators in spring 2013.

Note: For each district, we calculated the average number of practices aligned with the SIG application criteria that schools in that district reported using. Each line in this figure represents one district and presents its average (indicated by a dot) and +/- 1 standard deviation from the average (indicated by the black vertical lines). The analysis includes 280 schools implementing a SIG-funded intervention model in 2012–2013 and 190 schools not implementing a SIG-funded intervention model in 2012–2013. Ten of the 60 districts had fewer than three schools, so they were excluded from this analysis.

APPENDIX D

DISTRICT REPORTED PRACTICES ALIGNED WITH THE SIG APPLICATION CRITERIA

In contrast to the main body of the report and Appendix C, which summarized the extent to which *schools* reported using practices promoted by a School Improvement Grant (SIG), this appendix summarizes the extent to which *district* administrators reported using the practices promoted by SIG in spring 2013. The overarching research question answered by these district findings is: *How are districts supporting schools' efforts to use practices promoted by SIG?* For example, some of the school survey questions asked schools if they received particular types of support from districts or states. The findings in this appendix shed light on the extent to which districts reported providing those types of support.

In this appendix, we focus on the same four topic areas addressed in Chapter IV: (1) adopting comprehensive instructional reform strategies, (2) developing and increasing teacher and principal effectiveness, (3) increasing learning time and creating community-oriented schools, and (4) having operational flexibility and receiving support. For each area, we first present a table that shows the practices from the district interview that aligned with the SIG application criteria. We then present a series of figures that display the results. The first figure displays the results of the overall analysis for the area. The figures that follow display the results for each subtopic within that topic area.

The data presented in this appendix came from structured telephone interviews with administrators in the 60 districts where the SIG-sample schools were located. The interviews, conducted in spring 2012 and 2013, documented the school turnaround practices being used and addressed both state- and district-level supports for those practices. The district interview protocols for spring 2012 and 2013 are available at http://www.mathematica-mpr.com/~/media/publications/pdfs/spring_2012_district_interview_protocol.pdf and http://www.mathematica-mpr.com/~/media/publications/pdfs/education/spring_2013_district_interview_protocol.pdf. We used the same methods to summarize findings from the district interview that we used to summarize findings from the school survey (which are described in Appendix C).

One important difference between the figures shown in Chapter IV and the figures shown in this appendix is that the latter have no comparison group. All districts in the study sample included schools that were and were not implementing a SIG-funded intervention model. Therefore, in this appendix, we are not presenting comparisons between districts; instead, we are presenting descriptive information about the practices that study districts reported using.

A. Comprehensive instructional reform strategies

The spring 2013 district interview asked about seven practices aligned with SIG objectives on comprehensive instructional reform strategies (Table D.1).

Table D.1. Practices aligned with SIG objectives on comprehensive instructional reform strategies, by subtopic

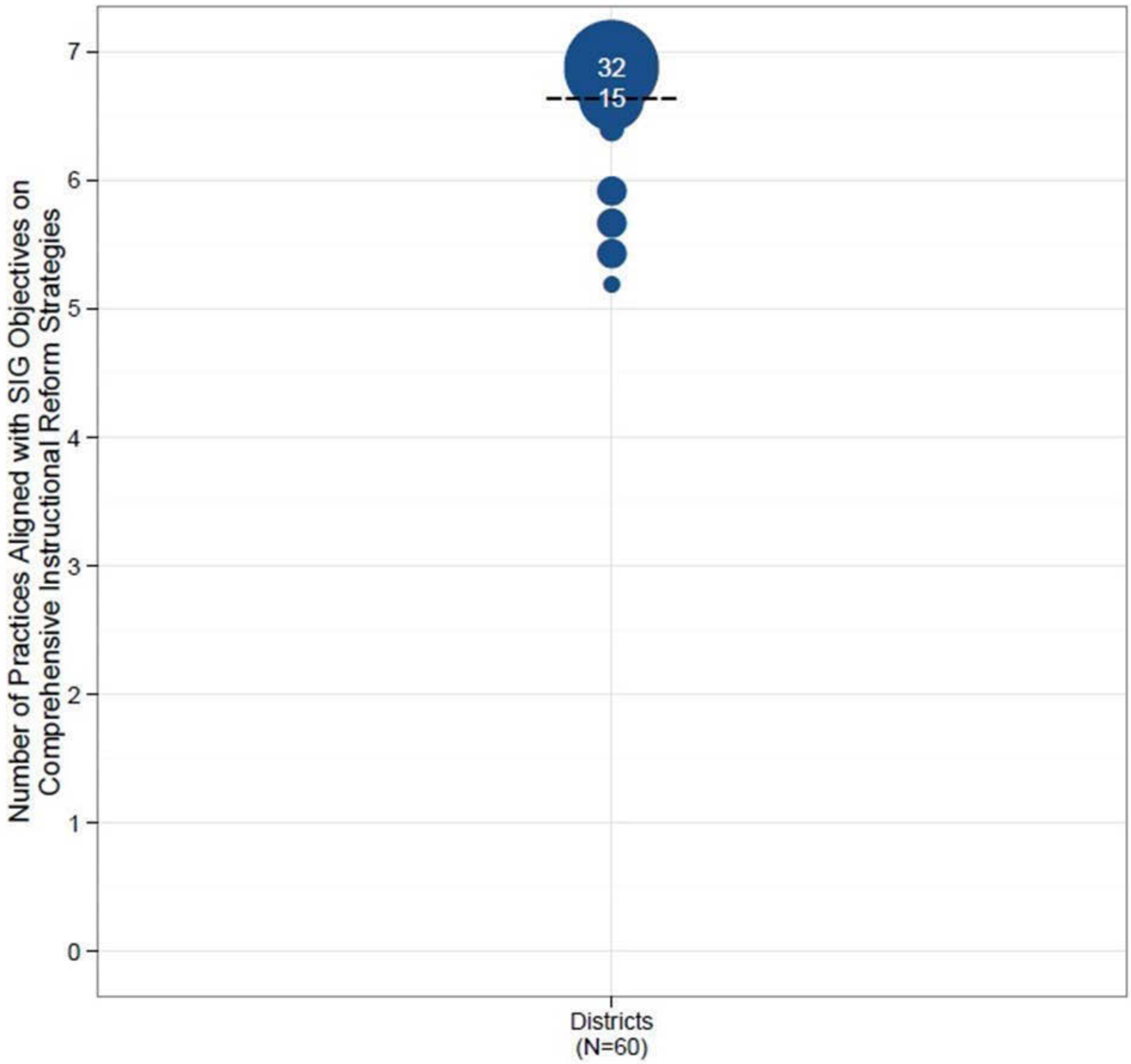
Using data to identify and implement an instructional program
Using data to evaluate instructional programs (for example, measuring program effectiveness)
English language arts or math curricula were aligned with state academic standards
Promoting the continuous use of student data
Using data to track or monitor the performance of SIG schools or to inform and differentiate instruction
Using interim or benchmark tests for English language arts or math
Providing supports and professional development to staff to assist ELLs and students with disabilities
Implementing strategies (including additional supports or professional development) to ensure that limited English proficient students acquire language skills to master academic content
Providing additional supports and programs to students with disabilities
Tailoring strategies for secondary schools
Using data to track attendance, graduation rates, or student progress toward grade promotion or graduation

Source: SIG application; interviews with district administrators in spring 2013.

ELLs = English language learners.

Figure D.1 displays results of the analysis on the extent to which district administrators reported using the comprehensive instructional reform strategies aligned with the SIG application criteria. Figure D.2 displays the change over time in districts' reported use of the practices included in the analysis for this area. Figure D.3 displays the extent to which districts reported using the individual practices included in the analysis for this area. Figures D.4–D.7 display the results for each subtopic.

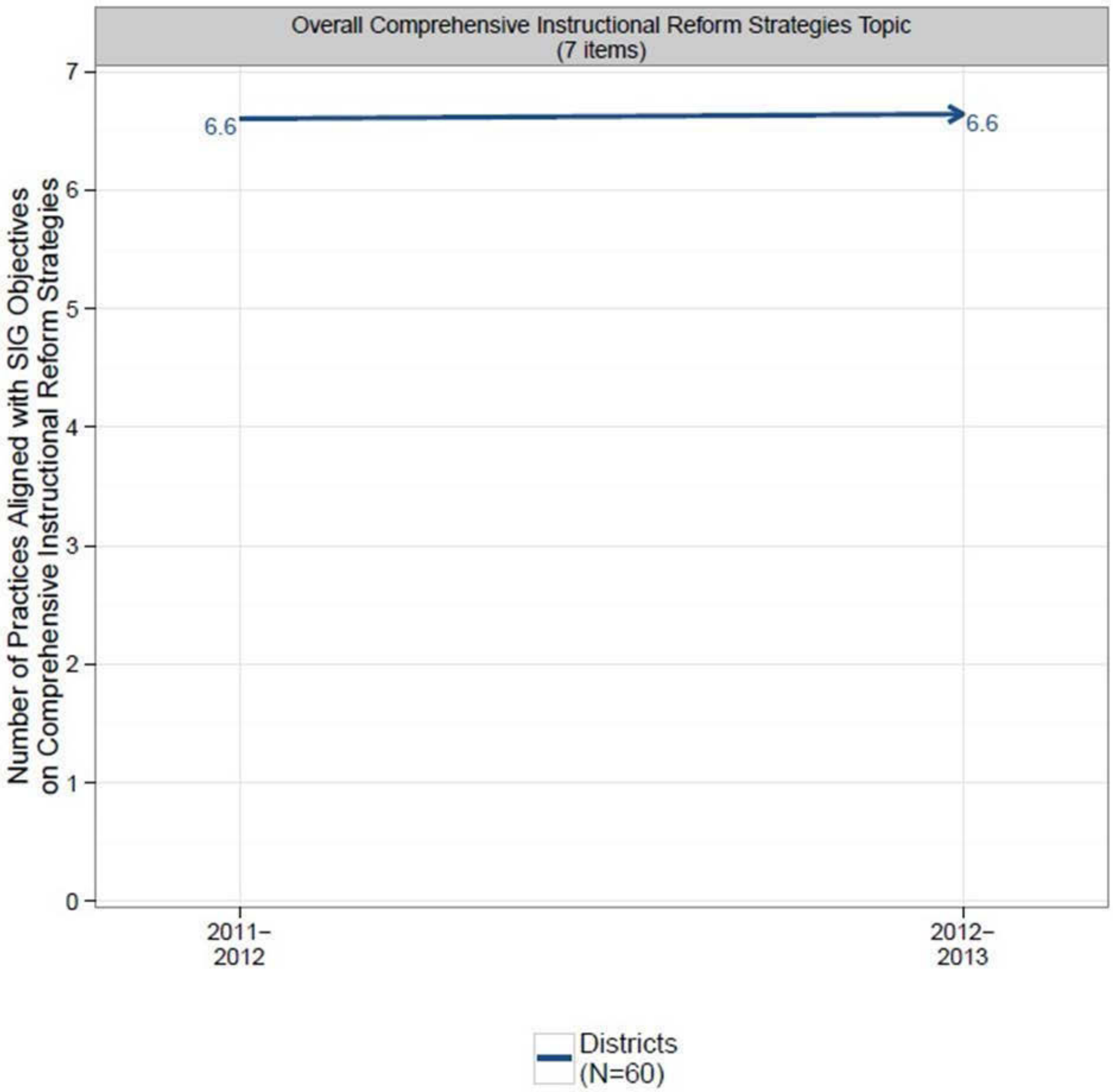
Figure D.1. Use of practices aligned with SIG objectives on comprehensive instructional reform strategies



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.1. Each dot in this figure represents the districts that reported using a particular number of practices (out of seven examined) that were aligned with the SIG application criteria. The number inside each dot is the number of districts represented by the dot; dots that represent fewer than 10 districts have no number inside. For two of the practices, a "yes" response received one point. In the other five cases, it was possible for a district to receive a fraction of one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

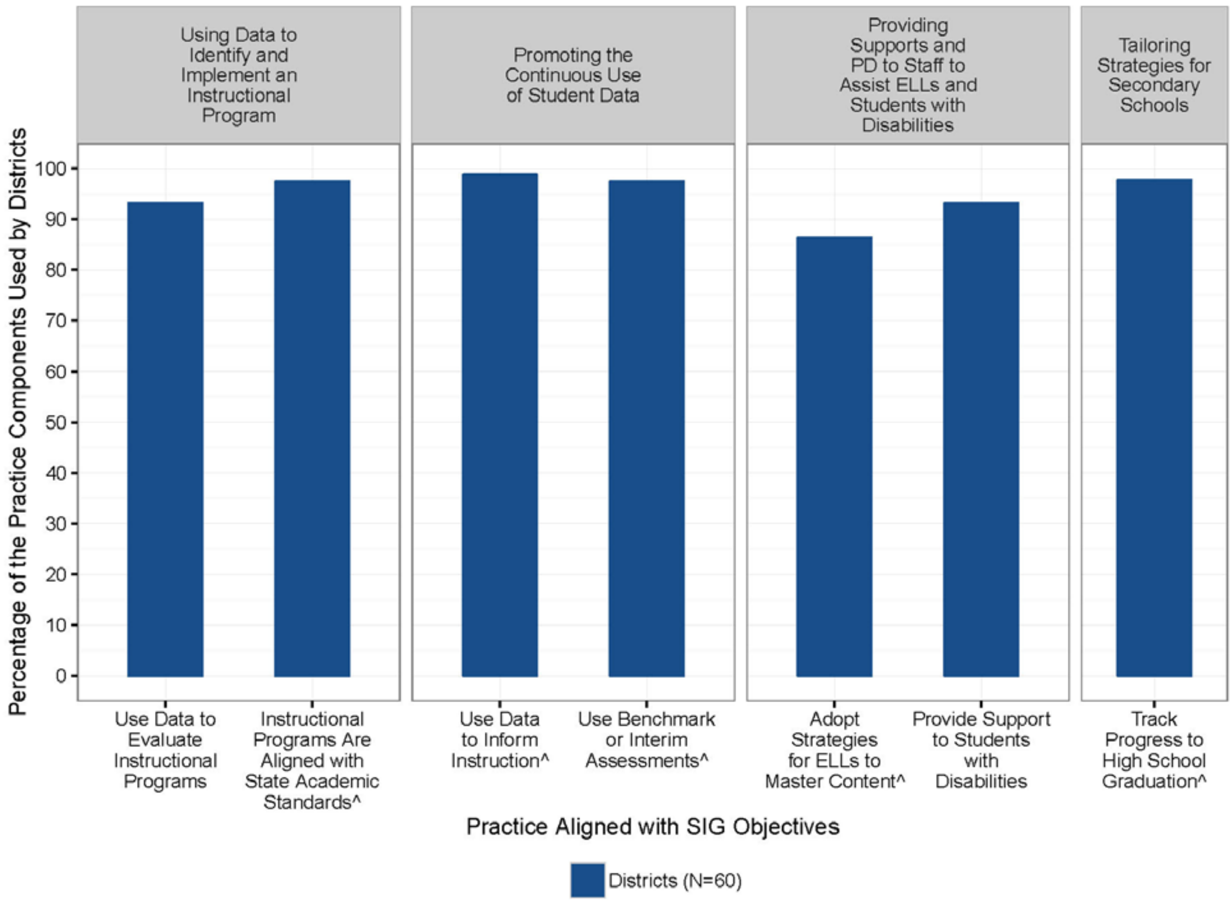
Figure D.2. Change in use of practices aligned with SIG objectives on comprehensive instructional reform strategies



Source: Interviews with district administrators in spring 2012 and spring 2013.

Note: This figure shows change over time for districts in the use of comprehensive instructional reform practices aligned with the SIG application criteria. The arrow starts at the average number of reported practices aligned with the SIG application criteria in spring 2012 and ends at the average number of reported practices aligned with the SIG application criteria in spring 2013.

Figure D.3. Use of individual practices aligned with SIG objectives on comprehensive instructional reform strategies



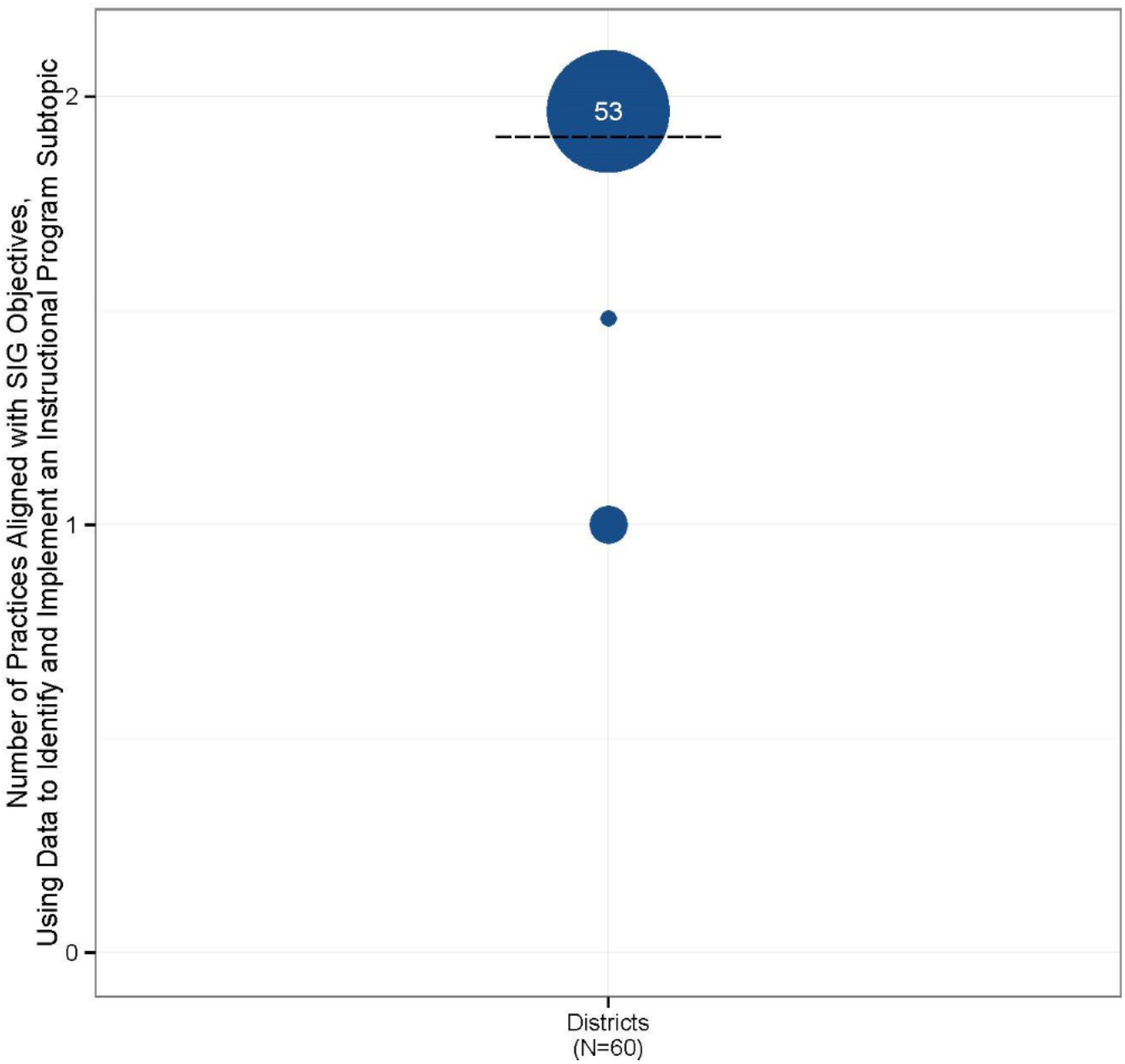
Source: Interviews with district administrators in spring 2013.

Note: This figure has a separate panel for each subtopic. We selected district interview questions that aligned with the practices described in the SIG application criteria. The practices shown on the horizontal axis of this figure are listed in Table D.1. For each practice in the SIG application criteria for which we identified one or more interview questions aligned with the practice, we calculated the percentage of interview questions with a “yes” response as a measure of the percentage of components each district used. The height of each bar represents the average percentage of the components of the practice that each group of districts used.

^Multiple district interview questions were used to assess whether districts used all of the components of this practice.

ELLs = English language learners; PD = professional development.

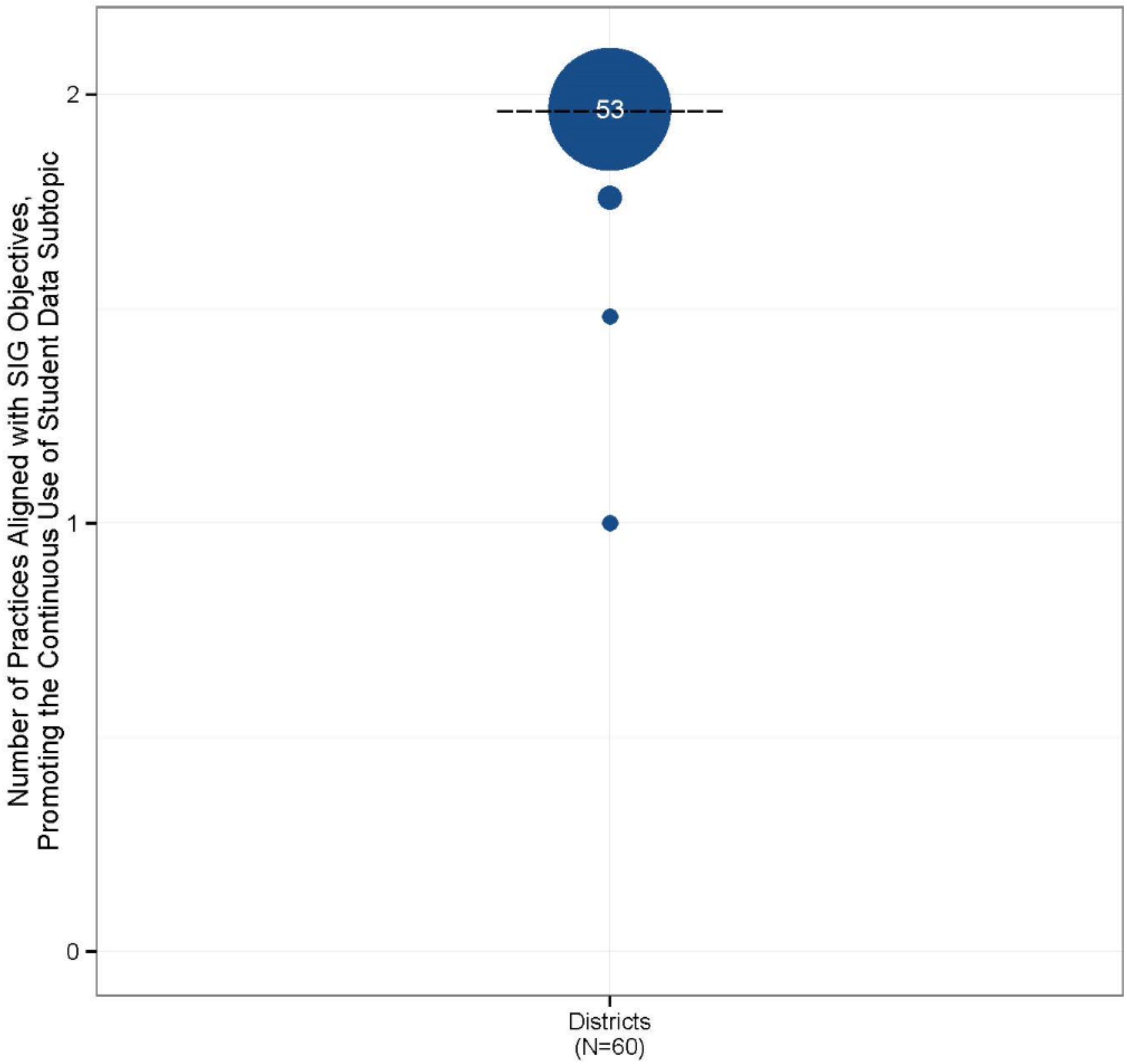
Figure D.4. Use of practices aligned with SIG, using data to identify and implement an instructional program subtopic



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.1. Each dot in this figure represents the districts that reported using a particular number of practices (out of two examined) that were aligned with the SIG application criteria. The number inside each dot is the number of districts represented by the dot; dots that represent fewer than 10 districts have no number inside. For one practice, a "yes" response received one point. For the other, it was possible for a district to receive a fraction of one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

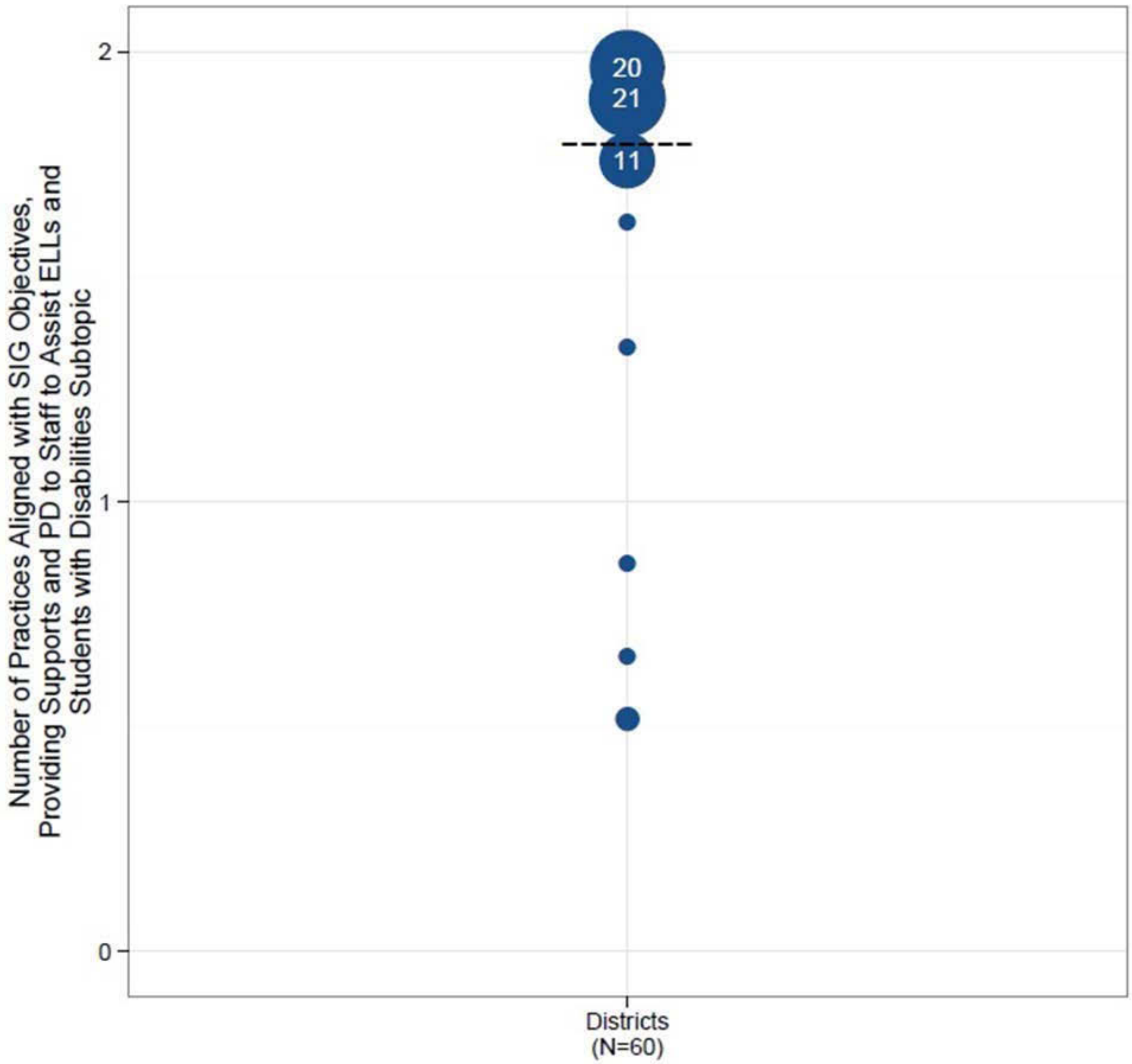
Figure D.5. Use of practices aligned with SIG, promoting the continuous use of student data subtopic



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.1. Each dot in this figure represents the districts that reported using a particular number of practices (out of two examined) that were aligned with the SIG application criteria. The number inside each dot is the number of districts represented by the dot; dots that represent fewer than 10 districts have no number inside. For both strategies, it was possible for a district to receive a fraction of one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

Figure D.6. Use of practices aligned with SIG, providing supports and professional development to staff to assist English language learners and students with disabilities subtopic

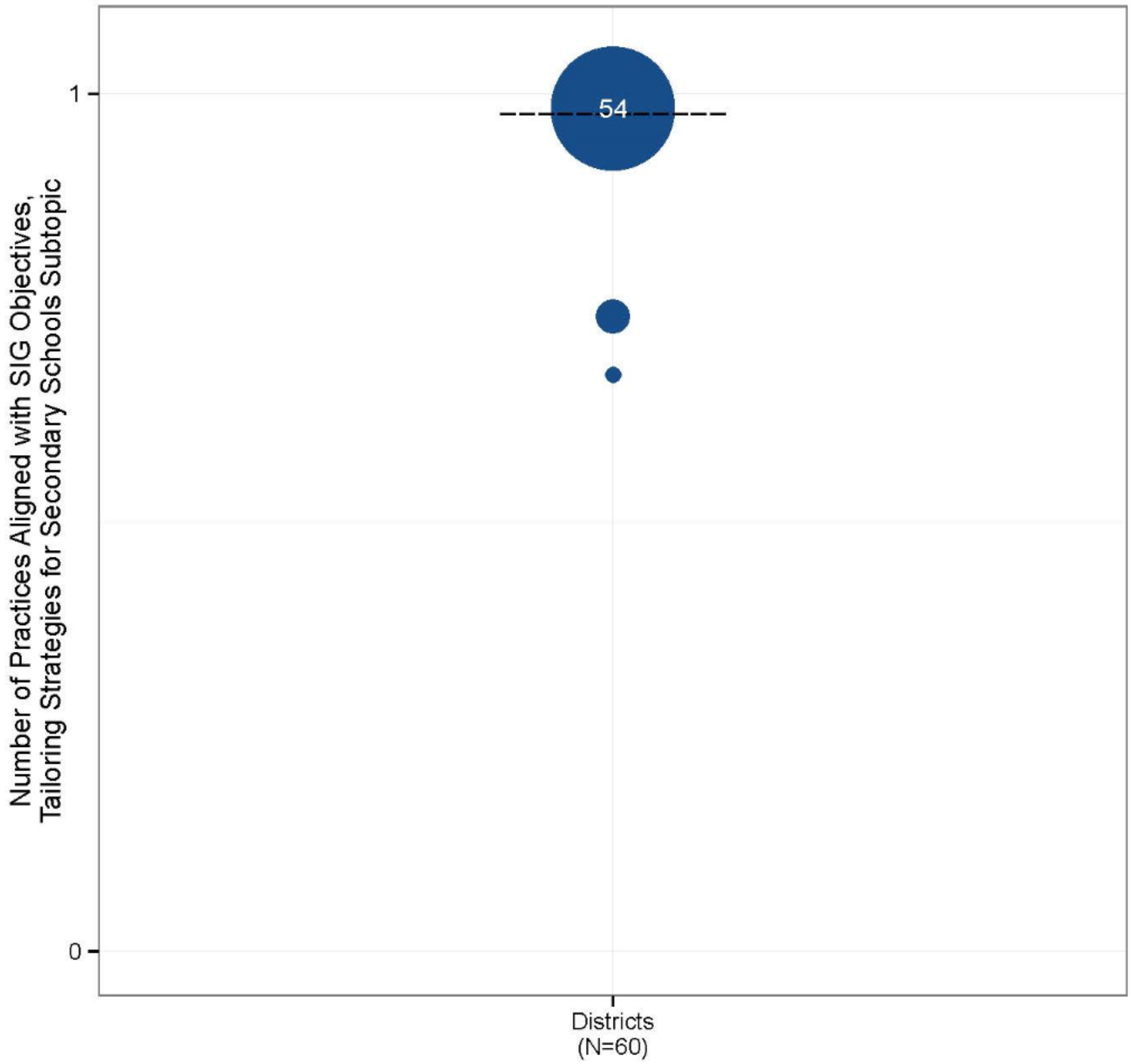


Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.1. Each dot in this figure represents the districts that reported using a particular number of practices (out of two examined) that were aligned with the SIG application criteria. The number inside each dot is the number of districts represented by the dot; dots that represent fewer than 10 districts have no number inside. For one strategy, a “yes” response received one point. For the other, it was possible for a district to receive a fraction of one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

ELLs = English language learners; PD = professional development.

Figure D.7. Use of practices aligned with SIG, tailoring strategies for secondary schools subtopic



Source: Interviews with district administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table D.1. This figure presents one practice described in the SIG application criteria to which multiple interview questions aligned. As described in Appendix C, whenever multiple interview questions aligned with a single practice from the application criteria, we used those questions to construct a variable ranging from zero to one, with districts receiving a fraction of a point for each question to which they responded "yes." Each dot in this figure represents the districts that reported using a particular proportion of the interview questions aligned to the practice described in the SIG application criteria. The number inside each dot is the number of districts represented by the dot; dots that represent fewer than 10 districts have no number inside. The dashed line denotes the average value across all districts.

B. Teacher and principal effectiveness

The spring 2013 district interview asked about 10 practices aligned with SIG objectives on teacher and principal effectiveness (Table D.2).

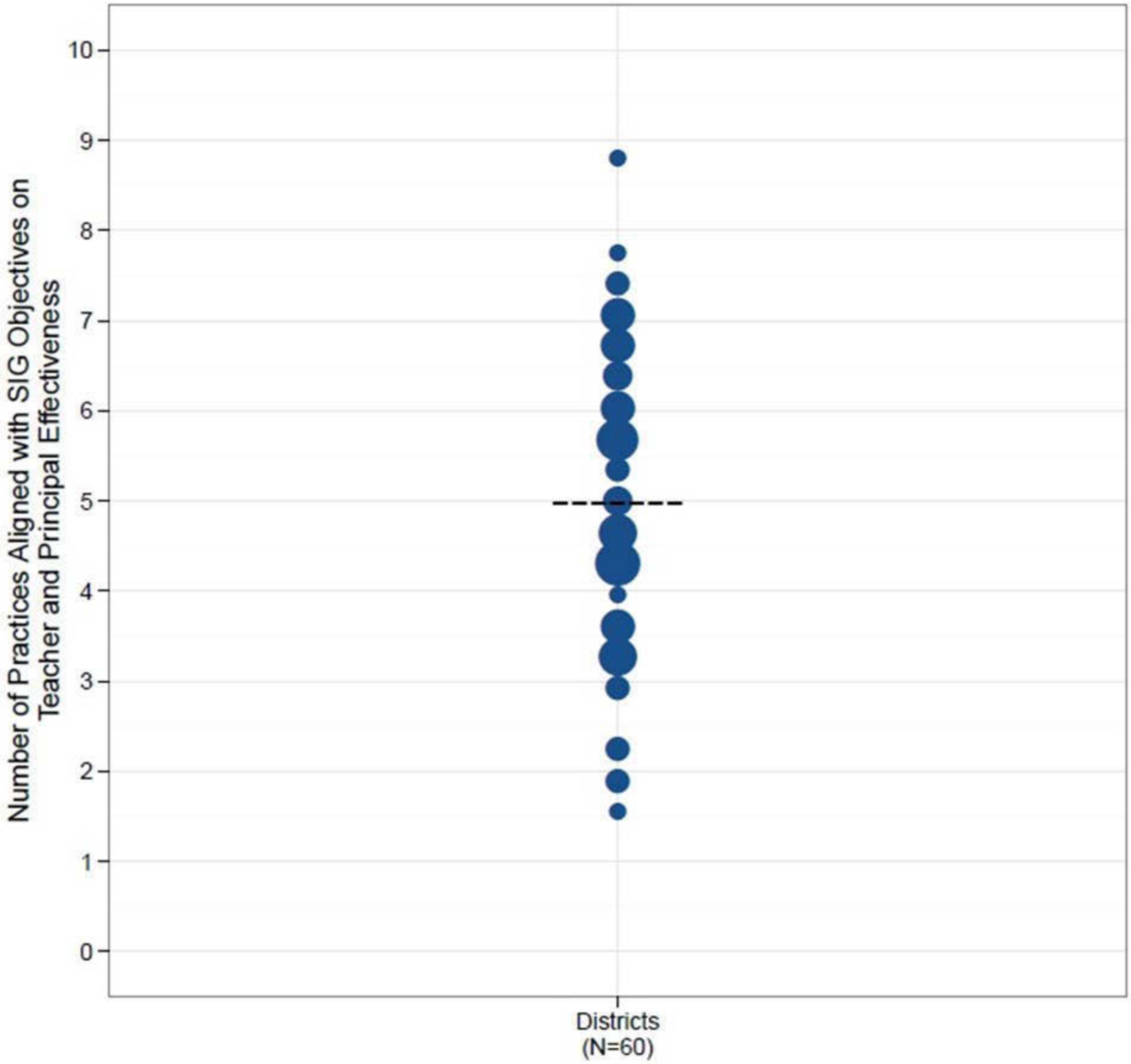
Table D.2. Practices aligned with SIG objectives on teacher and principal effectiveness, by subtopic

Teacher effectiveness
Using rigorous, transparent, and equitable evaluation systems
Requiring schools to use student achievement growth to evaluate teachers, specifying the extent to which student achievement growth must factor into teacher evaluations, or using state test scores to assess student growth for teacher evaluations
Requiring multiple performance measures for teacher evaluations
Identifying and rewarding effective teachers and removing ineffective ones
Using data to inform decisions such as tenure, retention, and bonuses for teachers
Providing high-quality, job embedded professional development or supports
Using data to inform professional development offerings for teachers
Using data to evaluate the success of professional development offerings for teachers
Implementing strategies to recruit, place, and retain staff
Implementing strategies, such as financial incentives or induction support for novice teachers, designed to recruit, place, and retain teachers in SIG schools
Modifying teacher tenure rules that affect placement in or removal from SIG schools or permitting principal discretion in hiring teachers for SIG schools
Principal effectiveness
Using rigorous, transparent, and equitable evaluation systems
Requiring schools to use student achievement growth to evaluate principals or using state test scores to assess student achievement growth for principal evaluations
Requiring multiple performance measures other than student growth for principal evaluations
Implementing strategies to recruit, place, and retain staff
Implementing strategies, such as financial incentives, that are designed to recruit, place, and retain principals in SIG schools

Source: SIG application; interviews with district administrators in spring 2013.

Figure D.8 displays results of the analysis on the extent to which district administrators reported using the teacher and principal effectiveness practices aligned with the SIG application criteria. Figure D.9 displays the change over time in districts' reported use of the practices included in the analysis for this area. Figures D.10 and D.11 display the extent to which districts reported using the individual teacher and principal effectiveness practices included in the analysis. Figures D.12–D.15 display the results for each subtopic.

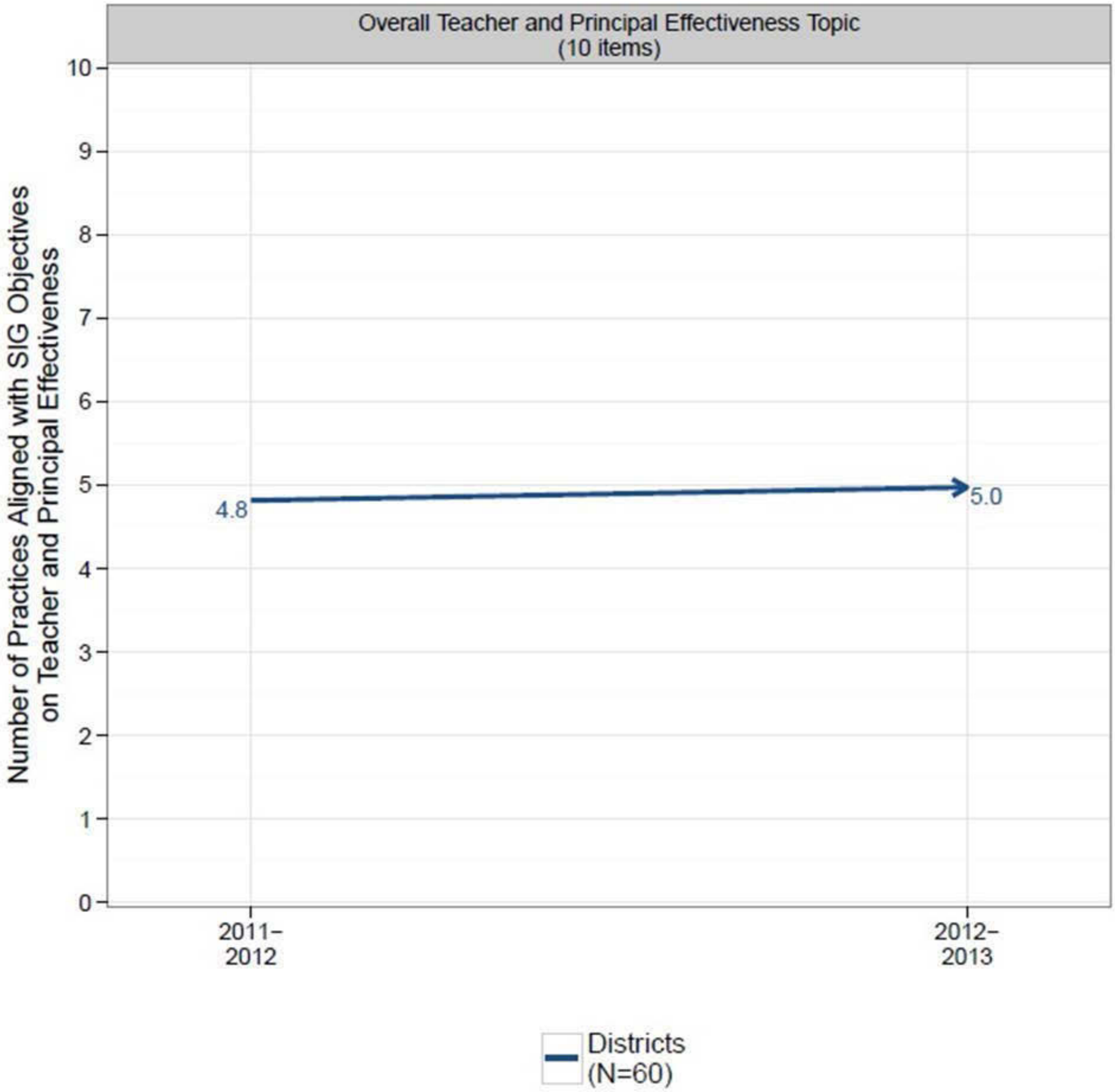
Figure D.8. Use of practices aligned with SIG objectives on teacher and principal effectiveness



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.2. Each dot in this figure represents the districts that reported using a particular number of practices (out of 10 examined) that were aligned with the SIG application criteria. Each dot in this figure represents fewer than 10 districts, so the numbers inside the dots have been removed to protect respondent confidentiality. For three of the practices, a “yes” response received one point. For the other seven practices, it was possible for a district to receive a fraction of one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

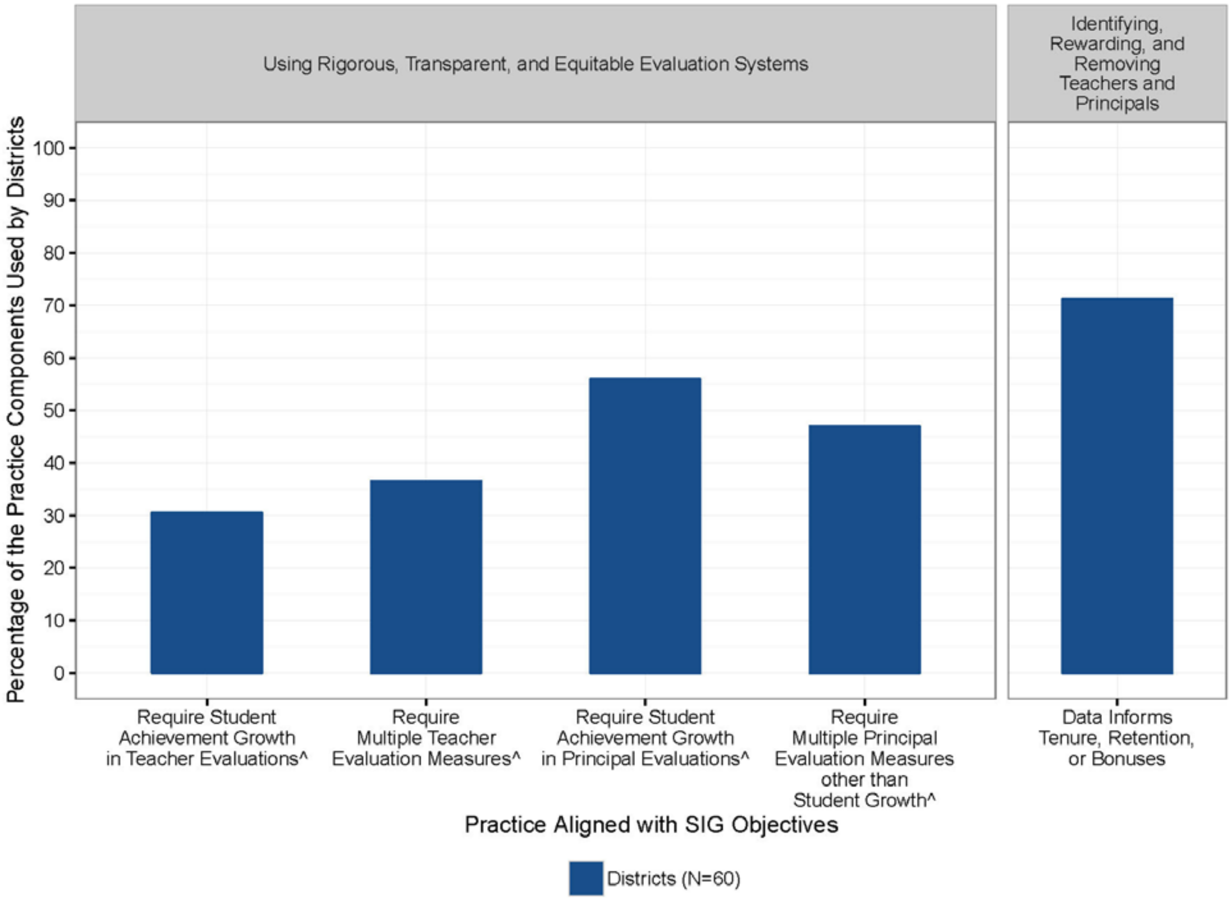
Figure D.9. Change in use of practices aligned with SIG objectives on teacher and principal effectiveness



Source: Interviews with district administrators in spring 2012 and spring 2013.

Note: This figure shows change over time for districts in the use of teacher and principal effectiveness practices aligned with the SIG application criteria. The arrow starts at the average number of reported practices aligned with the SIG application criteria in spring 2012 and ends at the average number of reported practices aligned with the SIG application criteria in spring 2013.

Figure D.10. Use of individual practices aligned with SIG; using rigorous, transparent, and equitable evaluation systems subtopic and identifying and rewarding or removing teachers and principals subtopic

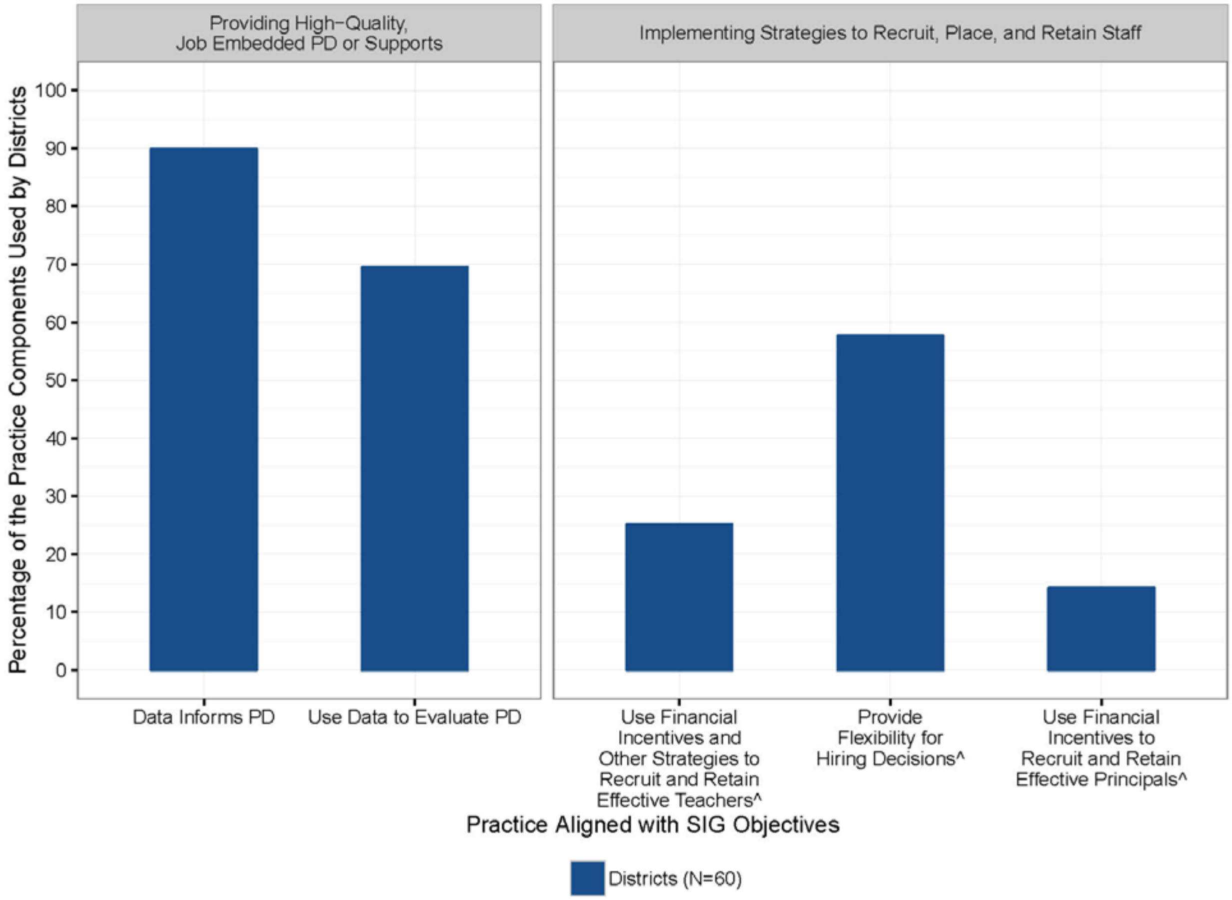


Source: Interviews with district administrators in spring 2013.

Note: This figure has a separate panel for each subtopic. We selected district interview questions that aligned with the practices described in the SIG application criteria. The practices shown on the horizontal axis of this figure are listed in Table D.2. For each practice in the SIG application criteria for which we identified one or more interview questions aligned with the practice, we calculated the percentage of interview questions with a “yes” response as a measure of the percentage of components each district used. The height of each bar represents the average percentage of the components of the practice that each group of districts used.

^Multiple district interview questions were used to assess whether districts used all of the components of this practice.

Figure D.11. Use of individual practices aligned with SIG; providing high-quality job-embedded professional development or supports subtopic and implementing strategies to recruit, place, and retain staff subtopic



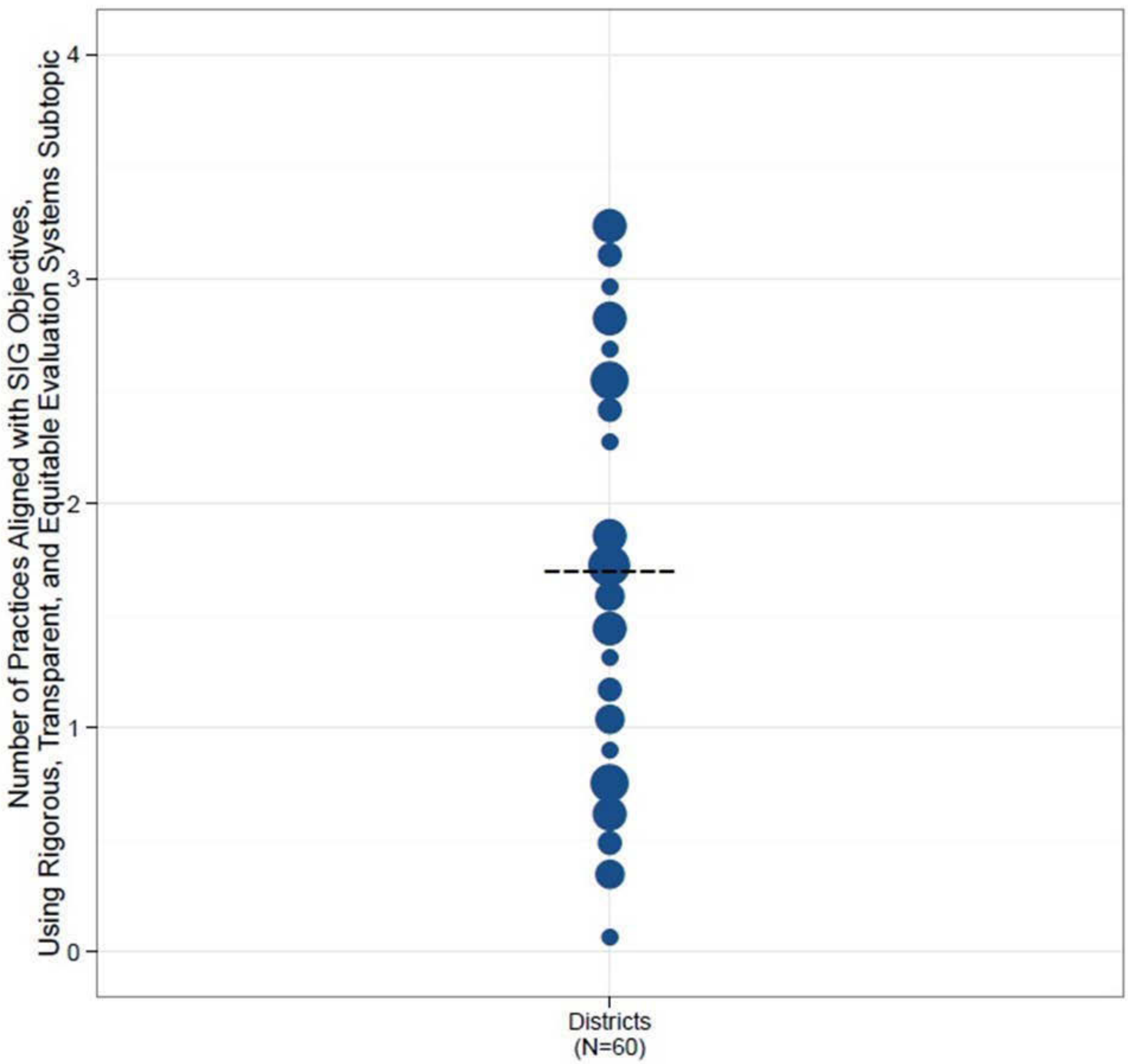
Source: Interviews with district administrators in spring 2013.

Note: This figure has a separate panel for each subtopic. We selected district interview questions that aligned with the practices described in the SIG application criteria. The practices shown on the horizontal axis of this figure are listed in Table D.2. For each practice in the SIG application criteria for which we identified one or more interview questions aligned with the practice, we calculated the percentage of interview questions with a “yes” response as a measure of the percentage of components each district used. The height of each bar represents the average percentage of the components of the practice that each group of districts used.

^Multiple district interview questions were used to assess whether districts used all of the components of this practice.

PD = professional development.

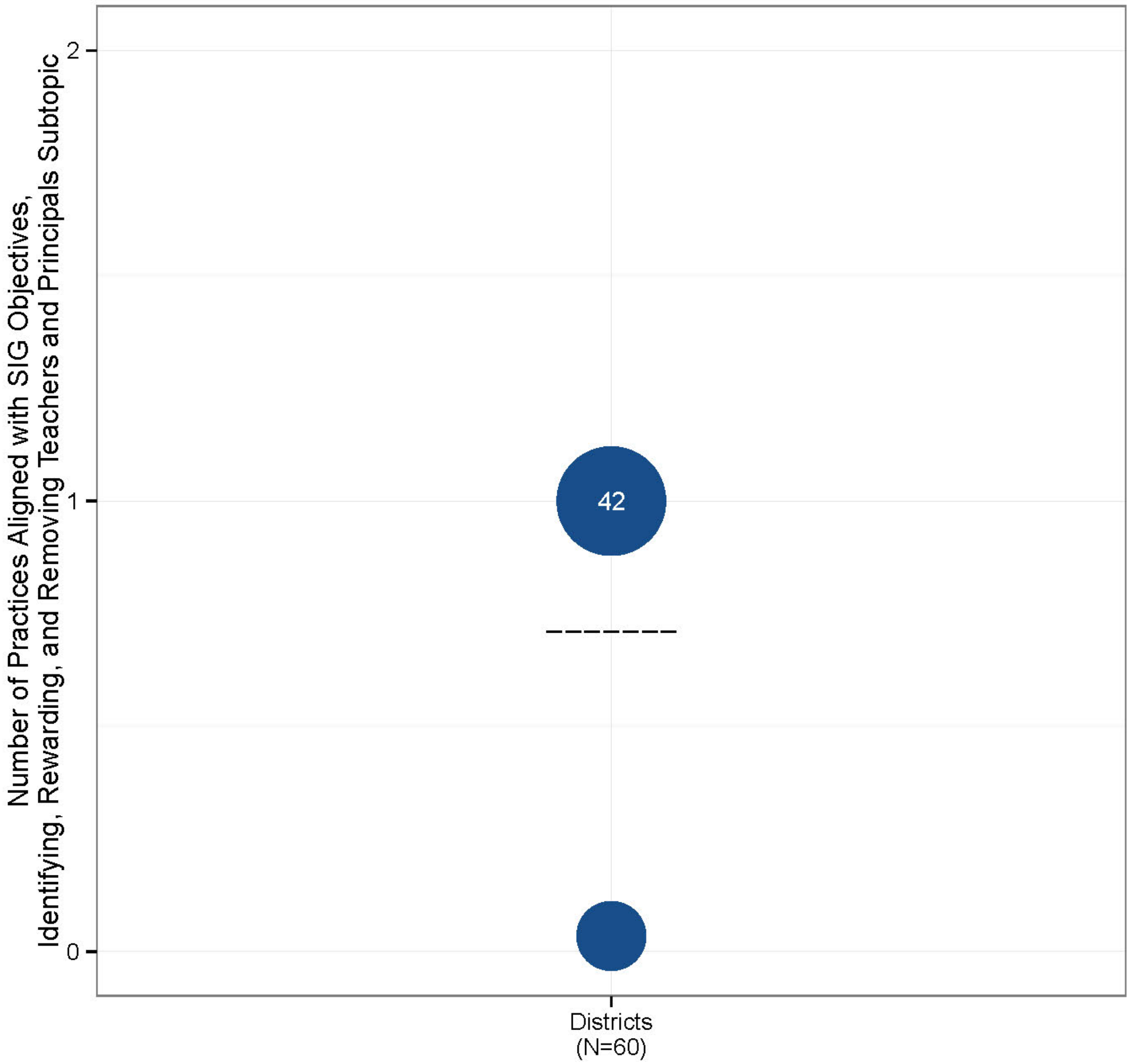
Figure D.12. Use of practices aligned with SIG, using rigorous, transparent, and equitable evaluation systems subtopic



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.2. Each dot in this figure represents the districts that reported using a particular number of practices (out of four examined) that were aligned with the SIG application criteria. Each dot in this figure represents fewer than 10 districts, so the numbers inside the dots have been removed to protect respondent confidentiality. For all four practices, it was possible for a district to receive a fraction of one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

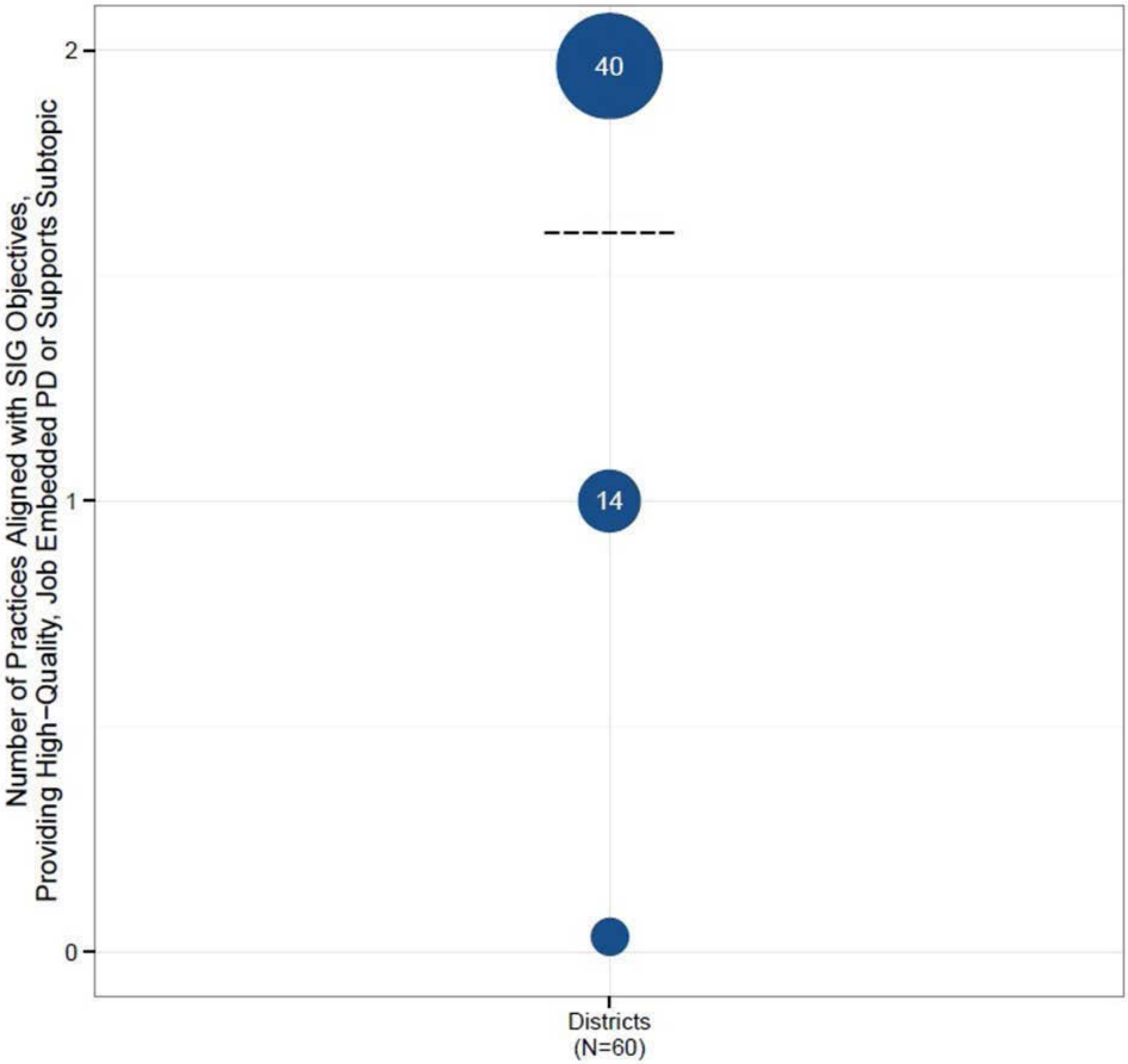
Figure D.13. Use of practices aligned with SIG, identifying and rewarding effective teachers and principals and removing ineffective ones subtopic



Source: Interviews with district administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table D.2. Each dot in this figure represents the districts that reported using the one practice that was aligned with the SIG application criteria. The number inside each dot is the number of districts represented by the dot. To protect respondent confidentiality, the number inside the smallest dot has been removed. For this practice, a “yes” response received one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

Figure D.14. Use of practices aligned with SIG, providing high quality, job-embedded professional development or supports subtopic

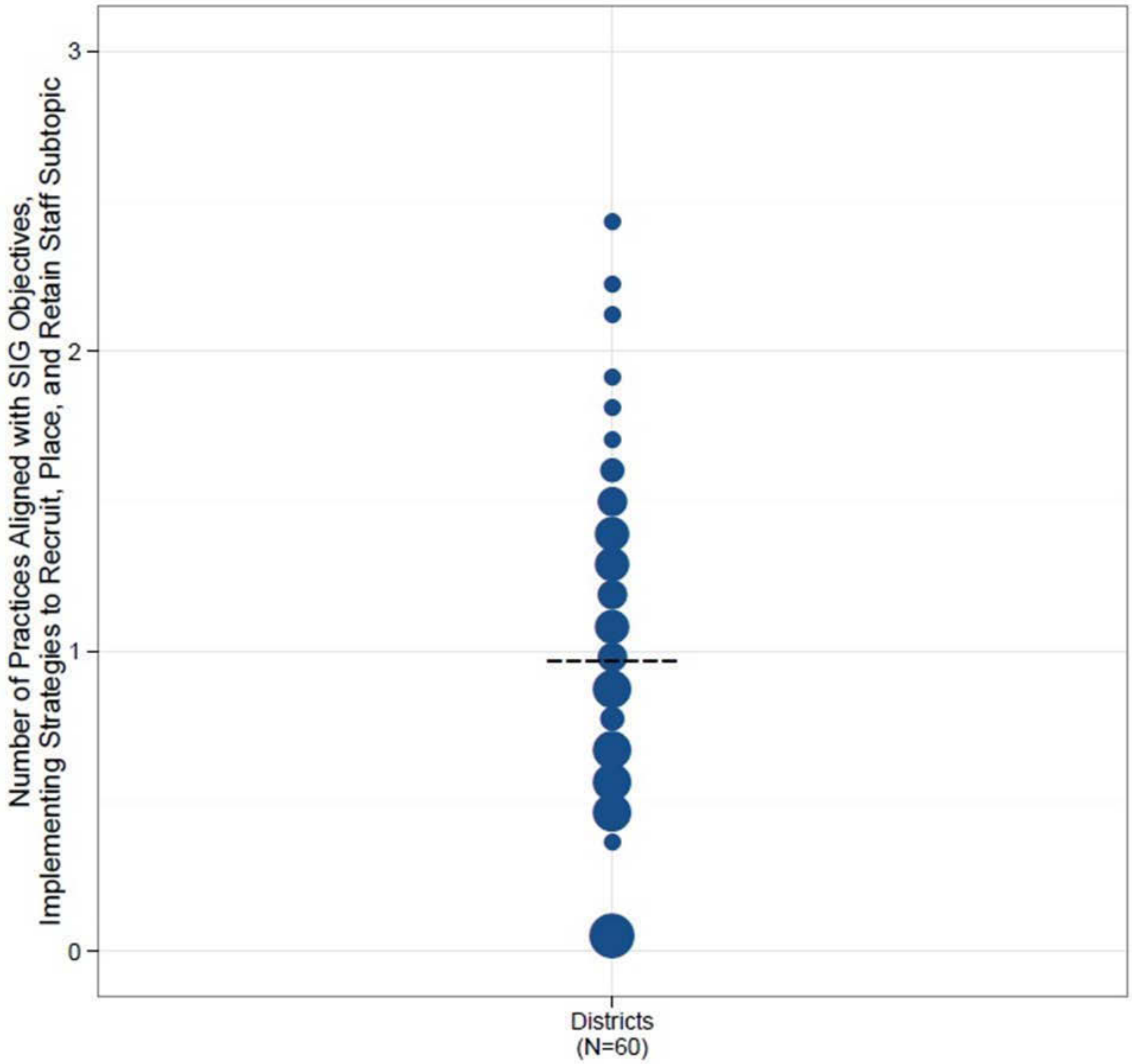


Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.2. Each dot in this figure represents the districts that reported using a particular number of practices (out of two examined) that were aligned with the SIG application criteria. The number inside each dot is the number of districts represented by the dot. To protect respondent confidentiality, the number inside the smallest dot has been removed. For both practices, a “yes” response received one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

PD = professional development.

Figure D.15. Use of practices aligned with SIG, implementing strategies to recruit, place, and retain staff subtopic



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.2. Each dot in this figure represents the districts that reported using a particular number of practices (out of three examined) that were aligned with the SIG application criteria. Each dot in this figure represents fewer than 10 districts, so the numbers inside the dots have been removed to protect respondent confidentiality. For all three practices, it was possible for a district to receive a fraction of one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

C. Learning time and community-oriented schools

The spring 2013 district interview asked about two practices aligned with SIG objectives on learning time and community-oriented schools (Table D.3).

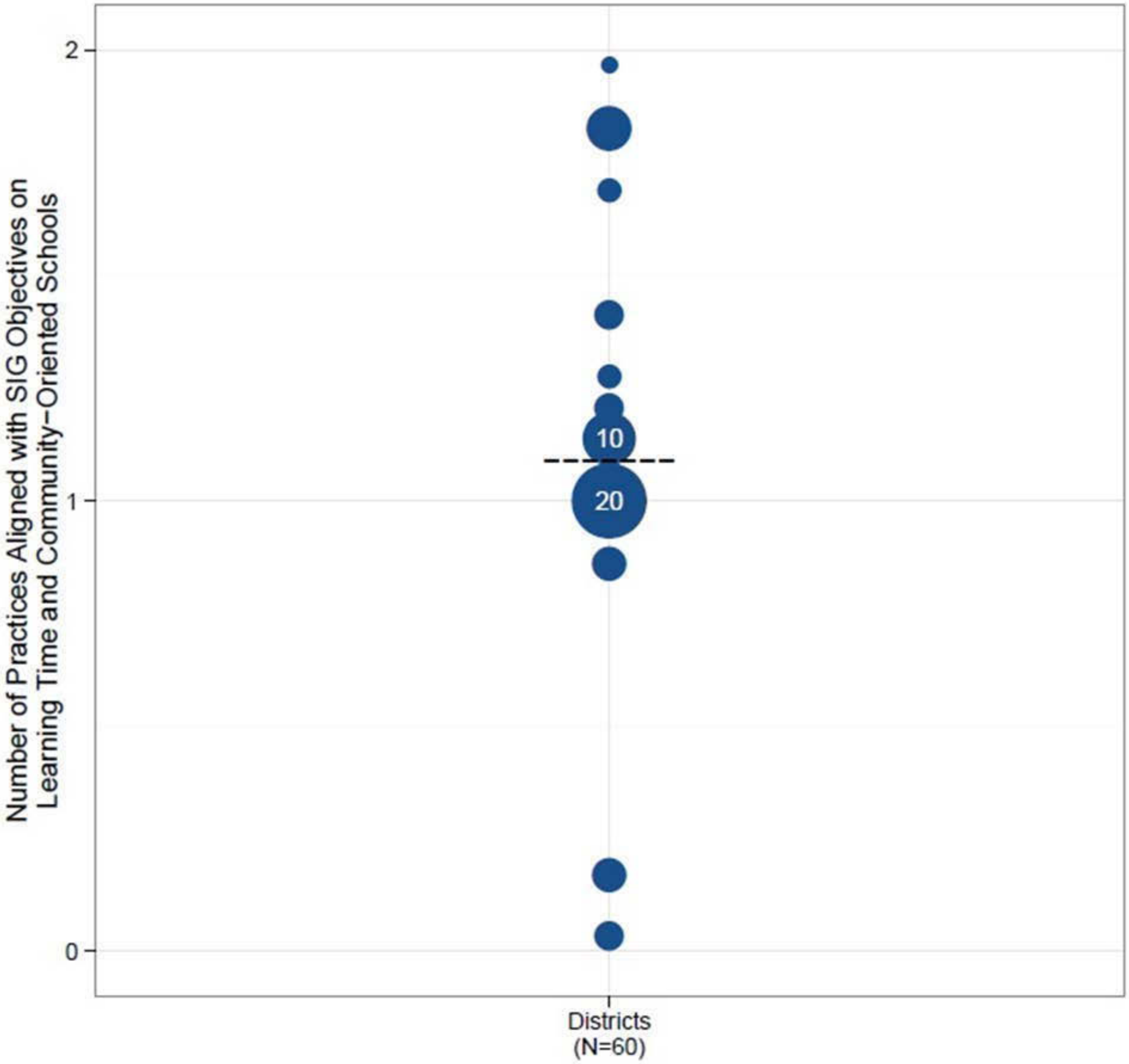
Table D.3. Practices aligned with SIG objectives on learning time and community-oriented schools, by subtopic

Increasing learning time
Increasing the minimum amount of time spent each week on English language arts or math instruction or increasing the number of instructional days in the school year
Engaging families and communities
Using data to guide the development and implementation of nonacademic supports or enrichment programs, for example, to identify how many and which students need counseling

Source: SIG application; interviews with district administrators in spring 2013.

Figure D.16 displays results of the analysis on the extent to which district administrators reported using the increasing learning time and creating community-oriented schools practices aligned with the SIG application criteria. Figure D.17 displays the change over time in districts' reported use of the practices included in the analysis for this area. Figure D.18 displays the extent to which districts reported using the individual practices included in the analysis for this area. Figures D.19 and D.20 display the results for each subtopic.

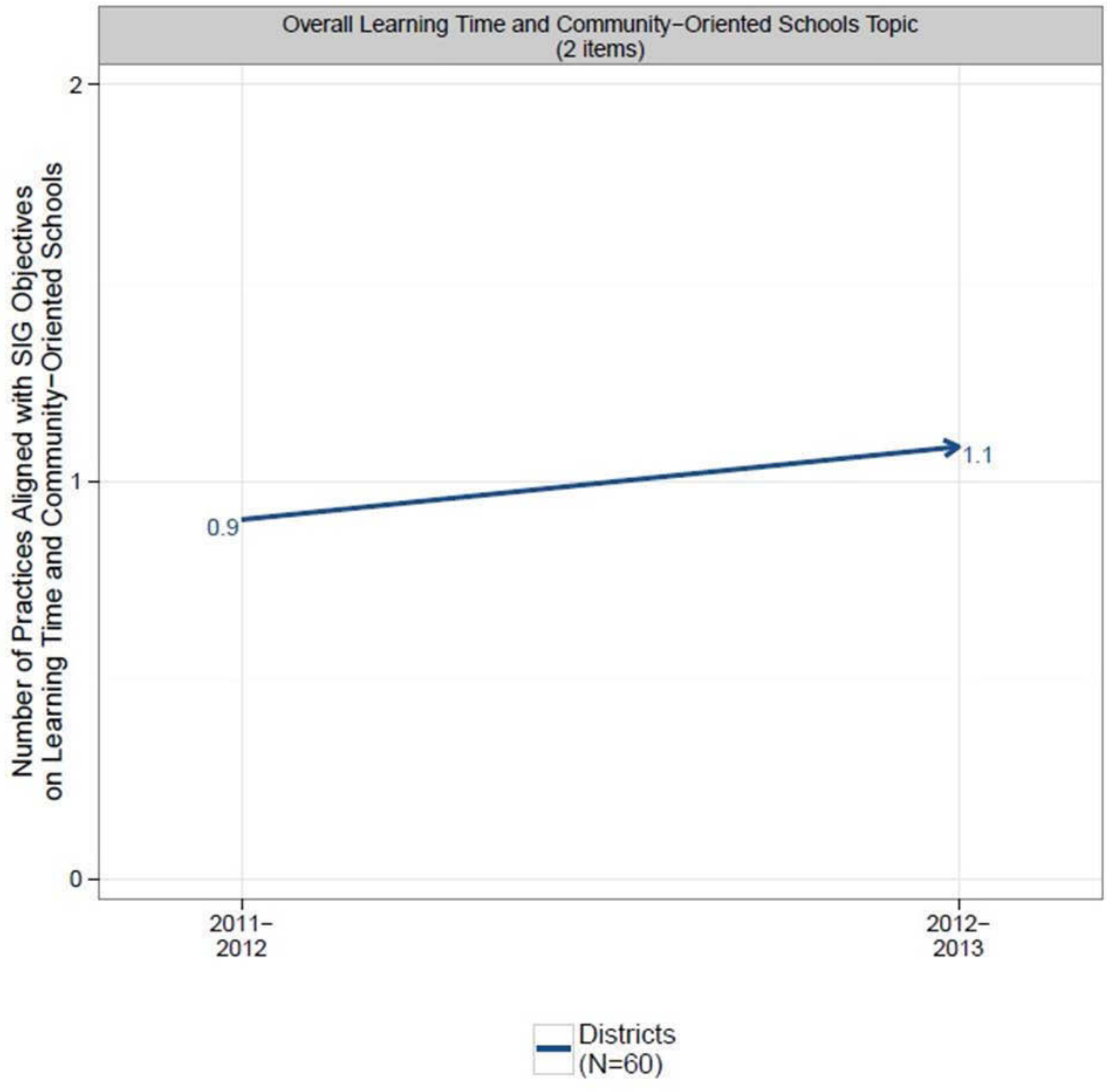
Figure D.16. Use of practices aligned with SIG objectives on learning time and community-oriented schools



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.3. Each dot in this figure represents the districts that reported using a particular number of practices (out of two examined) that were aligned with the SIG application criteria. The number inside each dot is the number of districts represented by the dot; dots that represent fewer than 10 districts have no number inside. For one practice, a "yes" response received one point. For the other practice, it was possible for a district to receive a fraction of one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

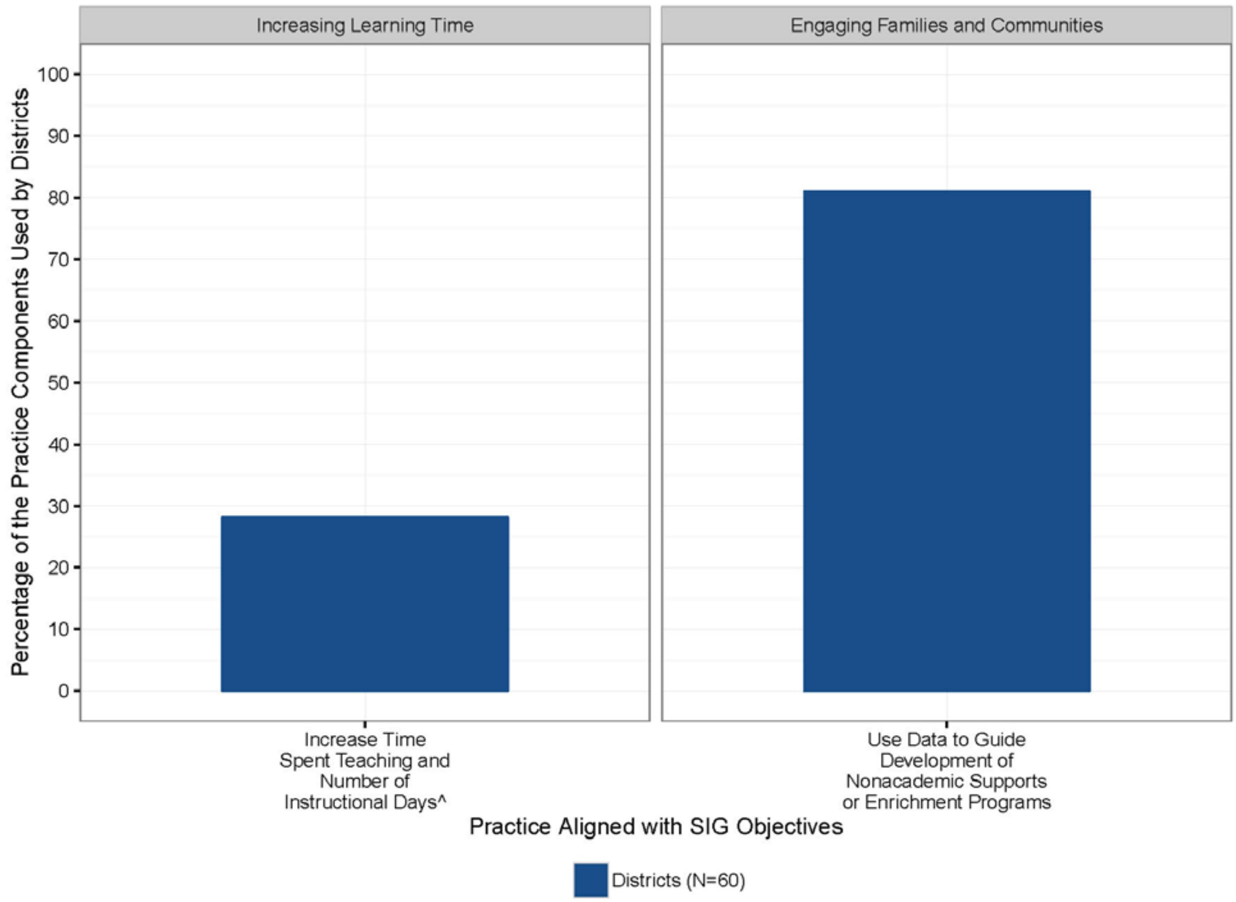
Figure D.17. Change in use of practices aligned with SIG objectives on learning time and community-oriented schools



Source: Interviews with district administrators in spring 2012 and spring 2013.

Note: This figure shows change over time for districts in the use of learning time and community-oriented schools practices aligned with the SIG application criteria. The arrow starts at the average number of reported practices aligned with the SIG application criteria in spring 2012 and ends at the average number of reported practices aligned with the SIG application criteria in spring 2013.

Figure D.18. Use of individual practices aligned with SIG objectives on learning time and community-oriented schools

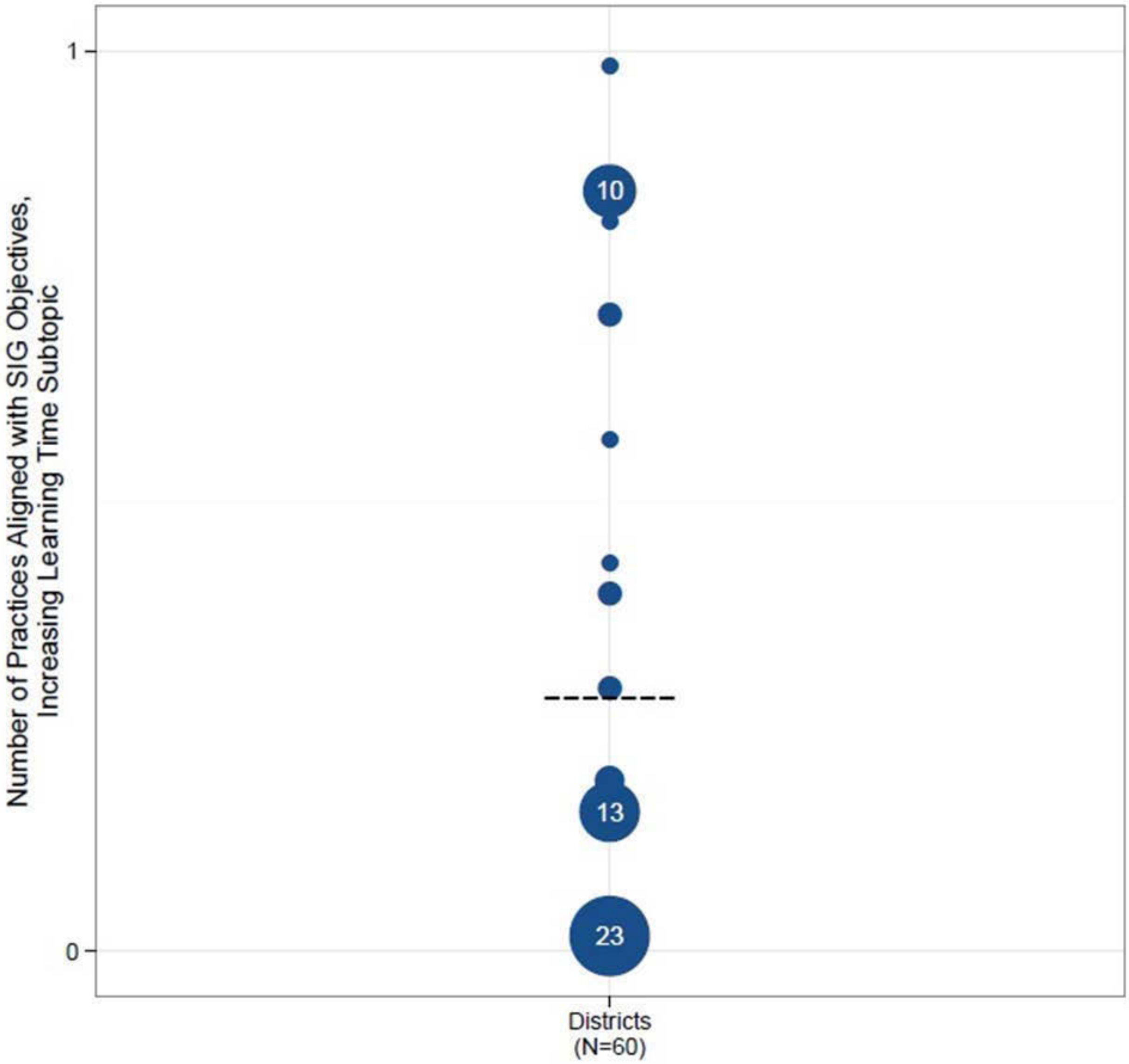


Source: Interviews with district administrators in spring 2013.

Note: This figure has a separate panel for each subtopic. We selected district interview questions that aligned with the practices described in the SIG application criteria. The practices shown on the horizontal axis of this figure are listed in Table D.3. For each practice in the SIG application criteria for which we identified one or more interview questions aligned with the practice, we calculated the percentage of interview questions with a “yes” response as a measure of the percentage of components each district used. The height of each bar represents the average percentage of the components of the practice that each group of districts used.

^Multiple district interview questions were used to assess whether districts used all of the components of this practice.

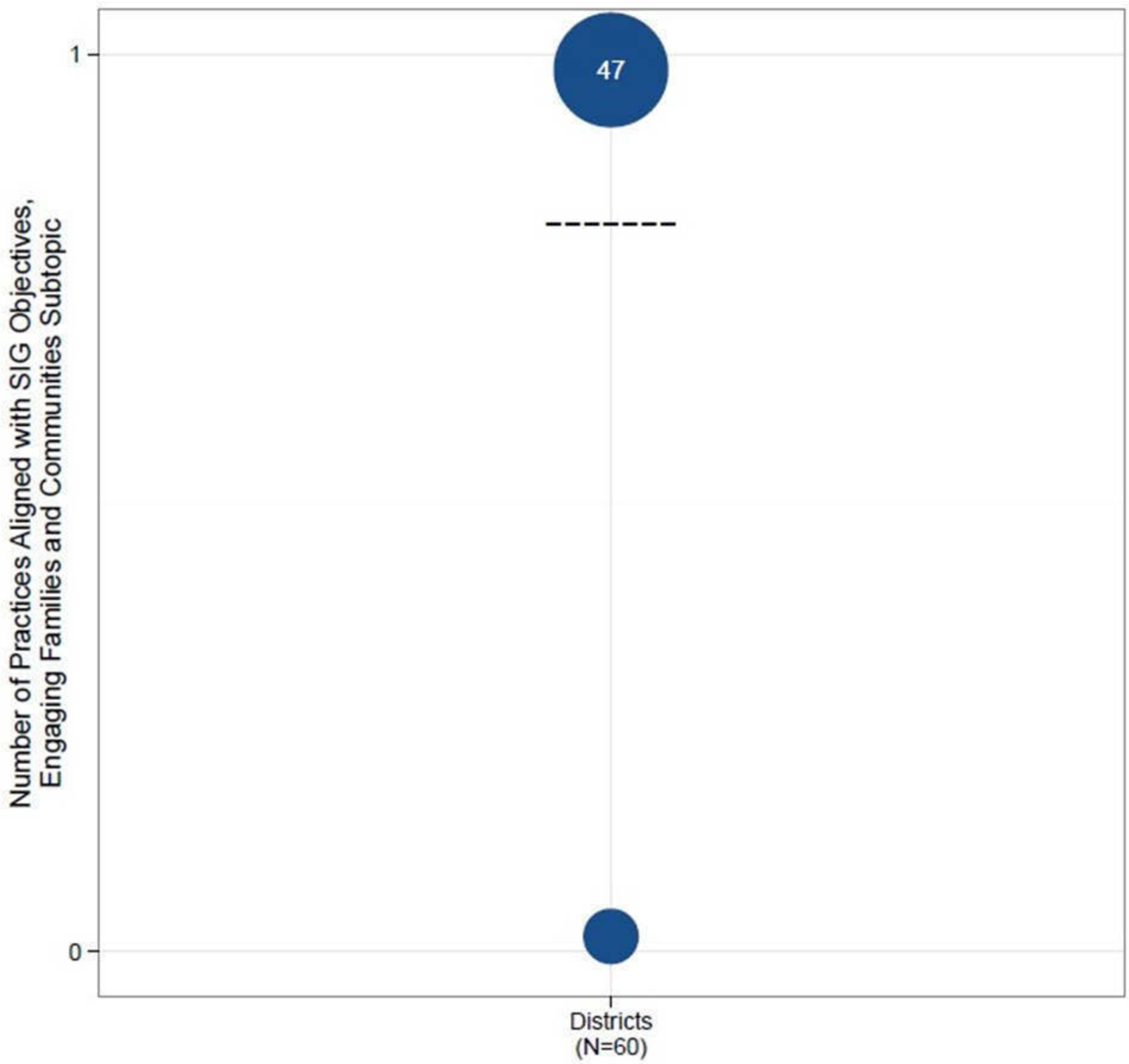
Figure D.19. Use of practices aligned with SIG, increasing learning time subtopic



Source: Interviews with district administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table D.3. This figure presents one practice described in the SIG application criteria to which multiple interview questions aligned. As described in Appendix C, whenever multiple interview questions aligned with a single practice from the application criteria, we used those questions to construct a variable ranging from zero to one, with districts receiving a fraction of a point for each question to which they responded "yes." Each dot in this figure represents the districts that reported using a particular proportion of the interview questions aligned to the practice described in the SIG application criteria. The number inside each dot is the number of districts represented by the dot; dots that represent fewer than 10 districts have no number inside. The dashed line denotes the average value across all districts.

Figure D.20. Use of practices aligned with SIG, engaging families and communities subtopic



Source: Interviews with district administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table D.3. Each dot in this figure represents the districts that reported using the one practice that was aligned with the SIG application criteria. The number inside each dot is the number of districts represented by the dot. To protect respondent confidentiality, the number inside the smallest dot has been removed. For this practice, a "yes" response received one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

D. Operational flexibility and support

The spring 2013 district interview asked about three practices aligned with SIG objectives on operational flexibility and support (Table D.4).

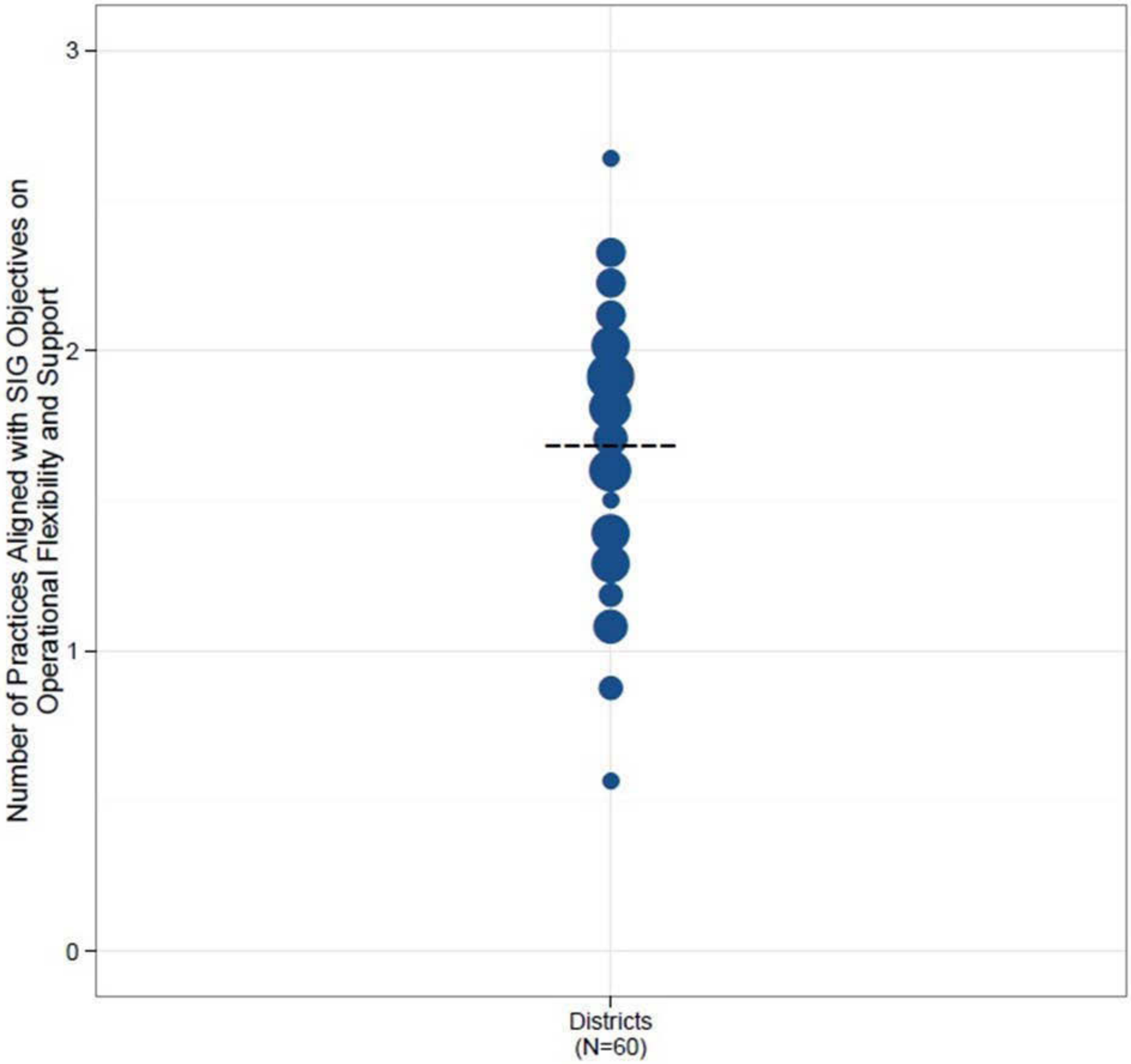
Table D.4. Practices aligned with SIG objectives on operational flexibility and support, by subtopic

Providing operational flexibility
Low-performing schools had primary responsibility for budget, hiring, discipline, or school year length decisions
Receiving technical assistance and support
Receiving training, technical assistance, or access to data from the state to support school improvement efforts or use data to improve instruction
Having a designated office or staff or contracting with external consultants to support school turnaround efforts

Source: SIG application; interviews with district administrators in spring 2013.

Figure D.21 displays results of the analysis on the extent to which district administrators reported using the operational flexibility and support practices aligned with the SIG application criteria. Figure D.22 displays the change over time in districts' reported use of the practices included in the analysis for this area. Figure D.23 displays the extent to which districts reported using the individual operational flexibility and support practices included in the analysis. Figures D.24 and D.25 display the results for each subtopic.

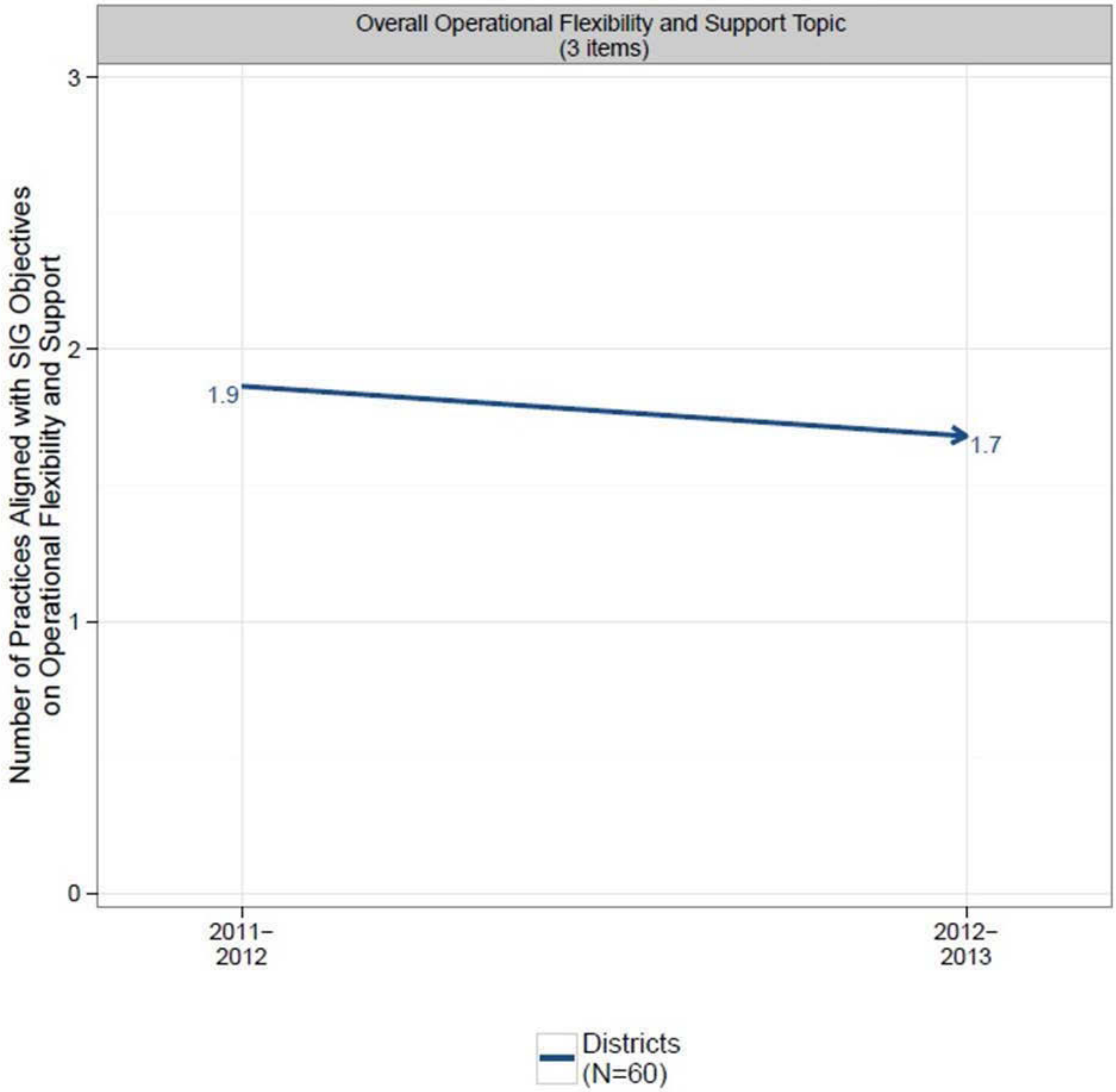
Figure D.21. Use of practices aligned with SIG objectives on operational flexibility and support



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.4. Each dot in this figure represents the districts that reported using a particular number of practices (out of three examined) that were aligned with the SIG application criteria. Each dot in this figure represents fewer than 10 districts, so the numbers inside the dots have been removed to protect respondent confidentiality. For all three practices, it was possible for a district to receive a fraction of one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

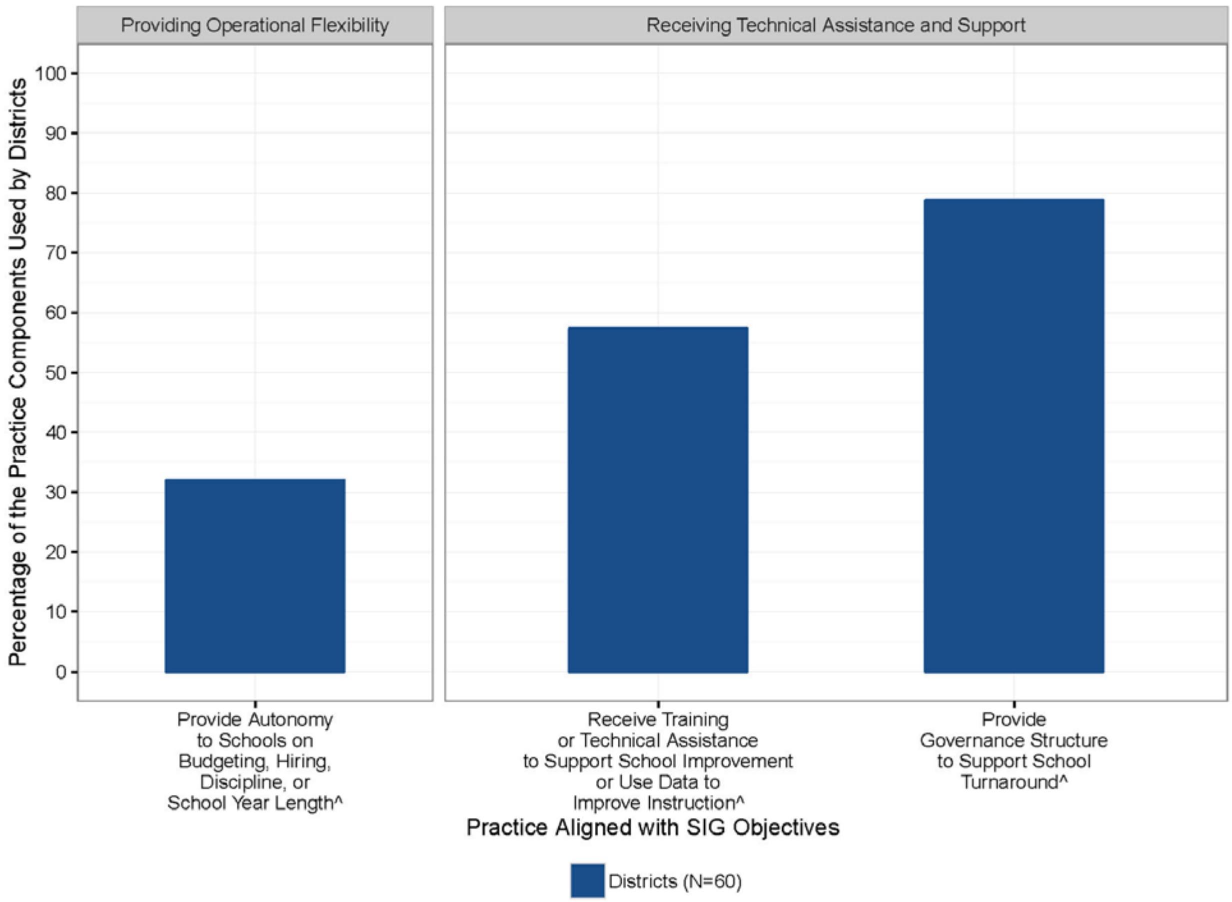
Figure D.22. Change in use of practices aligned with SIG objectives on operational flexibility and support



Source: Interviews with district administrators in spring 2012 and spring 2013.

Note: This figure shows change over time for districts in the use of operational flexibility and support aligned with the SIG application criteria. The arrow starts at the average number of reported practices aligned with the SIG application criteria in spring 2012 and ends at the average number of reported practices aligned with the SIG application criteria in spring 2013.

Figure D.23. Use of individual practices aligned with SIG objectives on operational flexibility and support

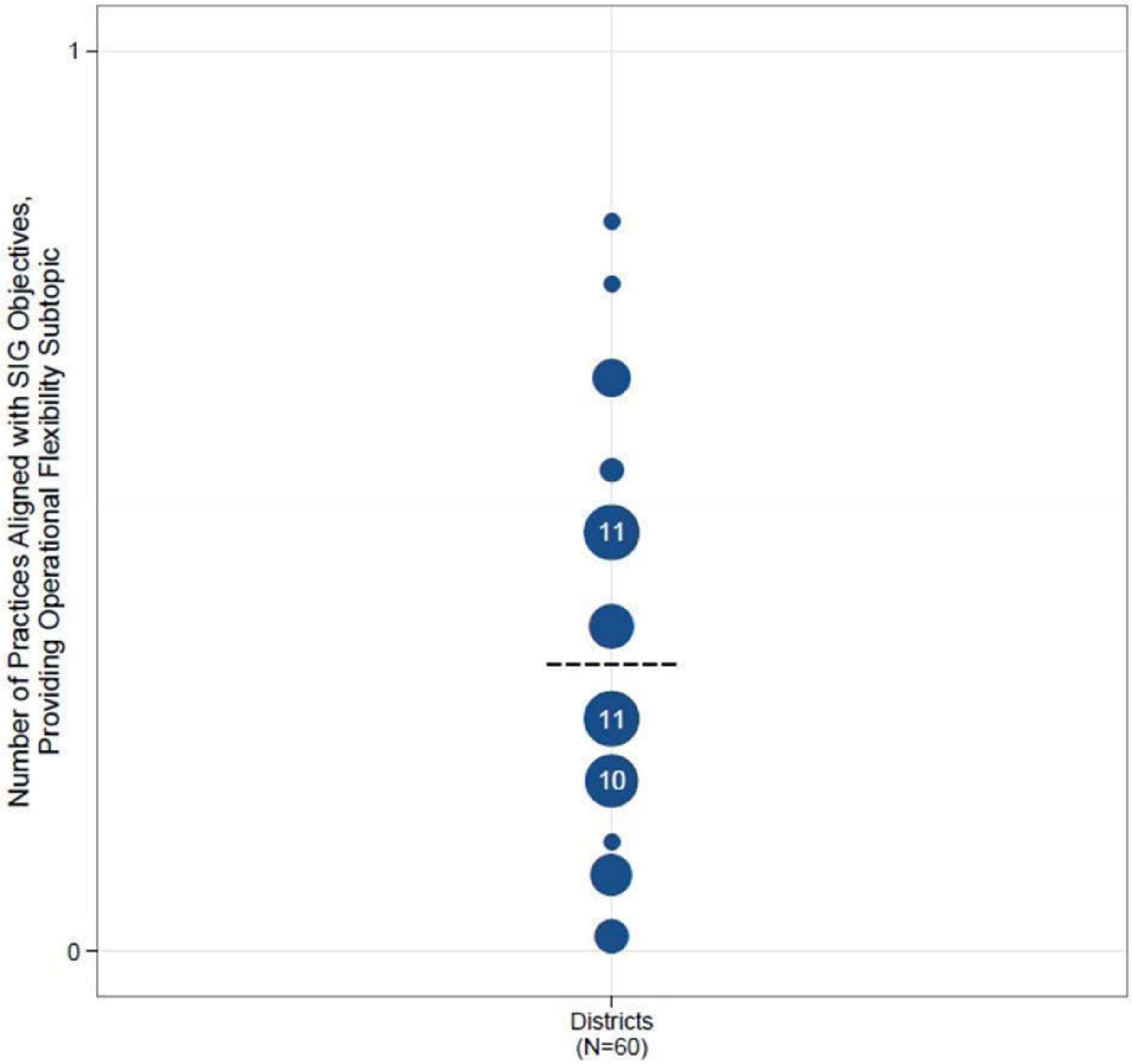


Source: Interviews with district administrators in spring 2013.

Note: This figure has a separate panel for each subtopic. We selected district interview questions that aligned with the practices described in the SIG application criteria. The practices shown on the horizontal axis of this figure are listed in Table D.4. For each practice in the SIG application criteria for which we identified one or more interview questions aligned with the practice, we calculated the percentage of interview questions with a “yes” response as a measure of the percentage of components each district used. The height of each bar represents the average percentage of the components of the practice that each group of districts used.

^Multiple district interview questions were used to assess whether districts used all of the components of this practice.

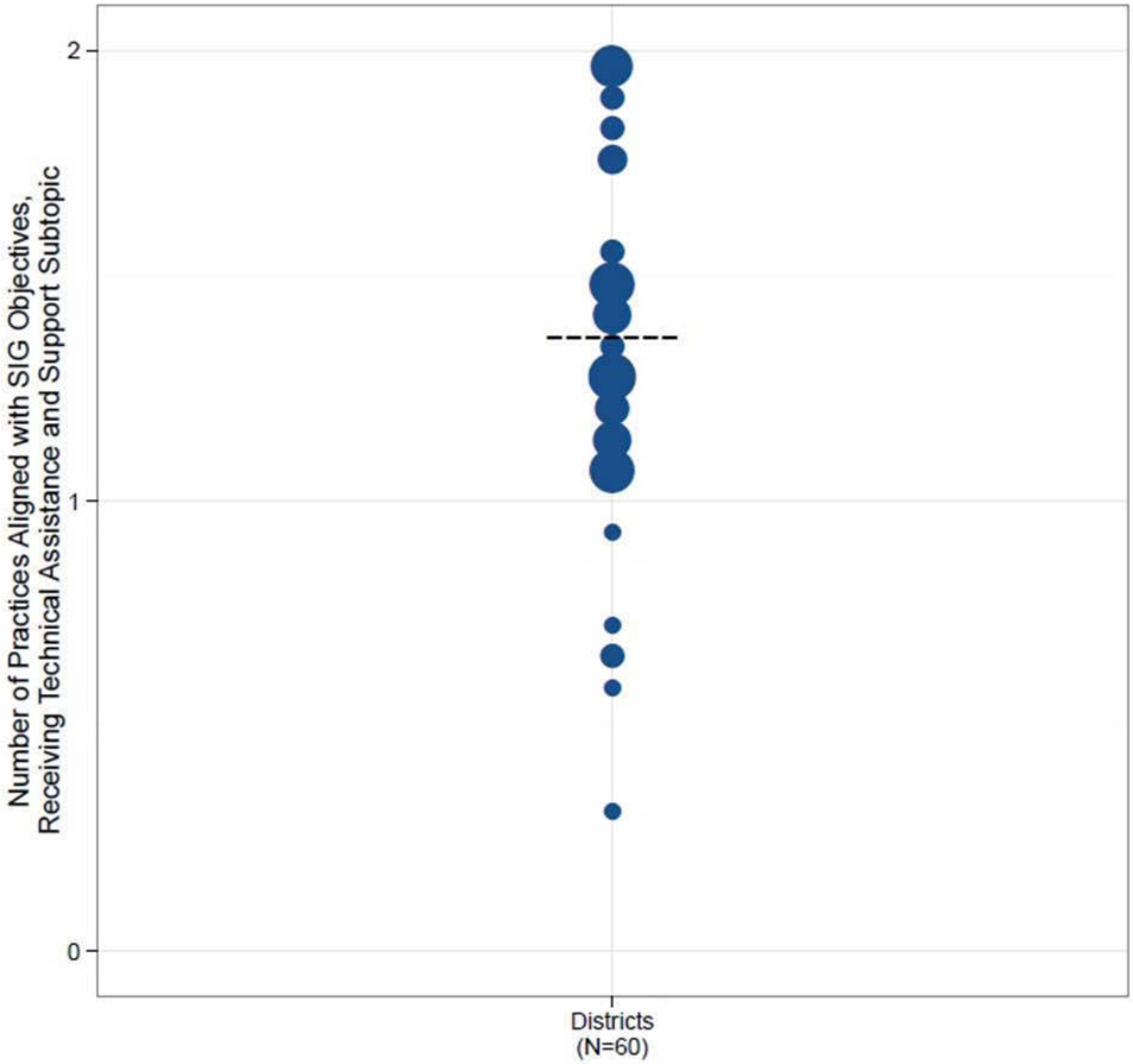
Figure D.24. Use of practices aligned with SIG, providing operational flexibility subtopic



Source: Interviews with district administrators in spring 2013.

Note: The practice summarized in this figure is presented in Table D.4. This figure presents one practice described in the SIG application criteria to which multiple interview questions aligned. As described in Appendix C, whenever multiple interview questions aligned with a single practice from the application criteria, we used those questions to construct a variable ranging from zero to one, with districts receiving a fraction of a point for each question to which they responded "yes." Each dot in this figure represents the districts that reported using a particular proportion of the interview questions aligned to the practice described in the SIG application criteria. The number inside each dot is the number of districts represented by the dot; dots that represent fewer than 10 districts have no number inside. The dashed line denotes the average value across all districts.

Figure D.25. Use of practices aligned with SIG, receiving technical assistance and support subtopic



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table D.4. Each dot in this figure represents the districts that reported using a particular number of practices (out of two examined) that were aligned with the SIG application criteria. Each dot in this figure represents fewer than 10 districts, so the numbers inside the dots have been removed to protect respondent confidentiality. For both practices, it was possible for a district to receive a fraction of one point. We determined the number of practices for each district using the same method that we used to determine the number of practices for each school (which is described in Appendix C). The dashed line denotes the average number of practices used across all districts.

APPENDIX E

DETAILED FINDINGS FROM INTERVIEWS AND SURVEYS

In Chapter IV and Appendix D, we summarized the extent to which schools and districts reported using SIG-promoted practices. In this appendix, we present detailed findings from the individual interview and survey questions underlying those practices, and describe how we analyzed those data. Specifically, we show the number of states and the percentage of districts, schools implementing a SIG-funded intervention model, and schools not implementing a SIG-funded model that responded "yes" to each question examined as part of this report. Readers interested in responses to individual interview and survey questions may, therefore, find this appendix useful.

The school-level data presented in this appendix are the same data used for the analyses presented in Chapter IV. The school-level tables in this appendix present results separately for schools implementing a SIG-funded model in 2012–2013 and schools not implementing a SIG-funded model in 2012–2013.

The state- and district-level data presented in this appendix came from structured telephone interviews with administrators in the 60 districts and 22 states where the SIG-sample schools were located. These interviews, conducted in spring 2012 and spring 2013, documented the state- and district-level supports for the school turnaround practices used by schools. The overarching research question answered by these findings is: How are states and districts supporting schools' efforts to use practices promoted by SIG? All 60 districts and 22 states in the SIG study sample included schools that were and were not implementing a SIG-funded model. Therefore, the state- and district-level tables in this appendix do not present comparisons; instead, they present descriptive information about the practices that districts and states reported using.

In Section A, we discuss how we analyzed data from closed- and open-ended questions and how we handled missing values. In Section B, we present findings from the interview questions in a series of tables, the titles of which are shown in the list of tables at the beginning of this report.

A. Analysis methods

Analyzing data from closed-ended questions. The evaluation's interviews and surveys comprised mostly closed-ended questions—that is, questions with yes-or-no responses or with a set of specific response categories from which to choose. As a result, these variables were already in a format that was suitable, or nearly suitable, for analysis.

Closed-ended questions sometimes included an "other-specify" response option so the interview or survey could progress smoothly when a respondent was uncertain about the response option that applied or could not find a response option that adequately captured the response he or she wished to provide. When a respondent chose this option, the interviewer asked the respondent to specify his or her response and recorded it. These "other-specify" responses were reviewed and either recoded into one of the existing structured response categories or coded into new response categories, as appropriate. Following reporting requirements established by the U.S. Department of Education's National Center for Education Statistics, we created a new response category only if at least three respondents (that is, states, districts, or schools) provided the same or similar response. If fewer than three respondents provided a particular response, the response remained part of the broad "other" category.

Analyzing data from open-ended questions. Whenever possible, we categorized the responses to open-ended questions into nominal categories (based on the themes that emerged) that could then be treated as quantitative, categorical data. This strategy enabled us to systematically identify and report on recurring themes mentioned frequently by respondents.

Handling missing values. Values can be missing for various reasons: (1) because the respondent did not complete the interview or survey; (2) because the respondent completed the interview or survey but did not complete the question; (3) because the respondent chose "don't know," "refused," or "not applicable"; or (4) because the question was logically skipped based on earlier responses. Generally, we excluded all missing values from our calculations regardless of the reason that the question was missing. That is, we did not recode a missing as a zero, with a few exceptions that are noted in the tables. In the tables presented in this appendix, we report the sample sizes for states, districts, and schools with nonmissing values on the given item. Percentages generally total 100 percent. In some cases, the number of states in a table totals more than 22 or the percentage of districts or schools totals more than 100 percent; we include a note in those tables explaining why. As one example, if the question asked the respondent to mark all responses that apply, respondents could choose multiple answers.

Selecting interview and survey questions aligned with the SIG application criteria. We reviewed the interview and survey questions and assigned those that aligned with the practices described in the SIG application criteria to specific topic areas and subtopics. We determined the subtopic into which each question fell based on the section of the SIG application criteria with which it aligned (see Appendix C for more details). In the tables presented in Section B, the last column of each table indicates whether each question was selected, and if it was selected, for which subtopic, by using the abbreviations shown in Table E.1.

Table E.1. Abbreviations for subtopics

Subtopic	Abbreviation
Topic area: comprehensive instructional reform strategies	
Using data to identify and implement an instructional program	IS-1
Promoting the continuous use of student data	IS-2
Providing supports and professional development to staff to assist both English language learners and students with disabilities	IS-3
Using and integrating technology-based supports	IS-4
Tailoring strategies for secondary schools	IS-5
Topic area: teacher and principal effectiveness	
Using rigorous, transparent, and equitable evaluation systems	TL-1
Identifying and rewarding effective teachers and principals and removing ineffective ones	TL-2
Providing high quality, job-embedded professional development or supports	TL-3
Implementing strategies to recruit, place, and retain staff	TL-4
Topic area: learning time and community-oriented schools	
Increasing learning time	TC-1
Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs	TC-2
Topic area: operational flexibility and support	
Having operational flexibility	FS-1
Receiving technical assistance and support	FS-2

Source: Surveys of school administrators in spring 2012 and spring 2013.

B. Detailed findings from interview and survey questions

The interviews and surveys were designed to cover policies and practices that were the focus of two grant programs—Race to the Top (RTT) and SIG. In this section we present findings from particular interview and survey questions that were most relevant to the SIG program. The tables are organized to follow the order of the modules in the interview and survey protocols, which was: (1) data systems, (2) teachers and leaders, (3) school turnaround, and (4) charter schools. The tables below indicate how the questions in each module align with the practices, topics, and subtopics described in the SIG application criteria and summarized in Table E.1. The protocols included a module on standards and assessments, but no questions from that module are presented here because it was designed to provide information about district and school use of standards and assessments policies promoted by RTT that were not a focus of SIG.

Table E.2. District reports of their schools' access to statewide longitudinal data systems and district data systems

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Reported that schools in their district have:		Yes (FS-2)
Both direct access to the SLDS data and access to district- or state-generated reports based on SLDS data		
2011–2012	65.5	
2012–2013	70.9	
Only direct access to SLDS data		
2011–2012	0.0	
2012–2013	0.0	
Only access to district- or state-generated reports based on SLDS data		
2011–2012	23.6	
2012–2013	14.5	
Access to neither type of information		
2011–2012	10.9	
2012–2013	14.5	
Reported that schools in their district have access to data from a district data system[a] that is distinct from the SLDS:		Yes (FS-2)
Both direct access to the district data and access to district-generated reports based on district data		
2011–2012	91.5	
2012–2013	98.3	
Only direct access to the district data, only access to district-generated reports based on district data, or access to neither type of information[b]		
2011–2012	8.5	
2012–2013	1.7	
Number of Districts	**60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

[a] District data system(s) (also identified as local instructional improvement systems) are defined by the U.S. Department of Education as technology-based tools and other strategies that provide teachers, principals, and administrators with meaningful support and actionable data to systemically manage continuous instructional improvement.

[b] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for "only direct access to the district data," "only access to district-generated reports based on district data," and "access to neither type of information."

SLDS = Statewide Longitudinal Data Systems; FS-2 = Receiving technical assistance and support.

Table E.3. District use of data analysis to monitor SIG school performance

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Reported tracking or monitoring the performance of SIG grantees implementing one of the four SIG intervention models by:		
Analyzing student achievement by grade level and/or by subject, by school		Yes (IS-2)
2011–2012	100.0	
2012–2013	100.0	
Analyzing student achievement data over time to identify trends		Yes (IS-2)
2011–2012	100.0	
2012–2013	100.0	
Examining other measures of student progress, such as benchmarks or diagnostic tests		Yes (IS-2)
2011–2012	96.6	
2012–2013	98.3	
Examining achievement gaps between groups of students, such as NCLB subgroups		Yes (IS-2)
2011–2012	96.6	
2012–2013	100.0	
Tracking graduation rates		Yes (IS-5)
2011–2012	100.0	
2012–2013	100.0	
Tracking student readiness for grade promotion or graduation		Yes (IS-5)
2011–2012	92.9	
2012–2013	92.9	
Tracking students' postsecondary enrollment and progress		Yes (IS-5)
2011–2012	54.0	
2012–2013	60.0	
Monitoring student attendance		Yes (IS-5)
2011–2012	100.0	
2012–2013	100.0	
Other analyses		No
2011–2012	75.9	
2012–2013	65.5	
Reported using different analyses for SIG schools compared with other schools in the district		No
2011–2012	18.6	
2012–2013	11.9	
Number of Districts	**50–60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample size because nonresponse varied across items.

IS-2 = Promoting the continuous use of student data; IS-5 = Tailoring strategies for secondary schools; NCLB = No Child Left Behind.

Table E.4. Purposes for which district staff use data

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Reported using data on all students from the SLDS, a district data system, or both, for the following purposes:		
To track overall school performance and identify areas for improvement		Yes (IS-2)
2011–2012	100.0	
2012–2013	100.0	
To evaluate instructional programs		Yes (IS-1)
2011–2012	93.2	
2012–2013	93.2	
To guide development and implementation of academic supports or enrichment programs		Yes (IS-2)
2011–2012	98.3	
2012–2013	98.3	
To guide development and implementation of nonacademic supports or enrichment programs (for example, counseling)		Yes (TC-2)
2011–2012	82.5	
2012–2013	80.7	
To track students' progress toward graduation		Yes (IS-5)
2011–2012	94.7	
2012–2013	100.0	
To track students' postsecondary enrollment and progress		No
2011–2012	56.1	
2012–2013	40.4	
To inform professional development offerings for teachers, principals, or other school leaders		Yes (TL-3)
2011–2012	89.8	
2012–2013	89.8	
To evaluate the success of professional development offerings for teachers, principals, or other school leaders		Yes (TL-3)
2011–2012	64.4	
2012–2013	69.5	
To inform other decisions regarding individual teachers, principals, or other school leaders (such as tenure, retention, or bonus decisions)		Yes (TL-2)
2011–2012	66.1	
2012–2013	71.2	
To inform resource allocation to improve instruction		Yes (IS-2)
2011–2012	93.2	
2012–2013	98.3	
For other purposes		No
2011–2012	39.7	
2012–2013	32.8	
Reported using data on ELLs from the SLDS, a district data system, or both, for the following purposes:		
To make decisions about students' entry into and/or exit from ELL status		Yes (IS-3)
2011–2012	98.3	
2012–2013	98.3	

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
To place ELLs into specialized programs and classes		Yes (IS-3)
2011–2012	98.3	
2012–2013	96.6	
To track the progress of current ELLs		Yes (IS-3)
2011–2012	98.3	
2012–2013	100.0	
To track the progress of former ELLs		Yes (IS-3)
2011–2012	89.1	
2012–2013	89.1	
To inform, improve, or differentiate instruction for ELLs		Yes (IS-2)
2011–2012	93.1	
2012–2013	96.6	
To identify professional development needs for teachers of ELLs		Yes (IS-3)
2011–2012	85.7	
2012–2013	78.6	
To assess teacher effectiveness with ELLs		Yes (IS-3)
2011–2012	57.9	
2012–2013	70.2	
For other purposes		No
2011–2012	48.2	
2012–2013	32.1	
Number of Districts	**60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

SLDS = Statewide Longitudinal Data System; ELLs = English language learners; IS-1 = Using data to identify and implement an instructional program; IS-2 = Promoting the continuous use of student data; IS-3 = Providing supports and professional development to staff to assist ELLs and students with disabilities; IS-5 = Tailoring strategies for secondary schools; TC-2 = Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs; TL-2 = Identifying and rewarding effective teachers and removing ineffective ones; TL-3 = Providing high-quality, job-embedded professional development or supports.

Table E.5. Purposes for which school staff use data

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not Implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported using data for the following purposes:			
To evaluate instructional programs			Yes (IS-1)
2011–2012	96.7	92.8	
2012–2013	95.9	94.0	
To guide development and implementation of academic supports or enrichment programs			Yes (IS-2)
2011–2012	100.0	98.2	
2012–2013	100.0	100.0	
To guide development and implementation of nonacademic supports or enrichment programs (for example, counseling)			Yes (TC-2)
2011–2012	91.4	82.0	
2012–2013	89.2	84.4	
To inform teachers' instructional practices			Yes (IS-2)
2011–2012	98.5	97.0	
2012–2013	100.0	97.0	
To inform professional development offerings			Yes (TL-3)
2011–2012	95.9	92.8	
2012–2013	97.0	94.0	
To evaluate the success of professional development offerings			Yes (TL-3)
2011–2012	80.7	73.9	
2012–2013	86.2	77.0	
To track individual student performance and identify areas of improvement for specific students			Yes (IS-2)
2011–2012	98.5	100.0	
2012–2013	98.9	98.8	
To track students' progress toward high school graduation[a]			Yes (IS-5)
2011–2012	97.6	98.7	
2012–2013	100.0	97.4	
To track students' preparation for college enrollment[a]			Yes (IS-5)
2011–2012	90.2	94.7	
2012–2013	87.8	96.1	
To track students' postsecondary enrollment and progress[a]			No
2011–2012	74.0	85.7	
2012–2013	66.7	79.2	
To inform resource allocation to improve instruction			Yes (IS-2)
2011–2012	88.8	82.0	
2012–2013	88.0	84.5	
Other purpose			No
2011–2012	12.0	13.0	
2012–2013	13.2	13.0	

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not Implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Among schools that reported having ELLs, reported using data on ELLs for the following purposes:			
To make decisions about students' entry into and/or exit from ELL status			Yes (IS-3)
2011–2012	96.6	95.9	
2012–2013	97.7	92.6	
To place ELLs into specialized programs and classes			Yes (IS-3)
2011–2012	94.2	90.8	
2012–2013	97.1	94.1	
To track the progress of current ELLs			Yes (IS-3)
2011–2012	97.2	96.8	
2012–2013	97.7	93.5	
To track the progress of former ELLs			Yes (IS-3)
2011–2012	75.9	68.6	
2012–2013	78.3	75.4	
To inform, improve, or differentiate instruction for ELLs			Yes (IS-2)
2011–2012	92.6	90.1	
2012–2013	93.7	90.1	
To identify professional development needs for teachers of ELLs			Yes (IS-3)
2011–2012	80.1	73.8	
2012–2013	80.7	80.3	
To assess teacher effectiveness with ELLs			Yes (IS-3)
2011–2012	78.5	75.0	
2012–2013	78.5	77.5	
Other purpose			No
2011–2012	0.0	3.6	
2012–2013	3.8	6.4	
Number of Schools	**120–270**	**80–170**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because missing data varied across items. For example, some questions shown in the table were only applicable to certain schools (such as high schools).

[a] The analysis for this row includes only high schools.

ELLs = English language learners; IS-1 = Using data to identify and implement an instructional program; IS-2 = Promoting the continuous use of student data; IS-3 = Providing supports and professional development to staff to assist ELLs and students with disabilities; IS-5 = Tailoring strategies for secondary schools; TC-2 = Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs; TL-3 = Providing high-quality, job-embedded professional development or supports.

Table E.6. Supports for data use

	Percentage of Low-Performing Schools (Unless Otherwise Specified)		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported receiving the following support to help school staff access and use data:			
Funds to support school investments related to data use			Yes (FS-2)
2011–2012	66.1	42.8	
2012–2013	67.3	41.6	
Hardware or software to facilitate data use			Yes (FS-2)
2011–2012	54.3	38.4	
2012–2013	53.4	43.3	
Materials on how to access and use data to differentiate or improve instruction			Yes (FS-2)
2011–2012	56.0	44.5	
2012–2013	56.0	49.4	
Other type of support			No
2011–2012	15.6	10.3	
2012–2013	12.7	10.3	
Reported having a designated staff person who supports the use of data by teachers			No
2011–2012	91.4	86.3	
2012–2013	88.0	81.5	
Reported providing scheduled time for teachers to examine data, either on their own or in collaboration with others			Yes (TL-3)
2011–2012	96.5	95.2	
2012–2013	98.8	97.0	
Reported that their school leaders coached teachers on the use of data to:			
Improve instruction			Yes (TL-3)
2011–2012	98.1	96.3	
2012–2013	99.3	97.6	
Among schools that reported having ELLs, improve instruction of ELLs			Yes (IS-3)
2011–2012	80.1	72.6	
2012–2013	79.5	72.6	
Reported receiving professional development, training, or technical assistance to help school staff access data, navigate data systems, or interpret and use data			Yes (TL-3)
2011–2012	90.5	85.5	
2012–2013	87.9	84.9	
Average reported number of hours this professional development, training, or technical assistance was provided to:[a]			
School administrators			No
2011–2012	19.2	14.3	
2012–2013	17.4	13.5	

	Percentage of Low-Performing Schools (Unless Otherwise Specified)		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Teachers			No
2011–2012	25.2	15.2	
2012–2013	22.8	14.6	
Among schools that reported having ELLs, reported receiving the following supports to help school staff access and use data related to ELLs:			
Supports to use data to track the performance of ELLs			Yes (IS-3)
2011–2012	60.6	57.0	
2012–2013	60.0	56.1	
Supports to use data to improve or differentiate instruction for ELLs			Yes (IS-3)
2011–2012	58.2	58.9	
2012–2013	61.2	53.6	
Other supports to use data about ELLs			No
2011–2012	35.2	29.8	
2012–2013	31.5	28.9	
Number of Schools	**170–270**	**110–170**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because missing data varied across items. For example, some questions shown in the table were only applicable to certain schools (such as schools that reported having ELLs).

[a] Schools that reported they did not receive professional development, training, or technical assistance to help school administrators and/or teachers access data, navigate data systems, or interpret and use data to improve and/or differentiate instruction are included in the analysis of this question as 0 responses.

ELLs = English language learners; FS-2 = Receiving technical assistance and support; IS-3 = Providing supports and professional development to staff to assist ELLs and students with disabilities; TL-3 = Providing high-quality, job-embedded professional development or supports.

Table E.7. District requirements for teacher evaluations

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Reported that all schools must use the same teacher evaluation model		No
2011–2012	89.8	
2012–2013	89.8	
Reported that student achievement growth was required		Yes (TL-4)
2011–2012	47.5	
2012–2013	42.4	
Reported that student achievement growth was required with the following weight[a]		Yes (TL-4)
No specific weight required or did not require student achievement growth, or other[b]		
2011–2012	81.0	
2012–2013	72.4	
1–34[c]		
2011–2012	6.9	
2012–2013	5.2	
35–50		
2011–2012	12.1	
2012–2013	22.4	
51 or more		
2011–2012	0.0	
2012–2013	0.0	
“Significant”, “Substantial,” or “Primary” factor		
2011–2012	0.0	
2012–2013	0.0	
Reported using the following number of rating levels for overall teacher evaluations		No
4 or more		
2011–2012	57.6	
2012–2013	66.1	
3 rating levels		
2011–2012	20.3	
2012–2013	20.3	
2 rating levels		
2011–2012	22.0	
2012–2013	13.6	
Do not specify minimum or no rating levels		
2011–2012	0.0	
2012–2013	0.0	
Number of Districts	**60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

[a] Districts that responded no to a question about whether the district required student achievement growth are included in the analysis of this question and are considered to not be using this practice.

[b] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for “no specific weight required or did not require student achievement growth” and “other.”

[c] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for “1–20” and “21–34.”

TL-4 = Using rigorous, transparent, and equitable evaluation systems.

Table E.8. District-reported requirements for performance measures (other than student achievement growth) for evaluations of teachers in *tested* grades and/or subjects

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Classroom observations		Yes (TL-4)
2009–2010[a]	93.2	
2011–2012	93.2	
2012–2013	98.3	
Self-assessment		Yes (TL-4)
2009–2010[a]	32.2	
2011–2012	42.4	
2012–2013	52.5	
Portfolios or other artifacts of teacher practice		Yes (TL-4)
2009–2010[a]	27.1	
2011–2012	32.2	
2012–2013	42.4	
Peer assessments other than classroom observations		Yes (TL-4)
2009–2010[a]	–[b]	
2011–2012	–[b]	
2012–2013	8.6	
Student work samples		Yes (TL-4)
2009–2010[a]	16.9	
2011–2012	25.4	
2012–2013	30.5	
Student surveys or other feedback		Yes (TL-4)
2009–2010[a]	–[b]	
2011–2012	8.6	
2012–2013	10.3	
Parent surveys or other feedback		Yes (TL-4)
2009–2010[a]	–[b]	
2011–2012	–[b]	
2012–2013	8.6	
Other measures		No
2009–2010[a]	16.9	
2011–2012	25.4	
2012–2013	16.9	
Number of Districts	**60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

[a] Data from 2009–2010 were collected retrospectively in spring 2012. All data provided by districts were self-reported and not independently verified by the research team. For these reasons and potential concerns about recall accuracy, readers should exercise caution when interpreting data from 2009–2010.

[b] This cell has been suppressed to protect respondent confidentiality.

TL-4 = Using rigorous, transparent, and equitable evaluation systems.

Table E.9. District-reported requirements for performance measures (other than student achievement growth) for evaluations of teachers *in nontested grades and subjects*

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Classroom observations		Yes (TL-4)
2009–2010[a]	93.2	
2011–2012	93.2	
2012–2013	96.6	
Self-assessment		Yes (TL-4)
2009–2010[a]	32.2	
2011–2012	40.7	
2012–2013	52.5	
Portfolios or other artifacts of teacher practice		Yes (TL-4)
2009–2010[a]	28.8	
2011–2012	33.9	
2012–2013	44.1	
Peer assessments other than classroom observations		Yes (TL-4)
2009–2010[a]	–[b]	
2011–2012	–[b]	
2012–2013	8.6	
Student work samples		Yes (TL-4)
2009–2010[a]	18.6	
2011–2012	25.4	
2012–2013	28.8	
Student surveys or other feedback		Yes (TL-4)
2009–2010[a]	5.2	
2011–2012	8.6	
2012–2013	8.6	
Parent surveys or other feedback		Yes (TL-4)
2009–2010[a]	–[b]	
2011–2012	–[b]	
2012–2013	8.6	
Other measures		No
2009–2010[a]	16.9	
2011–2012	25.4	
2012–2013	18.6	
Number of Districts	**60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

[a] Data from 2009–2010 were collected retrospectively in spring 2012. All data provided by districts were self-reported and not independently verified by the research team. For these reasons and potential concerns about recall accuracy, readers should exercise caution when interpreting data from 2009–2010.

[b] This cell has been suppressed to protect respondent confidentiality.

TL-4 = Using rigorous, transparent, and equitable evaluation systems.

Table E.10. District-reported policies for tenure and frequency of teacher evaluation

District-Reported Regulation	Percentage of Districts (Unless Otherwise Specified)	Item Aligned with SIG Application (Subtopic)
Allow teachers to earn tenure[a]		No
2011–2012	81.4	
2012–2013	81.4	
Have a probationary period for all or some teachers		No
2011–2012	96.6	
2012–2013	93.1	
Among districts with a probationary period, reported mean duration of probationary period (years)		No
2011–2012	2.8	
2012–2013	3.0	
Among districts with a probationary period, evaluate probationary teachers		No
Three or more times per year or other interval[b]		
2011–2012	17.0	
2012–2013	15.1	
Two times per year		
2011–2012	41.5	
2012–2013	37.7	
Annually		
2011–2012	41.5	
2012–2013	47.2	
Every other year		
2011–2012	0.0	
2012–2013	0.0	
Evaluate non-probationary teachers		No
Three or more times per year		
2011–2012	5.1	
2012–2013	5.1	
Two times per year		
2011–2012	10.2	
2012–2013	10.2	
Annually		
2011–2012	44.1	
2012–2013	49.2	
Every other year		
2011–2012	20.3	
2012–2013	23.7	
Other interval		
2011–2012	20.3	
2012–2013	11.9	
Number of Districts	**50–60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample size because missing data varied across items. For example, some questions shown in the table were only applicable to certain districts (such as districts that reported having a probationary period).

[a] This includes districts that provide teachers with some other continuing right to their job that the district does not refer to as “tenure.”

[b] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for “three or more times per year” and “other interval.”

Table E.11. School-reported policies for using student achievement growth in teacher evaluations

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (Subtopic)
Reported that student achievement growth was required			Yes (TL-4)
2011–2012	53.7	43.4	
2012–2013	57.9	57.2	
Reported that student achievement growth was required with a specific weight:[a]			Yes (TL-4)
No specific weight required or did not require student achievement growth			
2011–2012	53.5	66.0	
2012–2013	50.9	48.7	
1–20			
2011–2012	8.3	2.7	
2012–2013	10.4	9.3	
21–34			
2011–2012	3.9	5.3	
2012–2013	5.7	8.0	
35–50			
2011–2012	12.6	10.0	
2012–2013	10.4	9.3	
51 or more, or “Significant,” “Substantial,” or “Primary” factor[b]			
2011–2012	4.8	5.3	
2012–2013	7.4	8.7	
Other			
2011–2012	17.0	10.7	
2012–2013	15.2	16.0	
Number of Schools	**230–260**	**150–170**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Schools that responded no to a question about whether student achievement growth was required as a component of teacher evaluations are in included in the analysis of this question and are considered to not be using this practice.

[b] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for “51 or more” and “significant, substantial, or primary factor.”

TL-4 = Using rigorous, transparent, and equitable evaluation systems.

Table E.12. School-reported performance measures (other than student achievement growth) for teacher evaluations

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Classroom observations			Yes (TL-4)
2011–2012	98.1	98.8	
2012–2013	99.2	98.8	
Self-assessment			Yes (TL-4)
2011–2012	60.2	52.1	
2012–2013	59.8	59.5	
Peer assessment			Yes (TL-4)
2011–2012	21.5	18.0	
2012–2013	25.2	17.4	
Portfolios or other artifacts of teacher practice			Yes (TL-4)
2011–2012	43.1	36.8	
2012–2013	48.0	47.2	
Student work samples			Yes (TL-4)
2011–2012	47.3	42.7	
2012–2013	52.2	47.6	
Student surveys or other feedback			Yes (TL-4)
2011–2012	28.6	27.8	
2012–2013	26.6	32.7	
Parent surveys or other feedback			Yes (TL-4)
2011–2012	26.7	23.0	
2012–2013	21.7	27.3	
Other measures			No
2011–2012	10.7	8.9	
2012–2013	6.4	10.8	
Number of Schools	**230–260**	**160–170**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

TL-4 = Using rigorous, transparent, and equitable evaluation systems.

Table E.13. School-reported policies for tenure and frequency of teacher evaluation

	Percentage of Low-Performing Schools (Unless Otherwise Specified)		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Allow teachers to earn tenure[a]			No
2011–2012	72.9	72.0	
2012–2013	73.6	62.7	
Have a probationary period for teachers			No
2011–2012	95.8	97.5	
2012–2013	95.3	97.5	
Among schools reporting any duration for probationary period, mean duration of probationary period (years)			No
2011–2012	2.6	2.7	
2012–2013	2.6	2.6	
Among schools with a probationary period, evaluate probationary teachers			No
Three or more times per year			
2011–2012	22.0	16.7	
2012–2013	17.4	17.3	
Two times per year			
2011–2012	45.6	42.9	
2012–2013	49.4	40.4	
Annually			
2011–2012	21.6	31.4	
2012–2013	23.2	32.1	
Every other year or other interval[b]			
2011–2012	10.8	9.0	
2012–2013	10.0	10.3	
Evaluate non probationary teachers:			No
Three or more times per year			
2011–2012	4.6	12.0	
2012–2013	4.2	7.0	
Two times per year			
2011–2012	28.8	26.6	
2012–2013	34.2	35.4	
Annually			
2011–2012	35.8	33.5	
2012–2013	31.7	32.3	
Every other year			
2011–2012	20.0	22.2	
2012–2013	18.8	14.6	
Other			
2011–2012	10.8	5.7	
2012–2013	11.3	10.8	
Number of Schools	**180–260**	**120–160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items. For example, some questions shown in the table were only applicable to certain schools (such as schools that reported having a probationary period).

[a] This includes schools that provide teachers with some other continuing right to their job that is not referred to as “tenure.”

[b] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for “every other year” and “other interval.”

Table E.14. School-reported uses of teacher evaluation results

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported using teacher evaluation results to guide decisions about:			
Professional development and/or support			Yes (TL-2)
2011–2012	84.4	77.9	
2012–2013	82.5	77.9	
Annual salary increases			Yes (TL-2)
2011–2012	10.0	14.1	
2012–2013	11.1	17.8	
Bonuses or other performance-based compensation (other than annual salary increases)[a]			Yes (TL-2)
2011–2012	14.6	19.4	
2012–2013	17.3	13.3	
Career-advancement opportunities[b]			Yes (TL-4)
2011–2012	41.6	41.5	
2012–2013	34.5	40.2	
Reductions in force and excessing decisions			Yes (TL-4)
2011–2012	24.4	30.8	
2012–2013	27.2	34.6	
Number of Schools	**250–260**	**160–170**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Schools that responded no to a question about whether any teachers in their school have the opportunity to receive bonuses or other performance-based compensation (other than annual salary increases) are included in the analysis of this question and are considered to not be using this practice.

[b] Schools that responded no to a question about whether any teacher in their school have career-advancement opportunities available to them are included in the analysis of this question and are considered to not be using this practice.

TL-2 = Identifying and rewarding effective teachers and removing ineffective ones; TL-4 = Implementing strategies to recruit, place, and retain staff.

Table E.15. District principal evaluation requirements

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Reported that all schools must use the same principal evaluation model		No
2011–2012	96.6	
2012–2013	93.2	
Reported that student achievement growth was required		Yes (TL-1)
2011–2012	59.3	
2012–2013	55.9	
Reported using the following number of rating categories for overall performance:		No
4 or more rating levels		
2011–2012	69.0	
2012–2013	74.6	
3 rating levels		
2011–2012	13.8	
2012–2013	10.2	
2 rating levels, no rating levels, or do not specify minimum number of rating levels[a]		
2011–2012	17.2	
2012–2013	15.3	
Number of Districts	**60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

[a] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for “2 rating levels” and “do not specify minimum number of rating levels or no rating levels.”

TL-1 = Using rigorous, transparent, and equitable evaluation systems.

Table E.16. District-reported requirements for performance measures for principal evaluations

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Student achievement growth		Yes (TL-1)
2009–2010[a]	44.1	
2011–2012	59.3	
2012–2013	55.9	
Self-assessment		Yes (TL-1)
2009–2010[a]	49.2	
2011–2012	61.0	
2012–2013	66.1	
District administrator input		Yes (TL-1)
2009–2010[a]	89.8	
2011–2012	93.2	
2012–2013	84.7	
Staff input		Yes (TL-1)
2009–2010[a]	15.3	
2011–2012	20.3	
2012–2013	23.7	
Student input		Yes (TL-1)
2009–2010[a]	6.8	
2011–2012	13.6	
2012–2013	13.6	
Other measures		No
2009–2010[a]	22.0	
2011–2012	25.4	
2012–2013	30.5	
Number of Districts	**60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

[a] Data from 2009–2010 were collected retrospectively in spring 2012. All data provided by districts were self-reported and not independently verified by the research team. For these reasons and potential concerns about recall accuracy, readers should exercise caution when interpreting data from 2009–2010.

TL-1 = Using rigorous, transparent, and equitable evaluation systems.

Table E.17. School-reported performance measures for principal evaluations

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported that student achievement growth was used			Yes (TL-1)
2011–2012	76.0	75.2	
2012–2013	76.8	78.2	
Reported that student achievement growth was required with the following weight:[a]			Yes (TL-1)
No specific weight required or did not require student achievement growth			
2011–2012	31.9	31.7	
2012–2013	30.9	28.5	
1–20			
2011–2012	11.7	7.4	
2012–2013	11.7	5.7	
21–34			
2011–2012	2.6	3.3	
2012–2013	2.1	5.7	
35–50			
2011–2012	12.0	10.6	
2012–2013	15.6	16.3	
51 or more			
2011–2012	3.6	8.9	
2012–2013	2.6	4.1	
“Significant”, “Substantial” or “Primary” factor			
2011–2012	13.5	22.8	
2012–2013	16.7	16.3	
Other			
2011–2012	25.0	15.4	
2012–2013	20.8	24.4	
Reporting using other measures:			
Self-assessment			Yes (TL-1)
2011–2012	64.5	58.3	
2012–2013	70.1	66.3	
District administrator input			Yes (TL-1)
2011–2012	93.2	90.7	
2012–2013	91.6	92.0	
School staff surveys or other feedback			Yes (TL-1)
2011–2012	42.2	41.5	
2012–2013	44.2	60.4	
Student surveys or other feedback			Yes (TL-1)
2011–2012	31.5	31.3	
2012–2013	32.8	35.6	
Other measures			No
2011–2012	11.5	11.9	
2012–2013	12.3	14.5	
Number of Schools	**190–250**	**120–170**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Schools that responded no to a question about whether student achievement growth was used as a component of principal evaluations are included in the analysis of this question and are considered to not be using this practice.

TL-1 = Using rigorous, transparent, and equitable evaluation systems.

Table E.18. School-reported uses of principal evaluation results

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported using results to guide decisions about:			
Professional development and/or support			No
2011–2012	51.0	48.5	
2012–2013	49.0	47.2	
Annual salary increases			Yes (TL-3)
2011–2012	12.3	20.2	
2012–2013	11.9	19.0	
Bonuses or other performance-based compensation (other than regular salary increases)[a]			Yes (TL-2)
2011–2012	14.3	16.1	
2012–2013	12.7	16.1	
Number of Schools	**250–260**	**160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Schools that reported that the principal does not have the opportunity to receive a bonus or other performance-based compensation (other than regular salary increases) are included in the analysis of this question and are considered to not be using this practice.

TL-2 = Identifying and rewarding effective principals and removing ineffective ones; TL-3 = Providing high-quality, job-embedded professional development or supports.

Table E.19. District use of financial incentives to recruit or retain effective staff in SIG schools implementing one of the SIG-funded intervention models

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Reported offering financial incentives		Yes (TL-4)
2011–2012	64.4	
2012–2013	59.3	
Reported offering the following types of financial incentives:[a]		
Signing/recruitment bonuses for:		
Teachers		Yes (TL-4)
2011–2012	23.7	
2012–2013	20.3	
Principals		Yes (TL-4)
2011–2012	23.7	
2012–2013	20.3	
Retention bonuses for:		
Teachers		Yes (TL-4)
2011–2012	15.3	
2012–2013	16.9	
Principals		Yes (TL-4)
2011–2012	11.9	
2012–2013	15.3	
Performance bonuses for:		
Teachers		Yes (TL-4)
2011–2012	39.7	
2012–2013	41.4	
Principals		Yes (TL-4)
2011–2012	36.2	
2012–2013	36.2	
Increased annual compensation, other than bonuses, for:		
Teachers		Yes (TL-4)
2011–2012	28.8	
2012–2013	30.5	
Principals		Yes (TL-4)
2011–2012	15.3	
2012–2013	18.6	
Loan forgiveness for:		
Teachers		Yes (TL-4)
2011–2012	13.6	
2012–2013	13.6	
Principals		Yes (TL-4)
2011–2012	6.8	
2012–2013	6.8	
Tuition reimbursement for:		
Teachers		Yes (TL-4)
2011–2012	20.3	
2012–2013	25.4	

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Principals		Yes (TL-4)
2011–2012	11.9	
2012–2013	13.6	
Housing (purchase or rent) assistance for:		
Teachers		Yes (TL-4)
2011–2012		
2012–2013	5.1	
Principals		Yes (TL-4)
2011–2012	–[b]	
2012–2013	–[b]	
Financial incentives targeted toward increasing the number of staff with English language learner expertise in SIG schools for:		
Teachers		Yes (TL-4)
2011–2012	11.9	
2012–2013	10.2	
Principals		Yes (TL-4)
2011–2012	–[b]	
2012–2013	–[b]	
Other financial incentives for:		
Teachers		No
2011–2012	25.4	
2012–2013	15.3	
Principals		No
2011–2012	11.9	
2012–2013	6.8	
Number of Districts	**60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

[a] Districts that answered no to a question about whether the district offered any financial incentives to help recruit or retain effective teachers and/or principals are included in the analysis of this question and are considered to not be using this practice.

[b] This cell has been suppressed to protect respondent confidentiality.

TL-4 = Implementing strategies to recruit, place, and retain staff.

Table E.20. District use of nonfinancial strategies to recruit or retain effective staff in SIG schools implementing one of the SIG-funded intervention models

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Reported that principals had discretion to decide which staff to hire		Yes (TL-4)
2011–2012	81.4	
2012–2013	84.7	
Reported modifying teacher tenure rules that affect placement and/or removal		Yes (TL-4)
2011–2012	20.7	
2012–2013	29.3	
Reported using retention or recruitment efforts targeted toward increasing the number of staff with English language learner expertise		Yes (TL-4)
2011–2012	37.9	
2012–2013	46.6	
Reported increasing the amount of induction support for novice teachers (above and beyond that provided to all novice teachers in the district) with the goal of increasing retention		Yes (TL-4)
2011–2012	46.6	
2012–2013	43.1	
Other strategies		No
2011–2012	22.0	
2012–2013	18.6	
Number of Districts	**60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

TL-4 = Implementing strategies to recruit, place, and retain staff.

Table E.21. School-reported opportunities for staff to receive financial incentives

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported offering signing or recruitment bonuses for:			
Teachers			Yes (TL-4)
2011–2012	16.6	14.0	
2012–2013	13.0	12.0	
Principals			Yes (TL-4)
2011–2012	10.5	8.4	
2012–2013	10.9	10.5	
Reported offering retention bonuses for:			
Teachers			Yes (TL-4)
2011–2012	9.8	7.2	
2012–2013	10.7	5.2	
Principals			Yes (TL-4)
2011–2012	7.9	2.8	
2012–2013	5.7	4.2	
Reported offering performance bonuses for:			
Teachers			Yes (TL-4)
2011–2012	38.1	36.4	
2012–2013	38.5	29.1	
Principals			Yes (TL-4)
2011–2012	39.7	33.1	
2012–2013	38.4	32.4	
Reported increasing annual compensation other than bonuses, for:			
Teachers			Yes (TL-4)
2011–2012	33.3	24.5	
2012–2013	24.3	18.5	
Principals			Yes (TL-4)
2011–2012	29.3	20.8	
2012–2013	19.1	17.4	
Reported offering loan forgiveness for:			
Teachers			Yes (TL-4)
2011–2012	56.6	50.0	
2012–2013	54.5	52.0	
Principals			Yes (TL-4)
2011–2012	15.9	10.5	
2012–2013	15.9	9.8	
Reported offering tuition reimbursement for:			
Teachers			Yes (TL-4)
2011–2012	40.2	32.7	
2012–2013	42.7	35.3	
Principals			Yes (TL-4)
2011–2012	29.7	18.4	
2012–2013	29.3	19.1	

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported offering housing (purchase or rent) assistance for:			
Teachers			Yes (TL-4)
2011–2012	10.0	12.0	
2012–2013	9.5	12.0	
Principals			Yes (TL-4)
2011–2012	6.8	5.0	
2012–2013	5.0	5.8	
Among schools that reported having ELLs, reported offering financial incentives targeted toward increasing the number of staff with English language learner expertise in the school for:			
Teachers			Yes (TL-4)
2011–2012	5.4	9.2	
2012–2013	4.8	7.5	
Principals			Yes (TL-4)
2011–2012	-[a]	4.2	
2012–2013	3.1	3.1	
Reported offering other financial incentives for:			
Teachers			Yes (TL-4)
2011–2012	13.3	2.8	
2012–2013	6.6	-[a]	
Principals			Yes (TL-4)
2011–2012	6.9	0.0	
2012–2013	6.5	4.4	
Number of Schools	**130–250**	**100–150**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because missing data varied across items. For example, some questions shown in the table were only applicable to certain schools (such as schools that reported having ELLs).

[a] This cell has been suppressed to protect respondent confidentiality.

TL-4 = Implementing strategies to recruit, place, and retain staff.

Table E.22. School-reported use of nonfinancial strategies to recruit and retain staff

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported that principal had discretion to decide which staff to hire			Yes (TL-4)
2011–2012	66.4	68.5	
2012–2013	66.8	64.8	
Reported offering increased induction support for novice teachers in the school (above and beyond that provided to all novice teachers in the district)			Yes (TL-3)
2011–2012	56.4	52.2	
2012–2013	49.2	44.7	
Reported engaging in the following activities:			
Provided additional professional development, mentoring, and/or instructional coaching to teachers and/or school leaders[a]			Yes (TL-3)
2011–2012	98.0	95.7	
2012–2013	94.8	88.8	
Improved opportunities for collaboration (such as common planning time)			Yes (TL-3)
2011–2012	95.3	91.4	
2012–2013	97.2	92.0	
Improved the quality of school facilities			No
2011–2012	74.2	68.9	
2012–2013	73.8	70.8	
Increased availability of classroom or instructional supplies			No
2011–2012	89.6	80.4	
2012–2013	88.8	73.6	
Enhanced safety measures in the building			Yes (TC-2)
2011–2012	80.5	74.1	
2012–2013	82.9	76.5	
Increased access to technology for teachers			Yes (IS-4)
2011–2012	94.0	81.6	
2012–2013	91.6	81.0	
Offered more flexible work conditions (for example, flexible schedule)			Yes (TL-4)
2011–2012	29.4	30.4	
2012–2013	29.0	28.0	
Increased the use of aides/paraprofessionals			Yes (TL-4)
2011–2012	48.6	41.7	
2012–2013	42.9	37.4	
Increased the use of volunteers (for example, parents)			Yes (TC-2)
2011–2012	64.0	52.8	
2012–2013	58.3	52.8	
Other activities			No
2011–2012	10.2	5.7	
2012–2013	7.1	5.7	
Number of Schools	**230–250**	**160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Includes principals, assistant principals, or department heads.

TL-3 = Providing high-quality, job-embedded professional development or supports; TL-4 = Implementing strategies to recruit, place, and retain staff; TC-2 = Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs; IS-4 = Using and integrating technology-based supports.

Table E.23. Funds to support school improvement efforts

	Percentage of Low-Performing Schools		
	Implementing a SIG-Funded Intervention Model in 2012-2013	Not Implementing a SIG-Funded Intervention Model in 2012-2013	Item Aligned with SIG Application (Subtopic)
Reported receiving SIG funds in the following school years:			No
2010–2011	95.2	32.5	
2011–2012	87.2	35.6	
2012–2013	98.0[b]	24.4	
Reported being in a state that received an RTT grant			No
2011–2012	47.3	38.9	
Reported receiving RTT funds specifically for school improvement efforts in the following school years:			No
2010–2011[a,c]	24.7	11.5	
2011–2012[c]	29.1	16.6	
2012–2013	35.6	15.3	
Number of Schools	**250–260**	**160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Data from 2010–2011 were collected retrospectively in spring 2012. All data provided by schools were self-reported and not independently verified by the research team. For these reasons and potential concerns about recall accuracy, readers should exercise caution when interpreting data from 2010–2011.

[b] This number is less than 100 percent because we used several sources of information (other than the survey of school administrators) to identify the set of schools that implemented a SIG-funded intervention model in 2012–2013, and some schools that were identified as being in this group reported on the survey that they did not receive SIG funds.

[c] Schools that reported in spring 2012 that they were not in a state that received an RTT grant are included in the analysis of this question and are considered to not be using this practice.

Table E.24. District-reported school expenditures

	Schools implementing a SIG-funded intervention model in 2012–2013	Schools not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Mean total school expenditures in:			No
2009–2010[a]	$6,342,500	$7,569,226	
2011–2012	$6,875,879	$6,720,625	
2012–2013	$6,193,594	$6,554,238	
Mean percentage of school expenditures that go to wages, employee benefits, and other personnel expenditures in:			No
2009–2010[a]	86.4	83.9	
2011–2012	84.8	83.3	
2012–2013	84.4	87.9	
Number of Schools	**160**	**130**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

[a] Data from 2009–2010 were collected retrospectively in spring 2012. All data provided by districts were self-reported and not independently verified by the research team. For these reasons and potential concerns about recall accuracy, readers should exercise caution when interpreting data from 2009–2010.

Table E.25. School intervention models used in study schools

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-Funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported using one of the four SIG intervention models			No
2011–2012	94.0	23.1	
2012–2013	89.3	16.9	
Reported using the following SIG intervention models:[b]			
Turnaround model (replace the principal and rehire no more than 50 percent of staff; increase operational flexibility and learning time; make changes to the instructional program and professional development)			No
2011–2012	45.2	10.6	
2012–2013	39.6	6.9	
Transformation model (implement changes similar to those specified for the turnaround model, except [1] there are no limits on rehiring staff and [2] student growth must factor into teacher and principal evaluations)			No
2011–2012	44.8	10.0	
2012–2013	42.0	8.1	
Restart model (close the school and reopen under a charter or education management organization)			No
2011–2012	4.4	-[c]	
2012–2013	6.0	-[c]	
Closure model (close the school and send current students to higher-achieving schools in the district)			No
2011–2012	0.0	-[c]	
2012–2013	1.6	-[c]	
Restart or closure model			
2011–2012	-[c]	2.5	
2012–2013	-[c]	1.9	
Reported being a charter school			No
2011–2012	5.6	3.7	
2012–2013	5.6	3.1	
Number of Schools	**250**	**160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

[a] This number is less than 100 percent because we used several sources of information (other than the survey of school administrators) to identify the set of schools that implemented a SIG-funded Intervention Model in 2012–2013, and some schools that were identified as being in this group reported on the survey that they did not use one of the four SIG intervention models.

[b] Schools that responded no to the question in the prior row are included in the analysis of this question and are considered to not be using this practice.

[c] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for the restart and closure models for schools not implementing a SIG-funded intervention model.

Table E.26. Improvement strategies used in study schools

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported implementing changes to the following since July 2010/spring 2012:			
English language arts curriculum			No
2011–2012	68.7	63.3	
2012–2013	58.9	53.2	
Math curriculum			No
2011–2012	69.9	67.3	
2012–2013	61.0	60.9	
Instructional approaches in English language arts			No
2011–2012	92.3	77.1	
2012–2013	78.9	70.1	
Instructional approaches in math			No
2011–2012	91.7	78.6	
2012–2013	83.3	74.2	
Strategies to meet the needs of English language learners[a]			Yes (IS-3)
2011–2012	81.1	70.0	
2012–2013	70.9	61.5	
School administrative structure			No
2011–2012	83.5	56.9	
2012–2013	49.0	51.9	
Discipline policies			Yes (TC-2)
2011–2012	85.8	68.4	
2012–2013	65.9	57.6	
Nonacademic supports for students			Yes (TC-2)
2011–2012	81.3	65.8	
2012–2013	56.9	48.1	
Policies or strategies related to parent and/or community engagement			Yes (TC-2)
2011–2012	89.8	69.0	
2012–2013	66.8	58.9	
Policies around the use of data for instructional improvement			Yes (IS-2)
2011–2012	89.0	79.4	
2012–2013	71.5	64.4	
Monitoring of student readiness for grade promotion and/or high school graduation			Yes (IS-5)
2011–2012	75.9	70.3	
2012–2013	59.2	53.8	
Monitoring of students' college readiness[b]			Yes (IS-5)
2011–2012	90.1	87.8	
2012–2013	79.3	71.6	
Other changes			No
2011–2012	6.8	10.1	
2012–2013	2.7	5.4	
Number of Schools	**110–250**	**70–160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because missing data varied across items. For example, some questions shown in the table were only applicable to certain schools (such as high schools).

[a] Schools that reported that they did not have any English language learners are included in the analysis of this item and are considered to not be using this practice.

[b] The analysis for this row includes only high schools.

IS-2 = Promoting the continuous use of student data; IS-3 = Providing supports and professional development to staff to assist ELLs and students with disabilities; IS-5 = Tailoring strategies for secondary schools; TC-2 = Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs.

Table E.27. Instructional strategies used to meet the needs of English language learners

	Among Schools That Reported Having English Language Learners, Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	**Not implementing a SIG-funded intervention model in 2012–2013**	**Item aligned with SIG application (subtopic)**
Reported using the following strategies to meet the needs of these students:			
Used a curriculum that specifically addresses ELL needs			Yes (IS-3)
2011–2012	78.9	69.3	
2012–2013	75.8	67.5	
Implemented instructional strategies that specifically address ELL needs			Yes (IS-3)
2011–2012	87.3	84.1	
2012–2013	89.9	82.3	
Provided instructional programs specifically designed for ELLs			Yes (IS-3)
2011–2012	81.9	83.2	
2012–2013	72.5	75.2	
Provided specialized classes for ELLs			Yes (IS-3)
2011–2012	72.2	71.6	
2012–2013	71.5	74.3	
Provided additional services for ELLs			Yes (IS-3)
2011–2012	80.0	76.1	
2012–2013	81.3	76.1	
Provided professional development for teachers on providing instruction to ELLs			Yes (IS-3)
2011–2012	82.3	78.1	
2012–2013	84.2	80.7	
Used data on ELLs in school decision making			Yes (IS-3)
2011–2012	86.3	90.0	
2012–2013	86.9	90.0	
Other strategies			No
2011–2012	-[a]	-[a]	
2012–2013	-[a]	5.3	
Number of Schools	**130–160**	**100–110**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] This cell has been suppressed to protect respondent confidentiality.

ELLs = English language learners; IS-3 = Providing supports and professional development to staff to assist ELLs and students with disabilities.

Table E.28. District administrative supports for turnaround

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Reported having the following supports in place related to school turnaround:		
Staff explicitly designated to support school turnaround (but no designated turnaround office)		Yes (FS-2)
2009–2010[a]	44.0	
2011–2012	60.0	
2012–2013	56.0	
An office explicitly designated to support school turnaround (with designated staff)		Yes (FS-2)
2009–2010[a]	16.0	
2011–2012	36.0	
2012–2013	32.0	
Contracts with external consultants to support school turnaround		Yes (FS-2)
2009–2010[a]	55.9	
2011–2012	78.0	
2012–2013	67.8	
Other supports		No
2009–2010[a]	15.3	
2011–2012	27.1	
2012–2013	28.8	
Number of Districts	**50–60**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Data from 2009–2010 were collected retrospectively in spring 2012. All data provided by districts were self-reported and not independently verified by the research team. For these reasons and potential concerns about recall accuracy, readers should exercise caution when interpreting data from 2009–2010.

FS-2 = Receiving technical assistance and support.

Table E.29. Flexibility with or exemptions from collective bargaining agreements or staffing policies for SIG schools implementing one of the four SIG intervention models

	Percentage of Districts	Item Aligned with SIG Application (Subtopic)
Reported that SIG grantee schools had flexibility from the following aspects of collective bargaining agreements or policies that guide staffing in other district schools:		
Procedures for assigning or removing staff		Yes (FS-1)
2009–2010[a]	32.6	
2011–2012	63.0	
2012–2013	58.7	
Requirements or policies related to staff hours and responsibilities		Yes (FS-1)
2009–2010[a]	23.9	
2011–2012	67.4	
2012–2013	71.7	
Procedures related to the distribution of effective staff		Yes (FS-1)
2009–2010[a]	13.0	
2011–2012	41.3	
2012–2013	30.4	
Other types of flexibility or exemptions		No
2009–2010[a]	10.9	
2011–2012	34.8	
2012–2013	37.0	
Number of Districts	**50**	

Source: Interviews with district administrators in spring 2012 and spring 2013.

[a] Data from 2009–2010 were collected retrospectively in spring 2012. All data provided by districts were self-reported and not independently verified by the research team. For these reasons and potential concerns about recall accuracy, readers should exercise caution when interpreting data from 2009–2010.

FS-1 = Having operational flexibility.

Table E.30. School responsibility for decision making

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported having *primary* responsibility for making decisions in the following areas (rather than the state or district):			
Setting student discipline policies			Yes (FS-1)
2011–2012	38.7	32.9	
2012–2013	34.3	25.9	
Developing the school budget			Yes (FS-1)
2011–2012	53.9	54.4	
2012–2013	48.6	51.9	
Establishing the curriculum (including core texts)			Yes (FS-1)
2011–2012	16.7	15.4	
2012–2013	16.3	19.2	
Setting student assessment policies (on assessments other than state-mandated tests)			Yes (FS-1)
2011–2012	23.3	21.4	
2012–2013	31.8	28.3	
Staff hiring, discipline, and dismissal			Yes (FS-1)
2011–2012	38.5	41.5	
2012–2013	40.5	44.7	
Determining the length of the school day			Yes (FS-1)
2011–2012	17.2	12.0	
2012–2013	12.7	12.7	
Determining the length of the school year			Yes (FS-1)
2011–2012	5.7	3.1	
2012–2013	4.1	5.0	
Setting requirements for professional development			Yes (FS-1)
2011–2012	51.4	39.5	
2012–2013	44.0	34.4	
Number of Schools	**240–250**	**160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

FS-1 = Having operational flexibility.

Table E.31. Organization of instruction in schools

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported using the following methods to organize classes or other groups of students for instruction:			
Traditional grades or academic discipline-based departments			No
2011–2012	84.9	87.8	
2012–2013	84.1	87.8	
Grades or the school subdivided into small learning communities			Yes (IS-5)
2011–2012	58.1	53.8	
2012–2013	50.6	51.3	
Student groups that remain two or more years with the same teacher			No
2011–2012	29.7	32.5	
2012–2013	29.3	26.8	
Interdisciplinary teaching or paired/team teaching			No
2011–2012	58.6	48.1	
2012–2013	51.5	43.6	
Specialized classes for ELLs[a]			Yes (IS-3)
2011–2012	56.0	57.0	
2012–2013	55.2	58.9	
Block scheduling			Yes (TC-1)
2011–2012	47.9	56.8	
2012–2013	46.2	54.2	
Other methods			No
2011–2012	5.6	7.3	
2012–2013	5.6	6.0	
Number of Schools	**230–250**	**150–160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Schools that reported that they did not have any English language learners are included in the analysis of this item and are considered to not be using this practice.

ELLs = English language learners; IS-5 = Tailoring strategies for secondary schools; IS-3 = Providing supports and professional development to staff to assist ELLs and students with disabilities; TC-1 = Increasing learning time.

Table E.32. School instructional time

	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Mean number of instructional days that schools report being in session for students			Yes (TC-1)
2009–2010[a]	180.9	178.7	
2011–2012	181.5	178.9	
2012–2013	181.5	184.3	
Mean number of hours per day that schools report being in session for students			Yes (TC-1)
2009–2010[a]	6.8	6.9	
2011–2012	7.0	6.9	
2012–2013	7.0	7.0	
Mean number of minutes per week of instruction that schools report providing to the average student			
Mathematics			No
2011–2012	356.3	328.6	
2012–2013	347.6	341.2	
English language arts			No
2011–2012	370.8	353.3	
2012–2013	367.3	356.0	
Number of Schools	**180–230**	**110–150**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Data from 2009–2010 were collected retrospectively in spring 2012. All data provided by schools were self-reported and not independently verified by the research team. For these reasons and potential concerns about recall accuracy, readers should exercise caution when interpreting data from 2009–2010.

TC-1 = Increasing learning time.

Table E.33. School offerings outside the regular school day

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported using or offering the following:			Yes (TC-1)
Before- and/or after-school instruction			
2011–2012	90.3	77.7	
2012–2013	87.9	77.7	
Weekend instruction			
2011–2012	48.1	40.5	
2012–2013	43.8	43.8	
Summer instruction			
2011–2012	76.0	60.1	
2012–2013	73.6	59.5	
Number of Schools	**230–250**	**150–160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

TC-1 = Increasing learning time.

Table E.34. Common planning time

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported that all or some teachers have common planning time to meet in teams			Yes (TL-3)
2011–2012	94.3	93.6	
2012–2013	96.0	95.5	
Reported that all or some teachers have common planning time with the following frequency:[a]			
Daily			No
2011–2012	47.6	39.1	
2012–2013	46.4	49.1	
Several times per week			No
2011–2012	21.6	24.2	
2012–2013	26.0	22.4	
Once per week			No
2011–2012	22.8	26.7	
2012–2013	20.4	21.1	
Monthly, or a few times per year[b]			No
2011–2012	2.8	3.7	
2012–2013	3.2	3.1	
Reported that all or some teachers are *required* to participate in common planning time [a]			No
2011–2012	85.4	81.6	
2012–2013	84.1	85.0	
Number of Schools	**230–250**	**150–160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Schools that responded no to the question about whether teachers have common planning time to meet in teams are included in the analysis of this question and are considered to not be using this practice.

[b] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for "monthly" and "a few times per year."

TL-3 = Providing high-quality, job-embedded professional development or supports.

Table E.35. Frequency of use of benchmark tests in English language arts and math

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported that the typical English language arts teacher uses benchmark or interim assessments with the following frequency:			Yes (IS-2)
0–2 times per year[a]			
2011–2012	4.9	10.7	
2012–2013	6.5	13.8	
3–4 times per year			
2011–2012	55.5	52.8	
2012–2013	51.8	49.1	
5–6 times per year			
2011–2012	17.4	17.0	
2012–2013	18.6	18.2	
7–8 times per year			
2011–2012	6.5	7.5	
2012–2013	9.7	9.4	
More than 8 times per year			
2011–2012	15.8	11.9	
2012–2013	13.4	9.4	
Reported that the typical math teacher uses benchmark or interim assessments with the following frequency:			Yes (IS-2)
0–2 times per year[a]			
2011–2012	6.1	10.0	
2012–2013	5.7	10.0	
3–4 times per year			
2011–2012	51.6	50.0	
2012–2013	47.5	49.4	
5–6 times per year			
2011–2012	18.9	19.4	
2012–2013	17.6	18.1	
7–8 times per year			
2011–2012	7.8	8.1	
2012–2013	14.8	10.6	
More than 8 times per year			
2011–2012	15.6	12.5	
2012–2013	14.3	11.9	
Number of Schools	**240–250**	**160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] To comply with NCES statistical reporting requirements for small cell sizes, we aggregated the percentages for "zero times per year" and "one to two times per year."

IS-2 = Promoting the continuous use of student data.

Table E.36. Changes in staff implemented as part of school improvement efforts

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported getting a new principal			Yes (TL-2)
Between July 2010 and spring 2012	67.3	51.6	
Between spring 2012 and spring 2013	24.7	24.8	
Reported pursuing other significant leadership changes (aside from the principal)			No
Between July 2010 and spring 2012	51.8	31.0	
Between spring 2012 and spring 2013	30.1	24.1	
Reported having removed instructional staff through firing or counseling out			Yes (TL-2)
Between July 2010 and spring 2012			
Between spring 2012 and spring 2013	50.4	37.0	
Among schools that reported having removed instructional staff through firing or counseling out, average proportion of existing instructional staff that was removed			No
Between July 2010 and spring 2012	26.6	21.6	
Between spring 2012 and spring 2013	13.8	18.7	
Reported having hired a significant number of new staff (at least 50 percent of staff or more)			Yes (TL-2)
Between July 2010 and spring 2012	42.6	11.3	
Between spring 2012 and spring 2013	14.5	8.1	
Reported having reviewed the strengths and competencies of all existing instructional staff to assess the extent to which they were likely to be successful working in a school turnaround or improvement context			Yes (TL-2)
Between July 2010 and spring 2012	68.5	44.7	
Between spring 2012 and spring 2013	62.9	44.1	
Reported having assessed new hires for whether they possessed specific strengths or competencies deemed important to be successful working in a school turnaround or improvement context[a]			Yes (TL-2)
Between July 2010 and spring 2012	31.9	6.8	
Between spring 2012 and spring 2013	11.6	6.8	
Number of Schools	**90–250**	**30–160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because missing data varied across items. For example, some questions shown in the table were only applicable to certain schools (such as schools that reported having removed instructional staff through firing or counseling out).

[a] Schools that reported that they did not hire a significant number of new staff are included in the analysis of this item and are considered to not be using this practice.

TL-2 = Identifying and rewarding effective teachers and principals and removing ineffective ones.

Table E.37. School-reported training or technical assistance from the state or district

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported that the state and/or district provided the following types of training or technical assistance to the school since July 2010/spring 2012:			
Training or technical assistance on developing and implementing a school improvement plan			Yes (FS-2)
2011–2012	83.1	76.1	
2012–2013	77.0	73.5	
Training or technical assistance on identifying curricula, instructional strategies, or school reform models that have been shown to be effective at increasing student achievement			Yes (FS-2)
2011–2012	80.5	74.4	
2012–2013	73.9	64.7	
Training or technical assistance on identifying curricula, instructional strategies, or school reform models that have been shown to be effective at improving college readiness			Yes (FS-2)
2011–2012	65.0	60.6	
2012–2013	60.8	52.3	
Training or technical assistance on developing strategies to recruit and retain more effective teachers			Yes (FS-2)
2011–2012	45.5	37.7	
2012–2013	43.9	35.1	
Other assistance			No
2011–2012	5.9	4.6	
2012–2013	5.9	3.3	
Number of Schools	**240–250**	**150–160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

FS-2 = Receiving technical assistance and support.

Table E.38. Professional development for school instructional staff

	Percentage of Low-Performing Schools		
	Implementing a SIG-funded intervention model in 2012–2013	Not implementing a SIG-funded intervention model in 2012–2013	Item aligned with SIG application (subtopic)
Reported that their instructional staff received professional development on the following topics:			
Transitioning to the CCSS			Yes (TL-3)
2011–2012	70.0	71.5	
2012–2013	85.4	81.6	
Aligning instruction to state standards			Yes (TL-3)
2011–2012	85.1	78.1	
2012–2013	85.5	83.4	
Instructional strategies			Yes (TL-3)
2011–2012	93.6	82.6	
2012–2013	95.3	91.6	
Using data to improve and/or differentiate instruction			Yes (TL-3)
2011–2012	89.4	82.1	
2012–2013	92.3	83.4	
Meeting the needs of English language learners[a]			Yes (IS-3)
2011–2012	53.1	55.5	
2012–2013	51.9	49.7	
Strategies for turning around a low-performing school			Yes (TL-3)
2011–2012	74.4	43.4	
2012–2013	65.5	46.7	
Other topics			No
2011–2012	4.6	10.0	
2012–2013	9.7	9.3	
Reported that the following characteristics apply to at least half of the professional development activities provided to instructional staff:			
Single-session, one-time events			No
2011–2012	31.9	41.1	
2012–2013	28.8	40.4	
Multiple-session events			Yes (TL-3)
2011–2012	85.1	76.1	
2012–2013	83.5	78.7	
Involved practice in the classroom			Yes (TL-3)
2011–2012	83.5	79.6	
2012–2013	85.9	82.2	
Required for all instructional staff			No
2011–2012	93.5	96.8	
2012–2013	94.4	96.2	
Were designed with input from school staff			Yes (TL-3)
2011–2012	82.2	72.6	
2012–2013	78.9	75.2	
Number of Schools	**220–250**	**140–160**	

Source: Surveys of school administrators in spring 2012 and spring 2013.

Note: A range is provided for the sample sizes because nonresponse varied across items.

[a] Schools that reported that they did not have any English language learners are included in the analysis of this item and are considered to not be using this practice.

CCSS = Common Core State Standards; TL-3 = Providing high-quality, job-embedded professional development or supports; IS-3 = Providing supports and professional development to staff to assist ELLs and students with disabilities.

APPENDIX F

SURVEY QUESTIONS ALIGNED WITH SIG PRACTICES

This appendix provides crosswalks between practices aligned with the SIG application criteria, the school intervention model requirements, and the school administrator survey questions. The first table presents the required and permissible practices under the transformation and turnaround models that aligned with questions in the school administrator survey. (Restart model schools could choose to use any of these practices.) The next four tables—one for each of the four SIG topic areas addressed in Chapter IV—show the survey questions that aligned with the practices in that area. The school administrator survey protocol is available at http://www.mathematica-mpr.com/~/media/publications/pdfs/education/spring_2013_school_administrator_survey.pdf.

Table F.1. Required and permissible practices under the turnaround and transformation models

Required and Permissible Practices Under the Two Models	Required Under the SIG Transformation Model	Required Under the SIG Turnaround Model
Comprehensive instructional reform strategies		
Using data to identify and implement an instructional program		
Use data to evaluate instructional programs (for example, measuring program effectiveness)	X	X
Promoting the continuous use of student data		
Use data to inform and differentiate instruction	X	X
The typical English/language arts or math teacher used benchmark or interim assessments at least once per year	X	X
Providing supports and professional development to staff to assist ELLs and students with disabilities		
Implementing strategies (including additional supports or professional development) to ensure that students with limited English proficiency acquire language skills to master academic content		
Using and integrating technology-based supports		
Increased access to technology for teachers, or the typical English/language arts teacher used computer-assisted instruction		
Tailoring strategies for secondary schools		
Secondary school monitored students' college readiness (such as enrollment in Advanced Placement courses), including providing supports (such as project-based learning) so that low-achieving students can take advantage of these types of opportunities		
The school was divided, or grades within the secondary school were subdivided, into small learning communities or field/career-oriented academies		
Secondary school tracked student progress toward (and readiness for) high school graduation		
Teacher effectiveness		
Using rigorous, transparent, and equitable evaluation systems		
Student achievement growth was a required component of teacher evaluations, and the extent to which student achievement growth must factor into teacher evaluations, or state test scores were used to assess student growth for teacher evaluations was specified	X	
Using multiple performance measures for teacher evaluations	X	
Identifying and rewarding effective teachers and removing ineffective ones		
Use teacher evaluation results to inform decisions about compensation	X	
Reviewing the strengths and competencies of instructional staff for the purposes of hiring or removing staff		X

Required and Permissible Practices Under the Two Models	Required Under the SIG Transformation Model	Required Under the SIG Turnaround Model
Providing high-quality, job-embedded professional development or supports		
Providing instructional staff with PD that consisted mostly or entirely of multiple-session events	X	X
Providing instructional staff with PD focused on transitioning to Common Core State Standards, aligning instruction to state standards, or strategies for turning around a low-performing school	X	X
Providing staff with PD that involved educators working collaboratively or was facilitated by school leaders or coaches	X	X
Providing staff with PD that was focused on understanding and addressing student learning needs (including reviewing student work and achievement data, and collaboratively planning, testing, and adjusting instructional strategies based on data)	X	X
Providing staff with PD designed with input from school staff	X	X
Use data to evaluate the success of PD offerings		
Implementing strategies to recruit, place and retain staff		
Implementing strategies, such as financial incentives or more flexible work conditions, designed to recruit, place, and retain staff	X	X
Using teacher evaluation results as the primary consideration in reductions in force and excessing decisions, or having teacher assignment policies that allow for principal discretion in which staff to hire for the school		
Principal effectiveness		
Using rigorous, transparent, and equitable evaluation systems		
Measures of student achievement growth were used for principal evaluations and the extent to which student achievement growth must factor into principal evaluations was specified	X	
Using multiple performance measures for principal evaluations	X	
Identifying and rewarding effective principals and removing ineffective ones		
Using principal evaluation results to inform decisions about compensation	X	
School has a new principal	X	X
Providing high-quality, job-embedded professional development or supports		
State or district provides the principal or other school leaders with PD on analyzing and revising budgets or strategies for turning around a low-performing school	X	X
State or district provides the principal or other school leaders with PD on identifying effective instructional staff for leadership positions and supporting them in these positions	X	X
State or district uses principal evaluation results to develop the principal's PD or provides the principal with PD on aligning teachers' PD with evaluation results	X	X
Implementing strategies to recruit, place, and retain staff		
Provide financial incentives to recruit and retain effective principals	X	X
Increasing learning time		
Establish schedules and implement strategies to increase learning time	X	X
Engaging families and communities and providing a safe school environment that meets students' social, emotional, and health needs		
Changing policies or strategies related to parent or community engagement	X	X
State or district provided professional development on working with parents or creating a safe school environment		X
Changing discipline policies		
Guiding the development and implementation of, or making changes to, nonacademic supports or enrichment programs for students		X

Required and Permissible Practices Under the Two Models	Required Under the	
	SIG Trans-formation Model	SIG Turn-around Model
Having operational flexibility and receiving support		
School has primary responsibility for decisions on budget, hiring, discipline, or school year length	X	X
Receiving technical assistance and support		
State, district, or an external support provider sponsored by the state or district provided training or technical assistance to support school improvement efforts, or the school received support to help administrators and teachers use data to improve instruction	X	

Source: SIG application; surveys of school administrators in spring 2013.

Note: This table only lists the practices for which we had associated questions from the school administrator survey. An X in the SIG transformation model column means the practice in that row was required under the transformation model. When there is no X in that column, it means the practice in that row was permissible (but not required) under the transformation model. Similarly, an X in the SIG turnaround model column means the practice in that row was required under the turnaround model, and no X means that practice was permissible (but not required) under the turnaround model. Restart model schools could choose to use any of these practices.

Table F.2. Survey questions addressing the comprehensive instructional reform strategies topic area practices

Practice	Survey Questions Addressing the Comprehensive Instructional Reform Strategies Practice
Using data to evaluate instructional programs (for example, measuring program effectiveness)	DA1. During the current school year, for which of the following purposes has your school used data? a. To evaluate instructional programs (for example, measuring program effectiveness)
Using data to inform and differentiate instruction	DA1. During the current school year, for which of the following purposes has your school used data? b. To guide development and implementation of academic supports or enrichment programs (for example, identify how many and which students need academic support or enrichment, assign or reassign students to classes); d. To inform teachers' instructional practices (for example, identify areas for improvement, tailor instruction to meet student needs, manage instructional pacing); g. To track individual student performance and identify areas of improvement for specific students; k. To inform resource allocation to improve instruction (for example, which students participate in which programs, which staff work with which students) DA2. During the current school year, for which of the following purposes has your school used data on English language learners? e. To inform/improve/differentiate instruction for English language learners DA3. Within the past year, did any of the following activities related to data use occur in your school? If so, how often did they occur (daily, weekly, monthly, a few times per year, or once per year)? For item b below, if your school does not have English language learners, select —NA. a. District staff met with you and/or other school staff to review data on overall student performance; b. District staff met with you and/or other school staff specifically to review student performance data on English language learners; c. You or other school leaders reviewed student performance data to identify areas of improvement for the school; d. You or other school leaders met with teachers to discuss student performance data to identify areas in need of improvement for individual students or groups of students; h. After reviewing student performance data, teachers, administrators, and/or coaches formulated specific plans to update and revise instructional practice to address issues with specific students or specific classes. TA12. Since we last surveyed school administrators in spring 2012, did your school implement changes to any of the following? j. Policies around the use of data for instructional improvement
The typical English/language arts or math teacher used benchmark or interim assessments at least once per year	TA37. How often does the typical English language arts teacher in your school use benchmark or interim assessments? TA38. How often does the typical math teacher in your school use benchmark or interim assessments?

Practice	Survey Questions Addressing the Comprehensive Instructional Reform Strategies Practice
Implementing strategies (including additional supports or PD) to ensure that limited English proficient students acquire language skills to master academic content	TL29. During the current school year, have the state and/or district provided professional development or other support to the principal and/or other leaders of this school on any of the following topics? e. Ensuring that English language learners acquire the language skills needed to master academic content DA2. During the current school year, for which of the following purposes has your school used data on English language learners? If your school does not have English language learners, select —NA. a. To make decisions about students' entry into and/or exit from English language learner status; b. To place English language learners into specialized programs and/or classes; c. To track the progress of current English language learners; d. To track the progress of former English language learners; f. To identify professional development needs for teachers of English language learners; g. To assess teacher effectiveness with English language learners DA3. Within the past year, did any of the following activities related to data use occur in your school? If so, how often did they occur (daily, weekly, monthly, a few times per year, or once per year)? For item f below, if your school does not have English language learners, select —NA. f. School leaders coached teachers on the use of data specifically to improve instruction of English language learners. DA10.This school year, has your school received any of the following supports to help your school access and use data related to English language learners to improve and/or differentiate instruction for these students? For each type of support received, please describe the nature of the support received. If your school does not have English language learners, select —NA. a. Supports to help school staff use data to track the performance of English language learners (Please specify); b. Supports to help school staff use data to improve or differentiate instruction for English language learners (Please specify) TA12. Since we last surveyed school administrators in spring 2012, did your school implement changes to any of the following? For item e below, if your school does not have English language learners, select —NA. e. Strategies to meet the needs of English language learners TA22. Which of the following topics have been a focus of the professional development provided to instructional staff this school year? For item e below, if your school does not have English language learners, select —NA. e. Meeting the needs of English language learners TA32. Is your school currently using any of the following methods to organize classes or other groups of students for instruction? For item e below, if your school does not have English language learners, select —NA. e. Specialized classes for English language learners (such as newcomer class, English as a second language, sheltered content) TA36. Which of the following strategies/approaches does your school currently use to meet the needs of your school's English language learners? a. Use a curriculum that specifically addresses English language learners needs (Please specify); b. Implement instructional strategies that specifically address English language learners' needs, such as needs-based grouping, differentiated instruction, or increased progress testing of English language learners (Please specify); c. Provide instruction programs specifically designed for English language learners (such as English as a second language or bilingual programs) (Please specify); d. Provide specialized classes for English language learners (such as newcomer class, sheltered content class) (Please specify); e. Provide additional services for English language learners (such as tutors, bilingual aides, after-school program) (Please specify); f. Provide professional development for teachers on providing instruction to English language learners; g. Use data on English language learners in school decision making
Increased access to technology for teachers or that the typical English/language arts teacher used computer-assisted instruction	TA31. This school year, how often does the typical English language arts teacher in your school engage in the following activities? d. Use computer-assisted instruction TL28. Within the past year, has your school engaged in any of the following activities? f. Increased access to technology for teachers

Practice	Survey Questions Addressing the Comprehensive Instructional Reform Strategies Practice
Secondary school monitored students’ college readiness (such as enrollment in Advanced Placement courses), including providing supports (such as project-based learning) so that low-achieving students can take advantage of these types of opportunities	DA1. During the current school year, for which of the following purposes has your school used data? i. To track preparation for college enrollment (for example, participation in Advanced Placement courses or dual enrollment) TA12. Since we last surveyed school administrators in spring 2012, did your school implement changes to any of the following? l. Monitoring of students' college readiness (for example, participation in Advanced Placement courses, dual enrollment) TA31. This school year, how often does the typical English language arts teacher in your school engage in the following activities? a. Use project-based learning (for example, hands-on, inquiry-based activities) in classes; c. Use tiered interventions (for example, targeted/pull-out services for struggling students, intensive support to students who do not respond to interventions)
The school or grades within the secondary school were subdivided into small learning communities or field/career-oriented academies	TA32. Is your school currently using any of the following methods to organize classes or other groups of students for instruction? b. Grades or the school subdivided into small learning communities, such as "houses," "families," "teams," or field/career-oriented "academies" such as health or sciences
Secondary school tracked student progress towards (and readiness for) high school graduation	DA1. During the current school year, for which of the following purposes has your school used data? h. To track student progress toward high school graduation (for example, credits earned, required courses taken) TA12. Since we last surveyed school administrators in spring 2012, did your school implement changes to any of the following? k. Monitoring of student readiness for grade promotion and/or high school graduation

Source: Surveys of school administrators in spring 2013.

Note: DA indicates that the question came from the data systems module of the survey. TA indicates that the question came from the school turnaround module of the survey. TL indicates that the question came from the teachers and leaders module of the survey.

Table F.3. Survey questions addressing the teacher and principal effectiveness topic area practices

Practice	Survey Questions Addressing the Teacher and Principal Effectiveness Practice
Student achievement growth was a required component of teacher evaluations and the extent to which student achievement growth must factor into teacher evaluations or that state test scores were used to assess student growth for teacher evaluations was specified	TL2. Currently, to what extent does student growth evidence factor into the overall teacher evaluation? For example, student growth may be a "significant" factor in evaluations or have a specific weight (such as 20 percent) in the overall teacher evaluation. If this varies for different types of teachers, please describe this variation. TL3. Are any of the following measures used to assess student growth for teacher evaluations? a. State test scores; b. Scores on standardized assessments other than state tests; c. Some other measure of achievement (Please specify) [Note: TL1 (shown in the next row) was also used to address the practice in this row. Specifically, the practice in this row was coded as 0 if, among other things, the response to TL1 was “no teachers.”]
Using multiple performance measures for teacher evaluations	TL1. Currently, are measures of student growth a required component of teacher evaluations? TL8. Which of the following other measures of teacher performance are currently used by your school for teacher evaluations? If a particular measure is used only for some teachers, please specify the types of teachers for whom the measure is used. a. Classroom observations conducted by the principal; b. Classroom observations conducted by someone other than the principal (such as a peer or mentor teacher); c. Self-assessment; d. Peer assessments; e. Portfolios or other artifacts of teacher practice; f. Student work samples; g. Student surveys or other feedback; h. Parent surveys or other feedback
Using teacher evaluation results to inform decisions about compensation	TL14. Currently, do teacher evaluation results contribute to decisions about annual salary increases for teachers in your school? TL16. Currently, do teacher evaluation results contribute to the decision to provide bonuses or other performance-based compensation (other than annual salary increases) for teachers in your school?
Reviewing the strengths and competencies of instructional staff for the purposes of hiring or removing staff	TA16. Since spring 2012, did your school review the strengths and competencies of all existing instructional staff to assess the extent to which they were likely to be successful working in a school turnaround or improvement context? TA18. Since spring 2012, did your school remove instructional staff through firing or counseling out as part of school improvement efforts? TA20. Since spring 2012, did your school hire a significant number of new staff (at least 50 percent of staff or more) as part of school improvement efforts? TA21. Were these new hires assessed for whether they possessed specific strengths or competencies deemed important to be successful working in a school turnaround or improvement context?
Providing instructional staff with PD that consisted mostly or entirely of multiple-session events	TA23. How would you characterize the nature of the professional development activities provided to instructional staff in your school this year in terms of the following characteristics? For example, focusing on the first row below, would you say that all, most, roughly half, few, or none of the professional development provided to instructional staff this school year were single-session, one-time events? b. Multiple-session events
Providing instructional staff with PD that focused on transitioning to Common Core State Standards, aligning instruction to state standards, or strategies for turning around a low-performing school	TA22. Which of the following topics have been a focus of the professional development provided to instructional staff this school year? a. Transitioning to the Common Core State Standards; b. Aligning instruction to state standards; f. Strategies for turning around a low-performing school (Please specify)

Practice	Survey Questions Addressing the Teacher and Principal Effectiveness Practice
Providing staff with PD that involved educators working collaboratively or was facilitated by school leaders or coaches	TL27. Currently, does your school offer increased induction support (above and beyond that provided to all novice teachers in the district) for novice teachers in this school? TL28. Within the past year, has your school engaged in any of the following activities? a. Provided additional professional development, mentoring, and/or instructional coaching to teachers and/or school leaders (such as principals, assistant principals, or department heads); b. Improved opportunities for collaboration such as common planning time DA3. Within the past year, did any of the following activities related to data use occur in your school? If so, how often did they occur (daily, weekly, monthly, a few times per year, or once per year)? e. School leaders coached teachers on the use of data to improve instruction; g. Teachers met with each other to discuss data on their students/classes. DA6. Does your school provide scheduled time for teachers to examine data, either on their own or in collaboration with other teachers or school administrators? TA33. Currently, do all, some, or no teachers in your school have common planning time to meet in teams? If some (but not all) teachers have common planning time, please specify which teachers have common planning time. TA23. How would you characterize the nature of the professional development activities provided to instructional staff in your school this year in terms of the following characteristics? For example, focusing on the first row below, would you say that all, most, roughly half, few, or none of the professional development provided to instructional staff this school year were single-session, one-time events? c. Involved practice in the classroom
Providing staff with PD that was focused on understanding and addressing student learning needs (including reviewing student work and achievement data and collaboratively planning, testing, and adjusting instructional strategies based on data)	DA9. This school year, has your school received any professional development, training, or technical assistance to help school administrators and/or teachers access data, navigate data systems, or interpret and use data to improve and/or differentiate instruction? If so, please indicate the total number of hours of professional development, training, or technical assistance provided to school administrators and/or teachers this school year on these topics. DA1. During the current school year, for which of the following purposes has your school used data? e. To inform professional development offerings (for example, identify specific content or skills in which teachers need assistance or support) TL10. Currently, are teacher evaluation results used to guide decisions about what professional development and support is offered, recommended, or required for individual teachers in your school? TA22. Which of the following topics have been a focus of the professional development provided to instructional staff this school year? c. Instructional strategies (Please specify which instructional strategies were part of the professional development); d. Using data to improve and/or differentiate instruction (Please specify the specific strategies to improve and/or differentiate instruction that were part of the professional development)
Providing staff with PD designed with input from school staff	TA23. How would you characterize the nature of the professional development activities provided to instructional staff in your school this year in terms of the following characteristics? e. Were designed with input from school staff
Using data to evaluate the success of PD offerings	DA1. During the current school year, for which of the following purposes has your school used data? f. To evaluate the success of professional development offerings

Practice	Survey Questions Addressing the Teacher and Principal Effectiveness Practice
Implementing strategies, such as financial incentives or more flexible work conditions, that were designed to recruit, place, and retain staff	TL18. Currently, are teacher evaluation results used to guide decisions about career advancement for teachers in your school? TL26. Currently, do teachers and/or the principal at your school have the opportunity to receive any of the following financial incentives? a. Signing/recruitment bonuses for beginning to work in this school; b. Retention bonuses for continuing to work in the school; c. Performance bonuses; d. Increased annual compensation other than bonuses; e. Loan forgiveness; f. Tuition reimbursement; g. Housing; h. Financial incentives targeted towards increasing the number of staff with English language learner expertise in the school TL28. Within the past year, has your school engaged in any of the following activities? g. Offered more flexible work conditions (for example, more flexible schedule); h. Increased use of aides/paraprofessionals
Using teacher evaluation results as the primary consideration in reductions in force and excessing decisions or having teacher assignment policies that allow for principal discretion to decide which staff to hire for the school	TL13. Currently, are teacher evaluation results, rather than seniority, the primary consideration in reductions in force and excessing decisions for your school (if your school were to reduce the size of its faculty)? TL25. Do current teacher-assignment policies for your school allow for principal discretion or authority to decide which staff to hire for your school? If yes, please describe the discretion or authority available to your school's principal when making hiring decisions.
Measures of student achievement growth were used for principal evaluations and the extent to which student achievement growth must factor into principal evaluations was specified	TL20. Currently, to what extent does student growth factor into the overall principal evaluation? For example, student growth may be a "significant" factor in evaluations or have a specific weight (such as 20 percent) in the overall principal evaluation. [Note: TL19a (shown in the next row) was also used to address the practice in this row. Specifically, the practice in this row was coded as 0 if, among other things, the response to TL19a was "no."]
Using multiple performance measures for principal evaluations	TL19. Currently, which of the following measures are used to evaluate the performance of your school's principal? a. Student growth measures; b. Self-assessment; c. District administrator input; d. School staff surveys or other feedback; e. Student surveys or other feedback
Principal evaluation results were used to inform decisions about compensation	TL22. Currently, do principal evaluation results contribute to decisions about annual salary increases for the principal of your school? TL24. Currently, do principal evaluation results contribute to the decision to provide bonuses or performance-based compensation to the principal of your school?
School has a new principal	TA14. Did your school get a new principal since we last surveyed school administrators in spring 2012?
State or district provides the principal or other school leaders with PD on analyzing and revising budgets or strategies for turning around a low-performing school	TL29. During the current school year, have the state and/or district provided professional development or other support to the principal and/or other leaders of this school on any of the following topics? f. Analyzing and revising budgets to use resources effectively; g. Strategies for turning around a low-performing school
State or district provides the principal or other school leaders with PD on identifying effective instructional staff for leadership positions and supporting them in these positions	TL29. During the current school year, have the state and/or district provided professional development or other support to the principal and/or other leaders of this school on any of the following topics? b. Identifying effective instructional staff for leadership positions and supporting them in such positions

Practice	Survey Questions Addressing the Teacher and Principal Effectiveness Practice
State or district uses principal evaluation results to develop the principal's PD or provides the principal with PD on aligning teachers' PD with evaluation results	TL29. During the current school year, have the state and/or district provided professional development or other support to the principal and/or other leaders of this school on any of the following topics? a. Aligning professional development with teacher evaluation results TL21. Currently, are principal evaluation results used to develop professional development and/or support plans specifically for the principal of your school?
Principals have the opportunity to receive financial incentives designed to recruit, place, and retain staff	TL26. Currently, do teachers and/or the principal at your school have the opportunity to receive any of the following financial incentives? a. Signing/recruitment bonuses for beginning to work in this school; b. Retention bonuses for continuing to work in the school; c. Performance bonuses; d. Increased annual compensation other than bonuses; e. Loan forgiveness; f. Tuition reimbursement; g. Housing; h. Financial incentives targeted towards increasing the number of staff with English language learner expertise in the school

Source: Surveys of school administrators in spring 2013.

Note: DA indicates that the question came from the data systems module of the survey. TA indicates that the question came from the school turnaround module of the survey. TL indicates that the question came from the teachers and leaders module of the survey.

Table F.4. Survey questions addressing the learning time and community-oriented schools topic area practices

Practice	Survey Questions Addressing the Learning Time and Community-Oriented Schools Practice
Using schedules and strategies that provide increased learning time or increasing the number of hours per year that school was in session	TA24. Does your school schedule currently use or offer any of the following? a. block scheduling; b. Before- and/or after-school instruction; c. Weekend instruction; d. Summer instruction TA27. In the current school year, how many hours per day is your school in session for students? If the number of hours per day that your school is in session varies by day of the week, please record the number of hours per day that your school is in session for each day of the week in the box below. TA29. In the current school year, how many days per year is your school in session for students?
Changing policies or strategies related to parent or community engagement	TA12. Since we last surveyed school administrators in spring 2012, did your school implement changes to any of the following? i. Policies or strategies related to parent and/or community engagement
State or district provided professional development on working with parents or creating a safe school environment	TL29. During the current school year, have the state and/or district provided professional development or other support to the principal and/or other leaders of this school on any of the following topics? c. Working with parents; d. Integrating cultural sensitivity into the school environment TL28. Within the past year, has your school engaged in any of the following activities? e. Enhanced safety measures in the building; i. Increased use of volunteers (for example, parents)
Changing discipline policies	TA12. Since we last surveyed school administrators in spring 2012, did your school implement changes to any of the following? g. Discipline policies
Guiding the development and implementation of, or making changes to, nonacademic supports or enrichment programs for students	DA1. During the current school year, for which of the following purposes has your school used data? c. To guide development and implementation of nonacademic supports or enrichment programs (for example, identify how many and which students need counseling) TA12. Since we last surveyed school administrators in spring 2012, did your school implement changes to any of the following? h. Nonacademic supports (for example, mental health supports) for students

Source: Surveys of school administrators in spring 2013.

Note: DA indicates that the question came from the data systems module of the survey. TA indicates that the question came from the school turnaround module of the survey. TL indicates that the question came from the teachers and leaders module of the survey.

Table F.5. Survey questions addressing the operational flexibility and support topic area practices

Practice	Survey Questions Addressing the Operational Flexibility and Support Practice
School has primary responsibility for budget, hiring, discipline, or school year length decisions	TA40. Currently, does your school, the district, or the state have primary responsibility for decisions in each of the following areas for your school? a. Setting student discipline policies; b. Developing the school budget; c. Establishing the curriculum (including core texts); d. Setting student assessment policies (on assessments other than state-mandated tests); e. Staff hiring, discipline, and dismissal; f. Determining the length of the school day; g. Determining the length of the school year; h. Setting requirements for professional development
State, district, or an external support provider sponsored by the state or district provided training or technical assistance to support school improvement efforts or that the school received support to help administrators and teachers use data to improve instruction	DA8. This school year, has your school received any of the following types of support to help school administrators and/or teachers access and use data to improve and/or differentiate instruction? For each type of support received, please specify the nature of the support that your school received. For example, if funding was received, please specify how much funding and the purposes for which the funds were used (for example, to buy hardware or software, to develop or improve data systems, or to provide training to teachers on the analysis and use of data). a. Funds to support school investments related to data use; for example, funds to buy hardware or software, to develop or improve data systems, or to provide training to teachers on the analysis and use of data (Please specify); b. Hardware or software to facilitate data use (Please specify); c. Materials on how to access and use data to differentiate or improve instruction (Please specify)
	TA39. Since spring 2012, have the state and/or district provided any of the following types of training or technical assistance to your school? Please include assistance provided directly by state or district staff as well as assistance funded by the state or district but provided by someone other than state or district staff, for example, external consultants or staff from a regional office. a. Training or technical assistance on developing and implementing a school improvement plan; b. Training or technical assistance on identifying curricula, instructional strategies, or school reform models that have been shown to be effective at increasing student achievement; c. Training or technical assistance on identifying curricula, instructional strategies, or school reform models that have been shown to be effective at improving college readiness; d. Training or technical assistance on developing strategies to recruit and retain more effective teachers
	TA41. Does your school currently have a state- or district-sponsored external support provider(s) or consultant(s) that regularly provides technical assistance to your school administrators or instructional staff around school improvement efforts?

Source: Surveys of school administrators in spring 2013.

Note: DA indicates that the question came from the data systems module of the survey. TA indicates that the question came from the school turnaround module of the survey.

APPENDIX G

ADDITIONAL INFORMATION ABOUT ENGLISH LANGUAGE LEARNER-FOCUSED IMPLEMENTATION ANALYSES

This appendix contains additional information that is directly related to the English language learner (ELL)-focused analyses presented in Chapter VII. Section A of this appendix lists the 2013 school administrator survey questions that addressed the ELL-focused practices aligned with the School Improvement Grant (SIG) application criteria (Table G.1).

Section B of this appendix presents findings from an analysis of the extent to which *district* administrators reported using ELL-focused practices promoted by SIG in spring 2013. These findings shed light on the extent to which districts reported providing support to schools for ELL-focused practices. We first present the ELL-focused practices aligned with the SIG application criteria and the district interview questions that addressed them (Table G.2). For readers interested in districts' reported use of an individual ELL-focused practice listed in Table G.2, we present the extent to which districts reported using the individual ELL-focused practices aligned with the SIG application criteria (Figure G.1). We then present a series of figures that display the results. We present an analysis of districts' overall use of ELL-focused practices aligned with the SIG application criteria (Figure G.2) and findings on reported use of ELL-focused practices by districts with above- or below-median ELL populations (Figure G.3) and above- or below-median ELL/non-ELL achievement gaps (Figure G.4).

One important difference between the figures shown in Chapter VII of the report and Section B of this appendix is that the latter have no comparison group. All districts in the study sample included schools that were and were not implementing a SIG-funded intervention model. Therefore, in Section B, we are not presenting comparisons between districts; rather, we are presenting descriptive information about the ELL-focused practices that study districts reported using.

A. School survey questions addressing the ELL-focused practices

Table G.1. School survey questions addressing the ELL-focused practices

ELL-Focused Practice	Survey Questions Addressing the ELL-Focused Practices
Teachers have the opportunity to receive financial incentives designed to increase the number of staff with ELL expertise	TL26. Currently, do teachers and/or the principal at your school have the opportunity to receive any of the following financial incentives? h. Financial incentives targeted towards increasing the number of staff with ELL expertise in the school
Principals have the opportunity to receive financial incentives designed to increase the number of staff with ELL expertise	TL26. Currently, do teachers and/or the principal at your school have the opportunity to receive any of the following financial incentives? h. Financial incentives targeted towards increasing the number of staff with ELL expertise in the school
Using data on ELLs to inform and differentiate instruction	DA2. During the current school year, for which of the following purposes has your school used data on ELLs? If your school does not have ELLs, select "NA." a. To make decisions about students' entry into or exit from ELL status; b. To place ELLs into specialized programs and/or classes; c. To track the progress of current ELLs; d. To track the progress of former ELLs; e. To inform/improve/differentiate instruction for ELLs

ELL-Focused Practice	Survey Questions Addressing the ELL-Focused Practices
	DA3. Within the past year, did any of the following activities related to data use occur in your school? If so, how often did they occur (daily, weekly, monthly, a few times per year, or once per year)? For item b below, if your school does not have ELLs, select “NA.” b. District staff met with you and/or other school staff specifically to review student performance data on ELLs
	TA36. Which of the following strategies/approaches does your school currently use to meet the needs of your school’s ELLs? If your school does not have ELLs, select “NA.” g. Use data on ELLs in school decision making
Implementing strategies, supports, or professional development to meet the needs of ELLs	TA12. Since we last surveyed school administrators in spring 2012, did your school implement changes to any of the following? For item e below, if your school does not have ELLs, select “NA.” e. Strategies to meet the needs of ELLs TA22. Which of the following topics have been a focus of the professional development provided to instructional staff this school year? For item e below, if your school does not have ELLs, select “NA.” e. Meeting the needs of ELLs TA36. Which of the following strategies/approaches does your school currently use to meet the needs of your school’s ELLs? If your school does not have ELLs, select “NA.” a. Use a curriculum that specifically addresses ELLs needs (Please specify); b. Implement instructional strategies that specifically address ELLs instruction, or increased progress testing of ELLs (Please specify); c. Provide instruction programs specifically designed for ELLs (such English as a second language or bilingual programs) (Please specify); d. Provide specialized classes for ELLs (such as newcomer class, sheltered content class) (Please specify); f. Provide professional development for teachers on providing instruction to ELLs DA2. During the current school year, for which of the following purposes has your school used data on ELLs? If your school does not have ELLs, select “NA.” f. To identify professional development needs for teachers of ELLs; g. To assess teacher effectiveness with ELLs DA3. Within the past year, did any of the following activities related to data use occur in your school? If so, how often did they occur (daily, weekly, monthly, a few times per year, or once per year)? For item f below, if your school does not have ELLs, select “NA.” f. School leaders coached teachers on the use of data specifically to improve instruction on ELLs TL29. During the current school year, have the state and/or district provided professional development or other support to the principal and/or other leaders of this school on any of the following topics? For item e below, if your school does not have ELLs, select “NA.” e. Ensuring that ELLs acquire the language skills needed to master academic content TA32. Is your school currently using any of the following methods to organize classes or other groups of students for instruction? For item e below, if your school does not have ELLs, select “NA.” e. Specialized classes for ELLs (such as newcomer class, English as a second language, sheltered content).
Providing additional services for ELLs (such as tutors, bilingual aides, or an after-school program)	TA36. Which of the following strategies/approaches does your school currently use to meet the needs of your school’s ELLs? If your school does not have ELLs, select “NA.” e. Provide additional services for ELLs (such as tutors, bilingual aides, after-school program) (Please specify)
Receiving supports from the state or local education agency to use data on ELLs to improve or differentiate instruction	DA10. This school year, has your school received any of the following supports to help your school access and use data related to ELLs to improve and/or differentiate instruction for these students? For each type of support received, please describe the nature of the support received. If your school does not have ELLs, select “NA.” a. Support to help school staff use data to track the performance of ELLs (Please specify); b. Supports to help school staff use data to improve or differentiate for ELLs (Please specify)

Source: Surveys of school administrators in spring 2013.

Note: DA indicates that the question came from the data systems module of the survey. TA indicates that the question came from the school turnaround module of the survey. TL indicates that the question came from the teachers and leaders module of the survey.

ELL = English language learner; NA = not applicable.

B. Analysis of districts' reported use of ELL-focused practices aligned with SIG application criteria

Table G.2. District interview questions addressing the ELL-focused practices

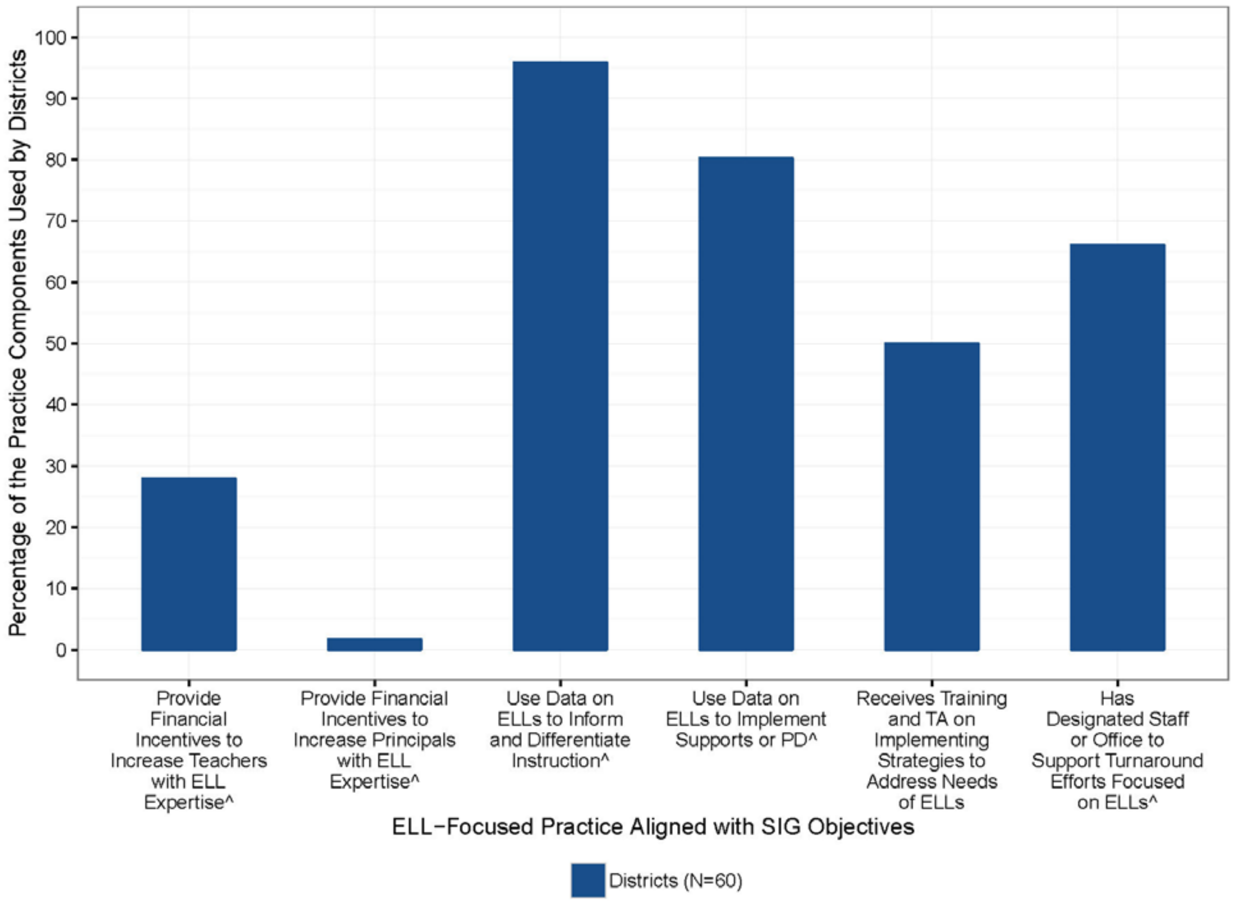
ELL-Focused Practice	District Interview Questions Addressing the ELL-Focused Practices
Teachers have the opportunity to receive financial incentives designed to increase the number of staff with ELL expertise	TL29a. Currently, which of the following types of financial incentives are offered by your district to teachers working in SIG grantee schools that are implementing one of the four intervention models specified by the U.S. Department of Education? h. Financial incentives targeted toward increasing the number of staff with ELL expertise in these schools TL30. Does your district currently use any of the following other strategies to help recruit and retain effective teachers and/or principals in SIG grantee schools implementing one of the four intervention models? c. Retention or recruitment efforts targets toward increasing the number of staff with ELL expertise in these schools
Principals have the opportunity to receive financial incentives designed to increase the number of staff with ELL expertise	TL29b. Currently, which of the following types of financial incentives are offered by your district to principals working in SIG grantee schools that are implementing one of the four intervention models specified by the U.S. Department of Education? h. Financial incentives targeted toward increasing the number of staff with ELL expertise in these schools
Using data on ELLs to inform and differentiate instruction	DA6. For which of the following purposes do district staff currently use data specifically on ELLs from either the state longitudinal data system or a district data system? a. To make decisions about students' entry into and/or exit from ELL status; b. To place ELLs into specialized programs and/or classes; c. To track the progress of current ELLs; d. To track the progress of former ELLs; e. To inform/improve/differentiate instruction for ELLs
Using data on ELLs to implement supports or professional development	DA6. For which of the following purposes do district staff currently use data specifically on ELLs from either the state longitudinal data system or a district data system? f. To identify professional development needs for teachers of ELLs; g. To assess teacher effectiveness with ELLs TA26. For which groups does the district provide this additional district-wide support and programs? a. ELLs
Receiving training and technical assistance on identifying and implementing strategies to address the needs of ELLs	TA42. This school year, which of the following types of training and/or technical assistance has the state provided to your district to support the improvement efforts of the persistently lowest-achieving schools in the district? Please report technical assistance provided directly by state staff as well as technical assistance funded by the state but provided by someone other than state staff, for example, an external consultant or staff from a regional office. g. Training or technical assistance on identifying and implementing strategies to address the needs of ELLs
Have designated staff or a designated office to support turnaround efforts focused on ELLs	TA9. Currently, does the district have any of the following organizational or administrative structures in place that are specifically intended to support school turnaround efforts focused on ELLs? a. District has explicitly designated staff to support school turnaround efforts focused on ELLs (but no designated office); b. District has an office explicitly designated to support school turnaround efforts focused on ELLs (with designated staff)

Source: Interviews with district administrators in spring 2013.

Note: DA indicates that the question came from the data systems module of the survey. TA indicates that the question came from the school turnaround module of the survey. TL indicates that the question came from the teachers and leaders module of the survey.

ELL = English language learner.

Figure G.1. District use of individual ELL-focused practices aligned with SIG objectives



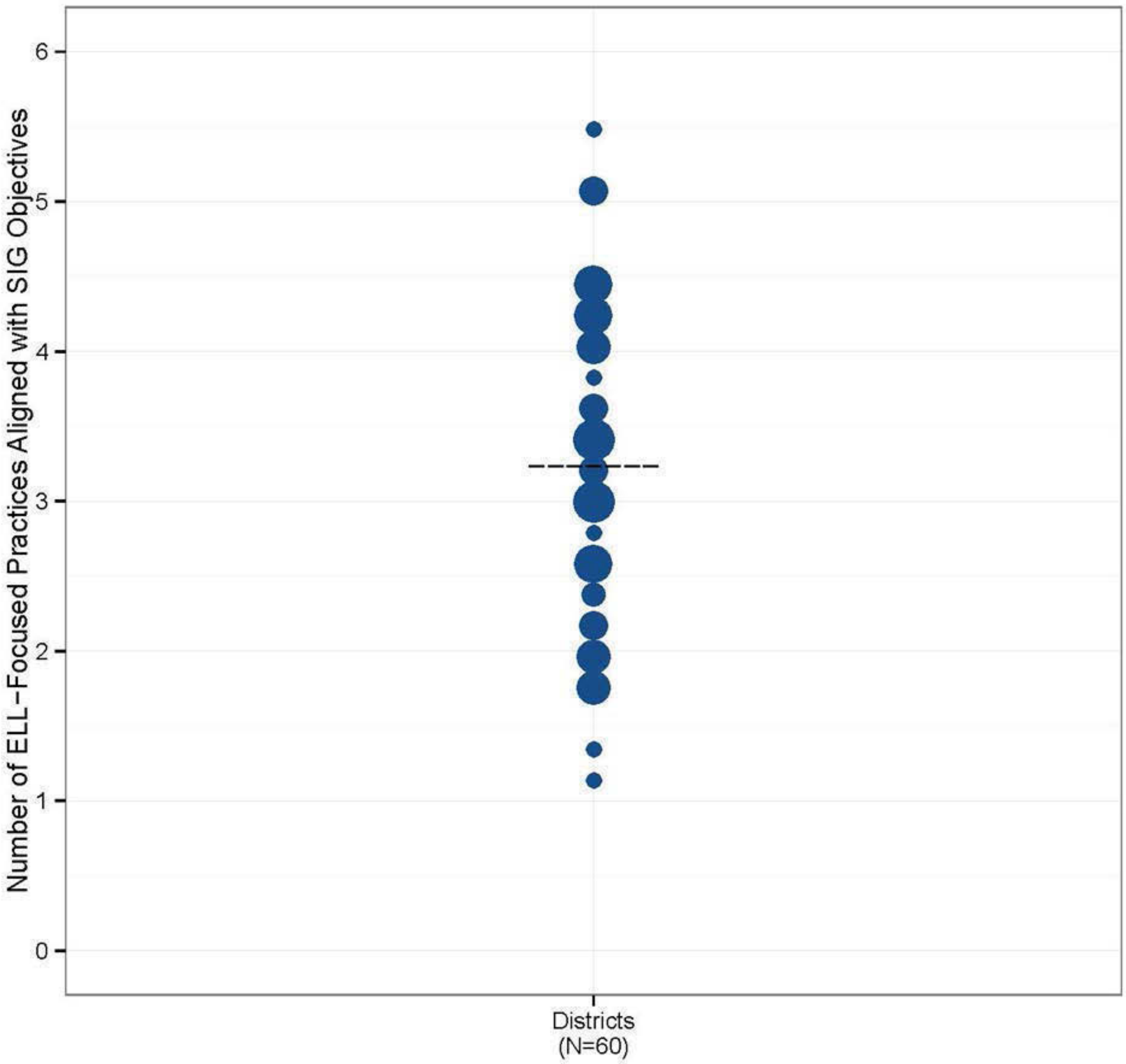
Source: Interviews with district administrators in spring 2013.

Note: As described in Chapter II, we selected interview questions that addressed the ELL-focused practices aligned with the SIG application selection criteria. The practices shown on the horizontal axis of this figure are listed in Table G.2. As described in Appendix C, for each ELL-focused practice in the SIG application criteria for which we identified one or more interview questions that addressed the practice, we calculated the percentage of interview questions with a “yes” response as a measure of the percentage of components each district used. The height of each bar represents the average percentage of the components of the ELL-focused practice that each group of districts used.

^Multiple district interview questions were used to assess whether districts used all of the components of this practice.

ELL = English language learner; PD = professional development; TA = technical assistance.

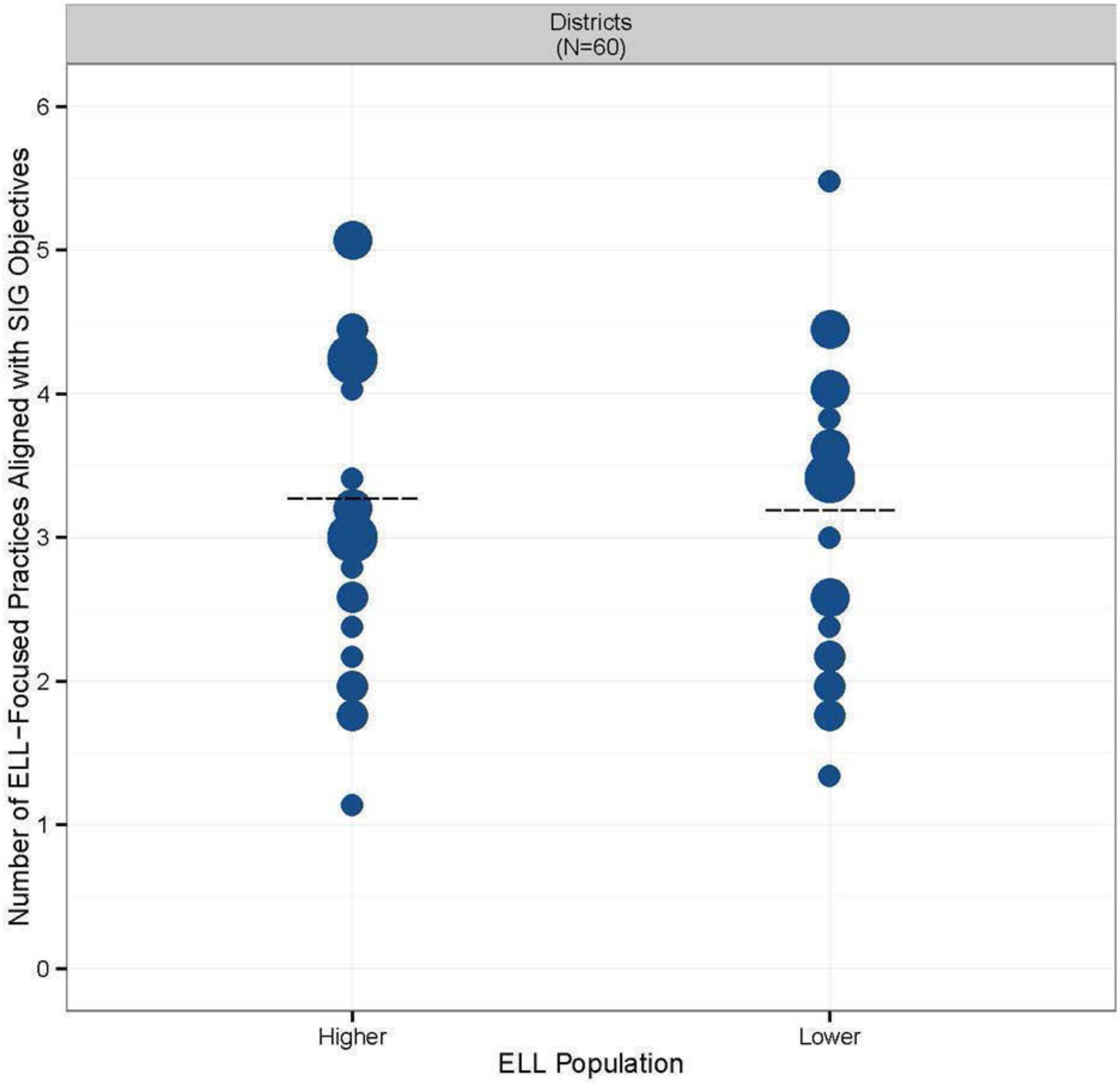
Figure G.2. District use of ELL-focused practices aligned with SIG objectives



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table G.2. Each dot in this figure represents the number of districts that reported using a particular number of ELL-focused practices (out of six examined) that were aligned with the SIG application criteria. Each dot in this figure represents less than 10 districts, so the numbers inside the dots have been removed to protect respondent confidentiality. For three of the ELL-focused practices, a “yes” response received one point. In the other three cases, it was possible for a district to receive a fraction of one point. See Appendix C for details on how we determined the number of ELL-focused practices for each district. The dashed line denotes the average number of ELL-focused practices across all districts.

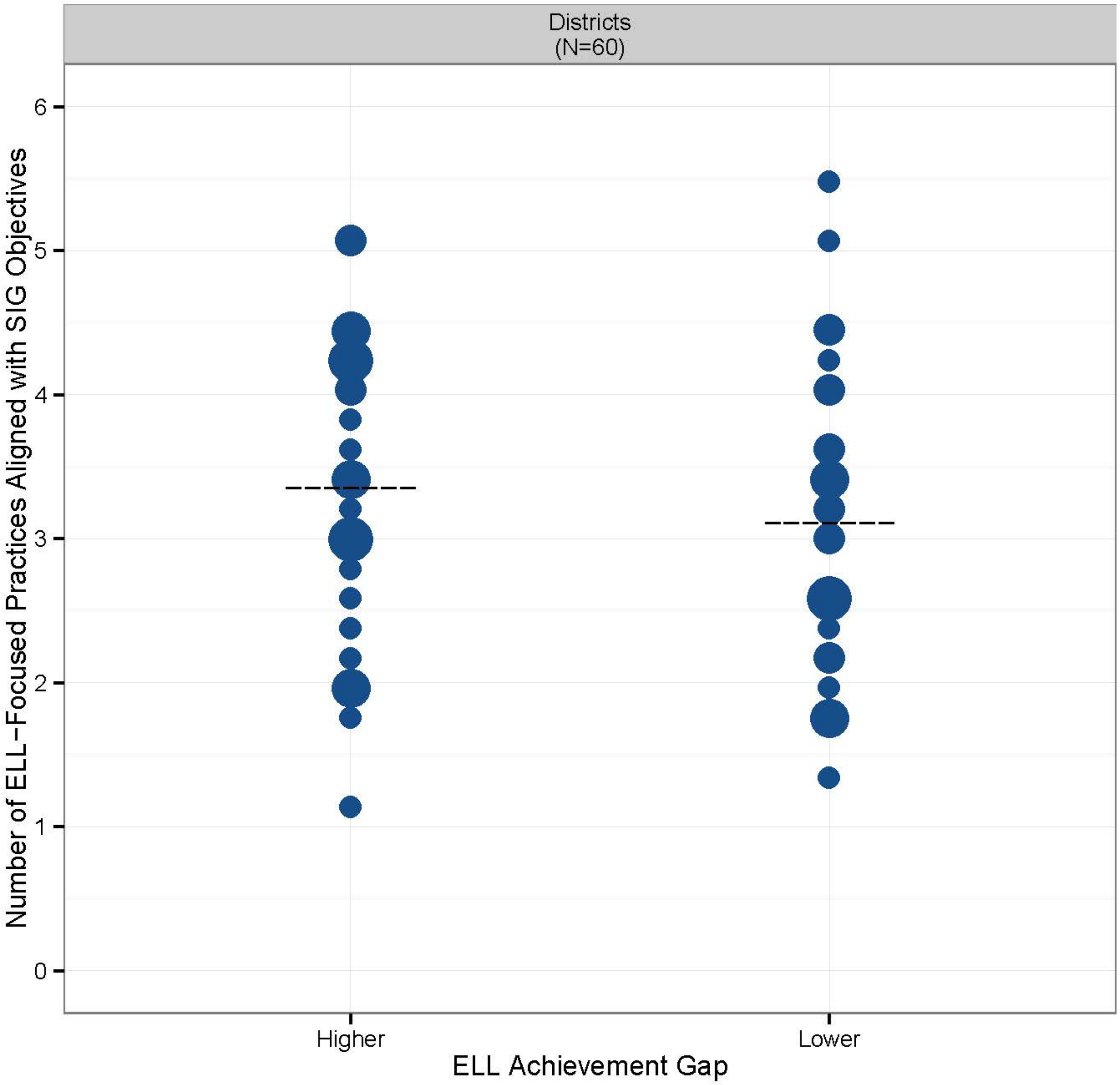
Figure G.3. District use of ELL-focused practices aligned with SIG objectives, by ELL population



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table G.2. The figure shows the number of ELL-focused practices that districts reported using, by districts that had above-median (higher) and below-median (lower) ELL populations. Each dot in this figure represents the number of districts that reported using a particular number of ELL-focused practices (out of six examined) that were aligned with the SIG application criteria. Each dot in this figure represents less than 10 districts, so the numbers inside the dots have been removed to protect respondent confidentiality. For three of the ELL-focused practices, a "yes" response received one point. In the other three cases, it was possible for a district to receive a fraction of one point. See Appendix C for details on how we determined the number of ELL-focused practices for each district. The dashed line denotes the average number of ELL-focused practices across all districts. There were no statistically significant differences between districts with higher and lower ELL populations at the 0.05 level using a two-tailed test.

Figure G.4. District use of ELL-focused practices aligned with SIG objectives, by ELL achievement gap



Source: Interviews with district administrators in spring 2013.

Note: The practices summarized in this figure are presented in Table G.2. The figure shows the number of ELL-focused practices that districts reported using, by districts that had above-median (higher) and below-median (lower) achievement gaps between ELL and non-ELLs. Each dot in this figure represents the number of districts that reported using a particular number of ELL-focused practices (out of six examined) that were aligned with the SIG application criteria. Each dot in this figure represents less than 10 districts, so the numbers inside the dots have been removed to protect respondent confidentiality. For three of the ELL-focused practices, a “yes” response received one point. In the other three cases, it was possible for a school to receive a fraction of one point. See Appendix C for details on how we determined the number of ELL-focused practices for each district. The dashed line denotes the average number of ELL-focused practices across all districts. There were no statistically significant differences between districts with higher and lower ELL/non-ELL achievement gaps at the 0.05 level using a two-tailed test.