USING INTEGRATED STUDENT SUPPORTS TO KEEP KIDS IN SCHOOL

A QUASI-EXPERIMENTAL EVALUATION OF COMMUNITIES IN SCHOOLS

Marie-Andrée Somers
Zeest Haider

APRIL 2017

mdrc
BUILDING KNOWLEDGE
TO IMPROVE SOCIAL POLICY

Using Integrated Student Supports to Keep Kids in School

A Quasi-Experimental Evaluation of Communities In Schools

Marie-Andrée Somers
Zeest Haider

April 2017

mdrc
BUILDING KNOWLEDGE TO IMPROVE SOCIAL POLICY

SOCIAL INNOVATION FUND

This report is funded by the Edna McConnell Clark Foundation's Social Innovation Fund, a White House initiative and program of the Corporation for National and Community Service (CNCS). The Edna McConnell Clark Foundation's Social Innovation Fund includes support from CNCS and 15 private coinvestors: The Edna McConnell Clark Foundation, The Annie E. Casey Foundation, The Duke Endowment, The William and Flora Hewlett Foundation, The JPB Foundation, George Kaiser Family Foundation, The Kresge Foundation, Open Society Foundations, The Penzance Foundation, The Samberg Family Foundation, The Charles and Lynn Schusterman Family Foundation, The Starr Foundation, Tipping Point Community, The Wallace Foundation, and the Weingart Foundation.

Dissemination of MDRC publications is supported by the following funders that help finance MDRC's public policy outreach and expanding efforts to communicate the results and implications of our work to policymakers, practitioners, and others: The Annie E. Casey Foundation, Charles and Lynn Schusterman Family Foundation, The Edna McConnell Clark Foundation, Ford Foundation, The George Gund Foundation, Daniel and Corinne Goldman, The Harry and Jeanette Weinberg Foundation, Inc., The JBP Foundation, The Joyce Foundation, The Kresge Foundation, Laura and John Arnold Foundation, Sandler Foundation, and The Starr Foundation.

In addition, earnings from the MDRC Endowment help sustain our dissemination efforts. Contributors to the MDRC Endowment include Alcoa Foundation, The Ambrose Monell Foundation, Anheuser-Busch Foundation, Bristol-Myers Squibb Foundation, Charles Stewart Mott Foundation, Ford Foundation, The George Gund Foundation, The Grable Foundation, The Lizabeth and Frank Newman Charitable Foundation, The New York Times Company Foundation, Jan Nicholson, Paul H. O'Neill Charitable Foundation, John S. Reed, Sandler Foundation, and The Stupski Family Fund, as well as other individual contributors.

For information about MDRC and copies of our publications, see our website: www.mdrc.org.

Overview

The Communities In Schools (CIS) Model of Integrated Student Supports aims to reduce dropout rates by providing students with integrated and tiered support services based on their levels of need. The model includes preventive services that are available to all students (Level 1 services) as well as intensive, targeted, and sustained services provided through case management (Level 2 services) for the 5 percent to 10 percent of students who display significant risk factors for dropping out, such as poor academic performance, high absentee rates, or behavioral problems. The CIS model posits that these tiered, integrated services will give students the skills and resources they need to succeed, which will lead to improvements in their outcomes.

In elementary schools, the CIS model focuses on improving attendance rates by engaging parents. In middle schools, the model begins to emphasize helping students improve their behavior. In high schools, the model focuses on services specifically intended to prevent students from dropping out, to help them progress through school, and to make sure they graduate.

This study, which is based on a quasi-experimental research design, examines the CIS model's effect on students' outcomes in elementary schools, middle schools, and high schools. The sample for this study includes 53 CIS schools in Texas and North Carolina (14 high schools, 15 middle schools, and 24 elementary schools) that started implementing the CIS model between 2005 and 2008. The study compares these CIS schools with 78 matched comparison schools (18 high schools, 24 middle schools, and 36 elementary schools). It is funded by the Edna McConnell Clark Foundation's Social Innovation Fund.

For the high schools, the main finding is that on-time graduation rates increased — and dropout rates decreased — in the study schools after the CIS model was launched. Graduation and dropout rates also improved in the comparison schools, so it is unclear whether the CIS model was *more* effective than the strategies used by the comparison schools. The findings do suggest that the CIS model may be at least as effective as these other approaches. In elementary schools, attendance rates (a central outcome the CIS model aims to effect in the elementary grades) improved in schools implementing the CIS model more than they did in a group of similar, comparison schools. There was no effect on attendance in middle and high schools. In middle schools, English/language arts test scores did not improve in schools implementing the CIS model, whereas they did improve in a group of similar, comparison middle schools. There was no effect on test scores in elementary and high schools. (It is important to note, however, that the CIS model does not attempt to improve state test scores or the quality of instruction provided during regular school hours.) Unfortunately, it was not possible to evaluate whether the CIS model improved middle school students' behavioral outcomes, which is the model's primary goal in those grades.

Contents

List of Exhibits

Table

Figure

Box

Preface

Although many services and other forms of support are available to students at risk of dropping out, far too many students still fail to complete high school. The problem may lie in the fact that services for at-risk students are spread across many different government agencies and nonprofit organizations, which makes it more challenging for schools to identify the services available to students and for students themselves to use them.

This report examines an integrated model of student support created by Communities In Schools (CIS), which is now working in about 2,400 schools and 360 school districts. The CIS model provides comprehensive and integrated services to students in different areas (academics, behavior, social skills and life skills, family outreach, health and wellness, etc.), delivered with varying intensity and duration based on students' level of need. Level 1 short-term, preventive services are broadly available to all students at a school, whereas Level 2 intensive, longer-term, targeted services are for students at higher risk of dropping out.

The schools in this study, located in Texas and North Carolina, started implementing the CIS model between 2005 and 2008. Using a quasi-experimental design, the study found mixed but promising results. In elementary schools, it appears that the CIS model may have improved attendance rates, which is consistent with the findings of other studies of CIS (including a randomized experiment in K-8 schools in Chicago). In middle schools, the CIS model does not appear to have improved any of the outcomes that could be measured in this study (attendance rates and test scores). In high schools, graduation rates improved after the CIS model was implemented, but it is not clear whether CIS caused these improvements.

Thus far, two quasi-experimental studies — the present one by MDRC and an earlier one by ICF International — have found that high schools implementing the CIS model have increased their graduation rates. In both studies, however, these findings have been inconclusive because of limitations related to finding a group of credible comparison schools. Therefore, a useful next step for CIS would be an evaluation based on a school-level randomized experiment. Such an experiment would provide the most rigorous evidence of the model's effects, and ultimately the most useful information for policymakers, districts, and schools.

Gordon L. Berlin
President, MDRC

Acknowledgments

This report is based on work supported by the Social Innovation Fund, a White House initiative and program of the Corporation for National and Community Service (CNCS). The Social Innovation Fund combines public and private resources with the goal of increasing the impact of innovative, community-based solutions that have compelling evidence of improving the lives of people in low-income communities throughout the United States.

The Edna McConnell Clark Foundation's Social Innovation Fund includes support from CNCS and 15 private coinvestors: The Edna McConnell Clark Foundation, The Annie E. Casey Foundation, The Duke Endowment, The William and Flora Hewlett Foundation, The JPB Foundation, George Kaiser Family Foundation, The Kresge Foundation, Open Society Foundations, Penzance Foundation, The Samberg Family Foundation, The Charles and Lynn Schusterman Family Foundation, The Starr Foundation, Tipping Point Community, The Wallace Foundation, and Weingart Foundation. This report would not have been possible without these organizations' support and commitment to the well-being of young people in low-income communities in the United States.

We owe special thanks to Communities In Schools (CIS) national and affiliate staff members for their support and cooperation throughout this study. At the national level, Heather Clawson provided important information about the organization. She also coordinated with other national and affiliate staff members to obtain data about the schools in the study, and we are grateful to them for providing this information.

We wish to thank several individuals outside of MDRC who gave us useful criticisms and encouragement throughout the study, and who provided insightful comments on earlier drafts of this report: Heather Clawson and Kevin Leary at CIS National; Gabriel Rhoads, Kelly Fitzsimmons, and Partheev Shah at The Edna McConnell Clark Foundation; Hilary Rhodes and Dara Rose at The Wallace Foundation; Jason Snipes and Bob Granger, who reviewed the report as members of the Evaluation Advisory Board for the Edna McConnell Clark Foundation's Social Innovation Fund; and the reviewers at CNCS.

At MDRC, William Corrin and Leigh Parise — the project director and manager for the study — provided extremely helpful suggestions, support, and guidance at critical times during the analysis of the data and the writing of this report. Kateryna Lashko processed and cleaned the many data files used for the analysis. Kelly Quinn and Daphne Chen contributed to report production, spending countless hours formatting tables and checking the accuracy of the information in the report, among other tasks. Kelly Granito helped our team organize its work on this report, with support from Kate Gualtieri. Kate was also a valuable link between the project team and the broader Edna McConnell Clark Foundation Social Innovation Fund community of program organizations and evaluators.

Gordon Berlin, Fred Doolittle, Howard Bloom, Robin Jacob, Rob Ivry, and Joshua Malbin carefully reviewed earlier drafts of the report, offered helpful ideas about how to present the findings, and provided valuable advice on conducting additional analyses. We also want to extend our thanks to MDRC's board members, who offered insightful critiques on how to interpret the findings from the study. Joshua Malbin edited the full report, and Carolyn Thomas and Ann Kottner prepared it for publication.

The Authors

Executive Summary

Every day more than 7,000 students drop out of school.[1] One-fifth of students who enter high school do not graduate within four years,[2] and more than two-fifths of Latino and African-American boys drop out.[3] Many students at risk of dropping out need academic and social services and other forms of support to make it through high school. However, these services are scattered across numerous government agencies and nonprofit organizations, which limits their potential to change the path of an at-risk student. Integrating student support services and connecting them with schools is viewed as a promising approach to assist school staff members and help students stay on track to graduate.[4]

This report presents the findings from a quasi-experimental study of the Communities In Schools (CIS) Model of Integrated Student Supports (referred to in this report as the "CIS model"), a promising whole-school approach that aims to reduce dropout rates by providing students with integrated and tiered support services based on their needs. Because some of the factors that put students on the path toward dropping out of high school are established well before ninth grade, CIS works with elementary schools, middle schools, and high schools. This study of the CIS model, which is funded by the Edna McConnell Clark Foundation's Social Innovation Fund, provides a unique opportunity to gain a better understanding of the potential effects of a tiered and integrated approach to dropout prevention that serves students of all ages.[5]

In high schools, the main finding from this study is that on-time graduation rates increased — and dropout rates decreased — in study schools after the CIS model was launched. However, it is not clear whether these improvements were caused by the CIS model. On the one

---

[1]Christopher B. Swanson, "Progress Postponed," *Education Week* 29, 34 (2010): 22-23.

[2]Richard J. Murnane, "U.S. High School Graduation Rates: Patterns and Explanations," *Journal of Economic Literature* 51, 2 (2013): 370-422; Marie C. Stetser and Robert Stillwell, *Public High School Four-Year On-Time Graduation Rates and Event Dropout Rates: School Years 2010-11 and 2011-12* (Washington, DC: National Center for Education Statistics, U.S. Department of Education, 2014).

[3]Schott Foundation for Public Education, *The Urgency of Now: The Schott 50 State Report on Public Education and Black Males* (Cambridge, MA: Schott Foundation for Public Education, 2012).

[4]Kristin A. Moore, Selma Caal, Rachel Carney, Laura Lippman, Weilin Li, Katherine Muenks, David Murphey, Dan Princiotta, Alysha Ramirez, Angela Rojas, Renee Ryberg, Hannah Schmitz, Brandon Stratford, and Mary Terzian, *Making the Grade: Assessing Evidence for Integrated Student Supports* (Bethesda, MD: Child Trends, 2014).

[5]The Social Innovation Fund is a program of the Corporation for National and Community Service. The Social Innovation Fund combines public and private resources to increase the impact of innovative, community-based solutions that have compelling evidence of improving the lives of people in low-income communities throughout the United States.

hand, graduation and dropout rates improved by greater amounts in the CIS high schools than in a group of comparison high schools, which suggests that the CIS model may have improved these rates more than they would have improved otherwise. On the other hand, the comparison schools and CIS schools had different graduation and dropout rates before the CIS model was implemented, so the comparison schools may not provide a credible reference point. For this reason, it is not possible to determine whether the CIS model was *more* effective than the strategies used by the comparison schools, although the study's findings suggest that the CIS model may be at least as effective as these other approaches.

In elementary schools, attendance rates (a central outcome measure CIS aims to affect in the elementary grades) improved in schools implementing the CIS model more than they did in a group of similar comparison schools. In middle schools, English/language arts (ELA) test scores did not improve in schools implementing the CIS model, whereas they did improve in a group of similar, comparison middle schools. Unfortunately, it was not possible to evaluate whether the CIS model improved middle school students' behavioral outcomes, which is the model's primary goal in those grades.

What Is the CIS Model?

Communities In Schools, which was founded in 1977 by children's advocate Bill Milliken, works with low-income K-12 students who are at risk of failing or dropping out of the nation's poorest-performing schools. CIS seeks to reduce dropout rates by integrating preventive services available to the entire school with intensive, targeted, and sustained services for the 5 percent to 10 percent of students who display significant risk factors for dropping out, such as poor academic performance, high absentee rates, or behavioral problems. CIS now serves 1.5 million students and their families in 25 states and the District of Columbia. It is active in approximately 2,300 schools and 360 school districts.[6]

CIS provides services and support to students in 10 different areas: academics, behavior, social skills and life skills, basic needs and resources, college and career preparation, enrichment and motivation, family outreach and engagement, health and physical wellness, community service, and mental health. These 10 categories of services are provided at two levels of intensity and duration, depending on students' needs:

- **Level 1 school-wide and preventive services:** Level 1 services are broadly available to all students at the school and are usually short-term, low-intensity activities or services (for example, making clothing or school sup-

[6]Communities In Schools, *2015 Annual Report* (Arlington, VA: Communities In Schools, 2015).

plies available to students, organizing a school-wide career fair, or hosting a financial aid workshop for twelfth-graders).

- **Level 2 targeted services:** In contrast, Level 2 services are intensive, often long-term, and targeted forms of support that are delivered to students who are displaying one or more significant risk factors for dropping out, such as poor academic performance, a high absentee rate, or behavioral problems. Level 2 services include forms of support such as individual and group counseling, tutoring, and after-school programs.

The CIS model's goals are different in high schools, middle schools, and elementary schools. In all three, the CIS model provides support in each of the 10 service categories, but the emphasis in each case is calibrated to the model's core goals for that type of school. In elementary schools, the model focuses on improving attendance rates by reaching out to and engaging parents. In middle schools, the model begins to emphasize helping students improve their behavior. In high schools, the model focuses on services specifically intended to prevent students from dropping out, help them progress through school, and make sure they graduate.

How Was the CIS Model Evaluated?

This study conducted by MDRC examines whether introducing the CIS model with all of its components improves schools' graduation rates, dropout rates, attendance rates, and state test scores.

The effect of the CIS model is evaluated using a comparative interrupted time series (CITS) design.[7] Two groups of schools are studied: one group that implemented the CIS model (CIS schools) and another group that did not implement the CIS model but was free to adopt some other reform or initiative (comparison schools). The first step in a CITS design is to determine the trends in school outcomes for the CIS schools and the comparison schools during the years *before* the intervention was launched. These are called the "baseline trends." The second step is to gauge how much the CIS and comparison schools "deviated" from their baseline trends *after* the intervention was launched. The outcomes of the two groups of schools are not compared directly; instead, the analysis compares the amount by which the two groups deviated from their separate baseline trends. If the CIS model is more effective than other programs or reforms available to schools (those used by the comparison schools), then the CIS

[7]For a discussion and history of CITS designs, see William R. Shadish, Thomas D. Cook, and Donald T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (Boston: Houghton Mifflin, 2002). For a discussion of these designs in the context of education research, see Howard S. Bloom, "Using 'Short' Interrupted Time-Series Analysis to Measure the Impacts of Whole-School Reforms, with Applications to a Study of Accelerated Schools," *Evaluation Review* 27, 3 (2003): 3-49.

schools should experience improvements relative to their baseline trend that exceed the improvements found in comparison schools.

The comparison schools play an important role in this study design. Their trends over time reveal what happened to similar schools that did not implement the CIS model and thus are intended to show how much CIS schools' outcomes would have improved had they not implemented it. For example, had they not implemented the CIS model, CIS schools' outcomes may have improved because they chose to implement some other school reform model (instead of CIS), or they may have improved due to a district-wide or statewide policy change. The CITS design makes it possible to identify the CIS model's effect over and above the effect of these alternative reforms and system-wide policy changes (that is, the model's *net* effect).

This study examined 53 schools in Texas and North Carolina (14 high schools, 15 middle schools, and 24 elementary schools) that started implementing the CIS model from 2005 to 2008. The study sample also includes 78 comparison schools (18 high schools, 24 middle schools, and 36 elementary schools). The comparison schools were chosen from non-CIS schools in counties in Texas and North Carolina where there is a CIS presence, ensuring that they are located in the kinds of districts where CIS typically operates. Matching methods were used to select comparison schools whose baseline characteristics and trends in school outcome measures were as similar as possible to those of the CIS schools.

Overall, the CIS schools and the comparison schools have similar baseline characteristics and outcome measure values, with one exception: The comparison high schools had higher graduation rates than the CIS schools (by about 9 percentage points) and lower dropout rates (by about 1.5 percentage points). These differences are small enough to meet commonly used criteria for baseline equivalence, such as the one used by the What Works Clearinghouse.[8] Substantively, however, these differences are large enough that the CIS schools may have had more incentive to turn themselves around. For example, had they not implemented the CIS model, the CIS schools may have chosen to implement different — but still intensive — whole-school interventions to improve their low graduation rates. In contrast, the comparison schools may have felt less need or pressure than the CIS schools to initiate a turnaround, and accordingly, they may have adopted less intensive strategies to improve their graduation rates. In that case, the comparison schools' deviations from trend would not provide the right information about what would have happened to the CIS schools had they not implemented the model. Any improvements in the comparison schools' graduation rates would understate the amount CIS

[8]The What Works Clearinghouse criterion is that differences in baseline characteristics should not exceed 0.25 standard deviations. See What Works Clearinghouse, *What Works Clearinghouse Procedures and Standards Handbook, Version 3.0* (Washington, DC: U.S. Department of Education, Institute of Education Sciences, 2014).

schools' own rates would have improved without the CIS model; by extension, the findings from this study would overstate the true effect of the CIS model. Various sensitivity tests were used to explore this possibility as well as other factors that could affect the credibility of the comparison schools as a reference point.

Data for the study were obtained from existing, publicly available school-level historical databases.[9] Using these data, this study is able to examine whether the CIS model was successful at meeting its main objectives: increasing on-time graduation rates and reducing high school dropout rates. The study can also examine whether the CIS model had a positive effect on attendance rates (the main outcome of interest in elementary schools) and performance on state tests. However, the study is *not* able to evaluate whether the CIS model improved students' behavioral outcomes, which is one of the model's main goals in middle schools. Another data-related limitation is that information is not available on what kinds of services and support are offered to students in the comparison schools, and how these services differ from those included in the CIS model. This limitation makes it more challenging to interpret the study's findings, because it is unclear what initiatives and services the CIS model is being compared with.

In addition to evaluating the effect of CIS' whole-school model, MDRC is also evaluating the effect of one component of the CIS model — Level 2 services — using a student-level random assignment research design. This study is being conducted in 28 secondary schools in North Carolina and Texas. Eligible at-risk students were randomly assigned to receive Level 2 case management and services or to continue with business as usual, with access to whatever other forms of support were available to them.

The first report from the random assignment study described the effect of Level 2 services on students' behavioral and academic outcomes after one year.[10] It found that Level 2 services had a positive and statistically significant impact on students' reports of having caring, supportive relationships with adults outside of home and school; the quality of their peer relationships; and their belief that education has positive value for their lives. But MDRC found no evidence that Level 2 services had improved students' attendance, course performance, or behavior. However, it is too early to make any definitive conclusions, because one year of case management may not be sufficient to improve the outcomes of these students, all of whom face serious academic and personal challenges. Accordingly, the next (and final) report will examine the effect of Level 2 case management after two years.

---

[9]These databases include the Common Core of Data and data sets maintained by the Texas Education Agency, the North Carolina Research and Data Center at Duke University, and the North Carolina State Department of Education.

[10]William Corrin, Leigh Parise, Oscar Cerna, Zeest Haider, and Marie-Andrée Somers, *Case Management for Students at Risk of Dropping Out: Implementation and Interim Impact Findings from the Communities In Schools Evaluation* (New York: MDRC, 2015).

Did the CIS Model Improve School Outcomes?

MDRC's study of the CIS model examines the following two research questions:

- After launching the CIS model, did the CIS schools have better outcomes than predicted by their baseline trend? This question is answered by looking at the deviations from baseline trend for the CIS schools in the study.

- Did the CIS schools' outcomes improve more than they would have in the absence of the CIS model? This question is answered by looking at the difference between the deviations from baseline trend for the CIS schools and the comparison schools. This difference represents the estimated net effect of the CIS model.

High School Graduation and Dropout Rates

- **After three years of implementing the CIS model, on-time graduation rates and dropout rates improved by statistically significant amounts in the CIS high schools, relative to what would have been expected given their baseline trends.**

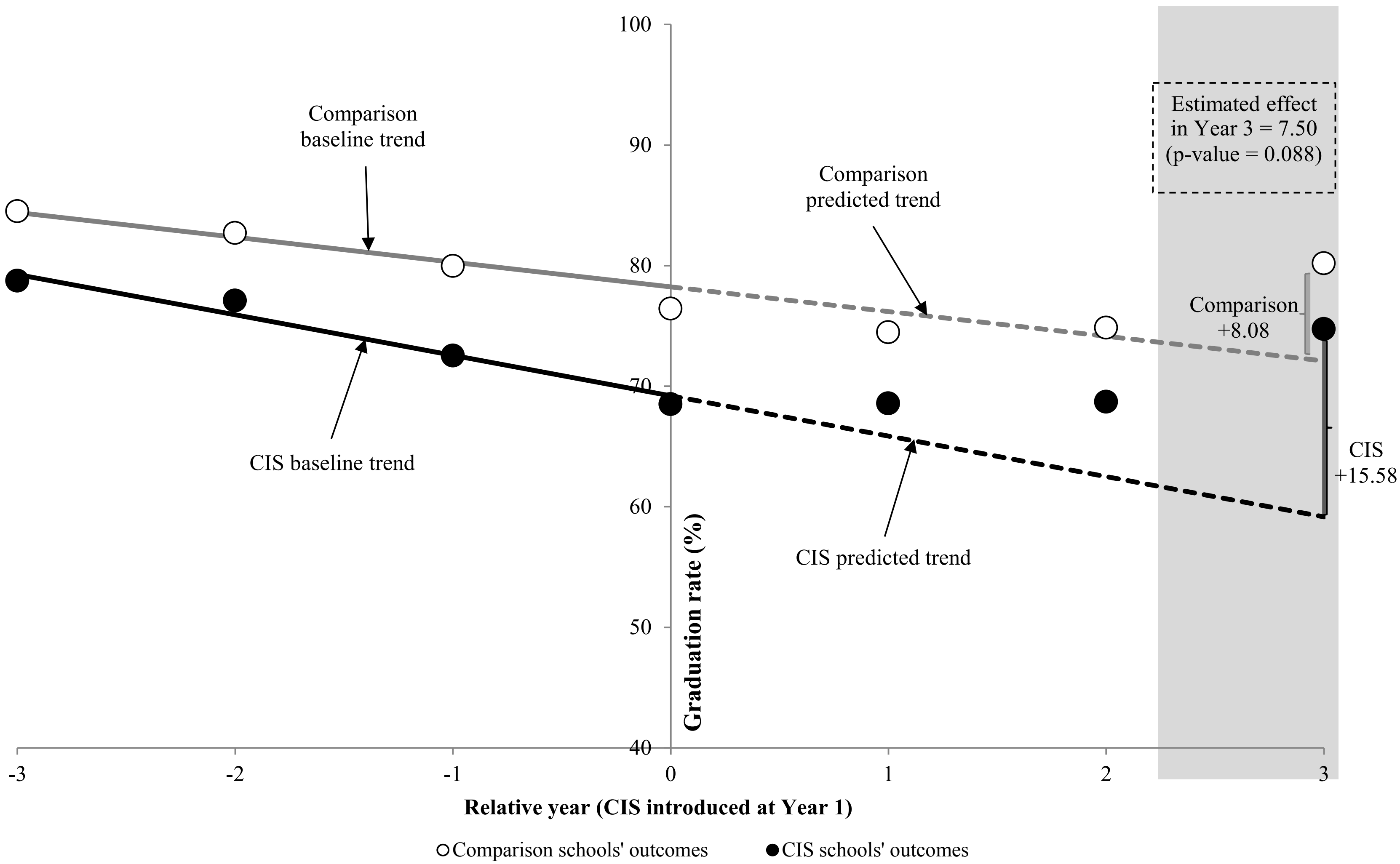
- **It is not possible to determine the extent to which these improvements can be attributed to the CIS model. Graduation and dropout rates improved more in the CIS schools than in the comparison high schools, but this pattern of results depends on the choice of comparison schools.**

Figure ES.1 plots the trend in graduation rates for the CIS high schools (solid black line) and the comparison schools (solid gray line), during the school years before the CIS model was launched, and also shows graduation rates after the model was launched. Before the CIS model was implemented, graduation rates had been steadily declining for both groups of schools. After the model was launched, graduation rates in the CIS high schools were higher than expected relative to these schools' baseline trend. Graduation rates were also higher than predicted in the comparison schools, perhaps because they chose to use a strategy other than the CIS model to turn themselves around. However, graduation rates improved more in the CIS schools than in the comparison schools, and therefore the estimated net effect of the CIS model is positive.

Table ES.1 summarizes the findings. As shown in the first panel, the estimated effect of the CIS model on graduation rates is consistently positive during the first three years of

Figure ES.1

Baseline Trends and Deviations from Trend in High School Graduation Rates



SOURCES: MDRC calculations based on school-level data on student outcomes from the Texas Education Agency.

Table ES.1

Estimated Effects on Selected Outcomes

	Deviation from Baseline Trend							
	CIS		Comparison		Estimated	Lower	Upper	P-Value for Estimated
Outcome and Follow-Up Year	Schools		Schools		Effect	90% CI	90% CI	Effect
High schools								
Graduation rate (%)								
Year 1	2.71	†	-1.74		4.44	-0.48	9.37	0.137
Year 2	6.19	†††	0.69		5.51	-0.77	11.78	0.148
Year 3	15.58	†††	8.08	††	7.50 *	0.27	14.73	0.088
Number of schools	8		11					
Dropout rate (%)								
Year 1	-0.75		0.30		-1.06	-2.40	0.29	0.197
Year 2	-1.85	††	-0.65		-1.19	-2.75	0.36	0.207
Year 3	-3.8	†††	-2.3	†††	-1.5	-3.4	0.5	0.211
Number of schools	14		18					
Middle schools								
ELA state test scores (z-scores)[a]								
Year 1	0.01		0.05	†	-0.05	-0.11	0.02	0.252
Year 2	0.02		0.10	†††	-0.08	-0.16	0.00	0.110
Year 3	0.00		0.11	†††	-0.11 *	-0.20	-0.01	0.061
Number of schools	8		15					
Elementary schools								
Attendance rate (%)								
Year 1	0.30	†††	0.14	†	0.16	-0.07	0.39	0.257
Year 2	0.44	†††	0.17		0.27	-0.01	0.54	0.111
Year 3	0.61	†††	0.20	†	0.41 **	0.10	0.72	0.030
Number of schools	24		36					

(continued)

Table ES.1 (continued)

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated deviations from baseline trend for each group of schools. The values in the "Estimated Effect" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. The values in the "Lower 90% CI" and "Upper 90% CI" columns are the 90 percent confidence intervals for the estimated effects. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed test was applied to estimated deviations and estimated differences between CIS schools and comparison schools. The statistical significance of estimated deviations is indicated as: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

implementation, and it becomes larger over time, as one would expect.[11] By the third year, the estimated effect is 7.5 percentage points, which represents an 11 percent increase in graduation rates relative to the last baseline year. An effect of this size means that in the third year of implementation, 55 additional students in each study school graduated from high school on time. Though the CIS model's estimated effect on dropout rates is not statistically significant, those results follow a similar pattern (as shown in the second panel of Table ES.1).

As explained earlier, a central assumption of the CITS design is that the comparison schools' deviations from their baseline trend represents what would have happened to the CIS schools in the absence of the intervention. As shown in Figure ES.1 and discussed earlier, the comparison schools had substantially higher baseline graduation rates (and lower dropout rates) than the CIS schools, so they may have had less incentive to turn their graduation rates around. This difference places an additional burden on the study to demonstrate that the comparison schools can still represent what would have happened to CIS schools in the absence of the CIS model. Various sensitivity tests were conducted to examine the credibility of this assumption. Some of these sensitivity tests suggest that the comparison schools do provide a good point of reference for the CIS schools, but others suggest that they may not, and that the net effects shown in Table ES.1 may therefore be overstated.

Given these results, it is not possible to determine whether the CIS model improved graduation and dropout rates more than they might have improved otherwise, or if it did, by

[11]Graduation rates are only available for 8 of the 14 CIS high schools, because data on this outcome are only available for high schools in Texas.

how much. However, the study does show that the average graduation and dropout rates for the CIS high schools improved relative what would have been expected given their prior trends. In addition, although it is not possible to determine whether the CIS model was *more* effective than the strategies used by the comparison schools, the study's findings suggest that the CIS model may be at least as effective.

Elementary School Attendance Rates

- **After three years, the average attendance rate increased by a statistically significant amount in the CIS elementary schools, relative to what would have been expected given their baseline trend.**
- **The improvement in attendance rates for these CIS schools was larger than the improvement in comparison schools by a statistically significant amount. The CIS model appears to have improved these schools' attendance rates more than they would have improved otherwise.**

Attendance is the main student outcome CIS aims to affect in elementary schools. After the CIS model was launched, attendance rates in the CIS elementary schools were higher than predicted by these schools' baseline trends by a statistically significant amount (as shown in the last panel of Table ES.1). Attendance rates in the comparison schools were also higher than predicted, perhaps because these schools were using other strategies to improve their attendance rates. However, in all three follow-up years, the CIS elementary schools deviated from their baseline trend by a greater amount than in the comparison schools, and that difference becomes statistically significant in the third year of implementation. This pattern of results holds across all of the sensitivity tests that were conducted to examine the credibility of the comparison schools as a reference point. Therefore, the CIS model may have improved attendance rates in the study's elementary schools more than the strategies being used by the comparison schools.

On its face, the CIS model's estimated effect on attendance rates may seem small: By the third year, the estimated effect is 0.4 percentage points, which represents an extra 0.7 days of school.[12] However, it is important to note that attendance rates were already high before the CIS model was launched (96.1 percent), so an effect of 0.4 percentage points gets schools 10 percent closer to perfect attendance.[13] In addition, average daily attendance rates can mask high rates of chronic absenteeism: 11 percent of elementary school students are chronically absent

[12]This number assumes a school year of 180 days.

[13]The gap between baseline attendance and perfect attendance is 3.9 percentage points (100 - 96.1). An effect of 0.4 percentage points is 10 percent of that gap (0.4 ÷ 3.9).

(that is, they are absent 15 or more days during the school year).[14] A small effect on overall attendance rates may represent a larger effect on chronic absenteeism.

Middle and High School Attendance Rates

In CIS middle schools, attendance rates did not improve by a statistically significant amount. In CIS high schools, attendance rates did improve by a statistically significant amount, but not by a greater amount than they would have improved otherwise. These results are consistent across the sensitivity tests that were conducted to examine the credibility of the comparison schools.

Middle School State Test Scores

- **After three years of implementation, state test scores did not improve in the CIS middle schools, relative to what would have been expected given their baseline trends.**
- **During the same period, state test scores did improve by a statistically significant amount in the comparison schools. CIS middle schools' test scores appear to be lower than they would have been otherwise.**

After the CIS model was launched, state test scores in the CIS middle schools were not better than predicted by their baseline trends. In contrast, the comparison schools performed better on state tests than predicted, perhaps because they implemented other reforms or initiatives to improve their students' performance. As a result, the CIS model's estimated effect on middle school students' ELA state test scores is consistently negative and it becomes more negative over time (as shown in the third panel of table ES.1).[15] By the third year, the estimated effect size is -0.11 standard deviations, which translates into about 14 weeks of learning.[16] These results hold across all of the sensitivity analyses that were conducted to examine the credibility of the comparison schools as a reference point. This result suggests that the CIS model may have been less successful at improving state test scores than the strategies used by

[14]Office of Civil Rights, *2013-2014 Civil Rights Data Collection, a First Look: Key Data Highlights on Equity and Opportunity Gaps in our Nation's Public Schools* (Washington, DC: U.S. Department of Education, Office of Civil Rights, 2016).

[15]The number of CIS and comparison schools for this analysis is less than the total number of middle schools in the study, because time-series data on state test scores are not available for all schools.

[16]The conversion from effect size to weeks of learning is based on data in Carolyn J. Hill, Howard S. Bloom, Alison Rebeck Black, and Mark W. Lipsey, "Empirical Benchmarks for Interpreting Effect Sizes in Research," *Child Development Perspectives* 2, 3 (2008):172-177. Middle school students make gains of about 0.008 standard deviations per week in reading. Therefore, an effect size of -0.11 standard deviations is equivalent to 14 weeks of learning (0.11 ÷ 0.008).

the comparison schools. It is important to note, however, that the CIS model does not attempt to improve state test scores or the quality of instruction provided during regular school hours.

Elementary School and High School State Test Scores

In CIS elementary schools, ELA state test scores improved by a statistically significant amount, but not by more than they would have improved otherwise. These results are consistent across the sensitivity tests that were conducted to examine the credibility of the comparison schools. In CIS high schools, ELA state test scores also improved by a statistically significant amount. However, it is not possible to determine the extent to which these improvements can be attributed to the CIS model due to the limitations in identifying a group of credible comparison schools described above.

Discussion

Because this study is based on a small number of purposefully selected schools that started implementing the CIS model a decade ago, its findings may not represent the effect of the CIS model nationally as it exists today. Therefore, the results from this study should be considered alongside the results of other evaluations of the CIS model, two of which are worth noting. The first is a national quasi-experimental study of the CIS model conducted by ICF International.[17] The second study is an (as yet) unpublished school-level random assignment evaluation of the CIS model in Chicago K-8 schools.[18]

All three studies of the CIS model conducted thus far have found positive effects on attendance rates for younger students. With respect to state test scores, the effect of the CIS model appears to depend on the local context: This study finds negative effects in middle schools, whereas the Chicago study finds positive effects. This difference suggests that in some settings, the CIS model *can* have positive effects on students' test scores and be more effective than the other strategies available.

The findings in high schools are promising but more difficult to interpret. The ICF study finds an improvement of 1.7 percentage points in ninth-grade students' probability of graduating from high school, but this effect is not statistically significant. In the present study, graduation rates improved for the CIS schools after they launched the model, but it is unclear to

[17]ICF International, *Communities In Schools National Evaluation Volume 1: School-Level Report. Results from the Quasi-Experimental Study, Natural Variation Study, and Typology Study* (Fairfax, VA: ICF International, 2008).

[18]David N. Figlio, "Experimental Evidence of the Effects of the Communities In Schools of Chicago Partnership Program on Student Achievement," Northwestern University Working Paper (Evanston, IL: Northwestern University, 2015).

what extent graduation rates would have improved had these schools not implemented the CIS model.

Therefore, it would be a useful next step for CIS to undertake a rigorous and large-scale study of its model in high schools in particular, based on a randomized experiment and supplemented by a cost study. A school-level random assignment research design would provide the best evidence of the CIS model's effect on student outcomes relative to that of other programs and strategies. By collecting cost data on the implementation of the CIS model as well as cost data on the strategies and interventions used by the control schools, one could also determine the CIS model's relative cost-effectiveness, which would ultimately provide the most useful information for school districts.

Chapter 1

Introduction

Every day more than 7,000 students drop out of school.[1] Among Latinos and African American boys, the dropout rates are 42 percent and 48 percent, respectively.[2] Even though high school graduation rates have risen this century, too many students who enter public high school — one in five — do not graduate within four years.[3] Compared with high school graduates, dropouts earn less money and are more likely to live in poverty, suffer from poor health, be involved in crime, and be dependent on social services.[4] Studies have found that each high school graduate brings a net economic benefit to society of around $127,000 and that the benefits of high school graduates are 2.5 times greater than the costs of educating them.[5]

One major obstacle to reducing dropout rates is that dropout-prevention services are delivered piecemeal. Support services for students at risk of dropping out are spread across government agencies and nonprofit organizations in communities with low-performing schools. As a result, students with multiple needs may receive one needed service but not others, limiting the ability of any given service to change that student's path.

This report presents the findings from a quasi-experimental study of the Communities In Schools (CIS) Model of Integrated Student Supports (referred to in this report as the "CIS model") — a whole-school approach that aims to reduce dropout rates by providing students with cohesively integrated and tiered support services based on their needs. Integrating student support services (that is, reducing their fragmentation across community agencies and organizations and connecting them with schools) is viewed as a way to provide necessary assistance to school staff members and to help students stay on track to graduation.[6] Similarly, providing students with varying levels of services based on their academic and personal needs is a promising strategy for making sure that the students who at the greatest risk of dropping out receive the intensive support they require to succeed.[7] In addition to working with high schools, CIS also works with middle schools and elementary schools in an effort to address some of the early factors that put students on the path toward dropping out of high school. This study of the CIS

[1]Swanson (2010).
[2]Schott Foundation for Public Education (2012).
[3]Murnane (2013); Stetser and Stillwell (2014).
[4]Child Trends (2015).
[5]Levin, Belfield, Muennig, and Rouse (2007).
[6]Moore et al. (2014).
[7]Simonsen, Sugai, and Negron (2006).

model aims to examine the effects of this tiered and integrated approach to dropout prevention that serves students of all ages.

This remainder of this chapter provides further information on the CIS model: the types of schools and students that it serves; the services and other forms of support it provides to students; and the types of outcomes that that it aims to improve. The chapter then provides a brief overview of this evaluation of the CIS model and its objectives.

The CIS Model

Communities In Schools works in low-income communities and in some of the nation's poorest-performing schools. It was founded in 1977 by children's advocate Bill Milliken and since then has gained extensive national reach. The program now serves 1.5 million students and their families in 25 states and the District of Columbia.[8]

CIS partners with schools located in areas of high poverty and low student performance. The program is now active in approximately 2,300 schools across 360 districts.[9] Around 82 percent of all students served by CIS are racial minorities and 94 percent are eligible for free or reduced-price lunches.[10] Because many of the factors that put students on the path toward dropping out of high school are established well before ninth grade, CIS serves schools of all grades: around 37 percent of CIS schools are elementary schools, 24 percent are middle schools, and 23 percent are high schools. The remaining 16 percent are combined and nontraditional schools (that is, alternative, charter, and magnet schools).[11]

The national office of CIS (CIS National) oversees a network of 161 independent nonprofit organizations operating as CIS local affiliates.[12] CIS National is responsible for developing and enhancing the CIS model; communicating about the model to national audiences and advocating for education reform that includes integrated student support; fostering collaboration among members of its network; supporting research into and evaluation of the model; and establishing national partnerships intended to generate resources and funding for members of the network. The local affiliates oversee and guide the implementation of the model in schools.

[8]Communities In Schools (2015).
[9]Communities In Schools (2015).
[10]Communities In Schools (2015).
[11]Communities In Schools (2011).
[12]Communities In Schools (2015).

They build community partnerships and develop local funding and resources to support model implementation in their schools.[13]

Features of the Model

The CIS model is a system of integrated services and other forms of support for students that addresses both the short-term and longer-term risk factors associated with dropping out of school (Figure 1.1). To meet its objectives, CIS places a site coordinator at a school who is given the task of providing tiered services to students in 10 different areas: academics; behavior; social skills and life skills; basic needs and resources; college and career preparation; enrichment and motivation; family outreach and engagement; health and physical wellness; community service; and mental health.

These 10 categories of services are provided at two levels of intensity and duration, depending on students' needs:

- **Level 1 school-wide preventive services:** Level 1 services are broadly available to all students at the school and are usually short-term, low-intensity activities or services (for example, making clothing or school supplies available to students, organizing a school-wide career fair, or hosting a financial aid workshop for twelfth-graders). Level 1 services also include short-term "crisis" interventions when an extreme event disrupts a student's life (for example, finding a solution if the power is turned off at a student's home or providing short-term counseling in response to a traumatic event). There is no enrollment process for these short-term services, and all students can participate.[14] CIS's internal standards say that Level 1 services must reach a minimum of 75 percent of students at the school.

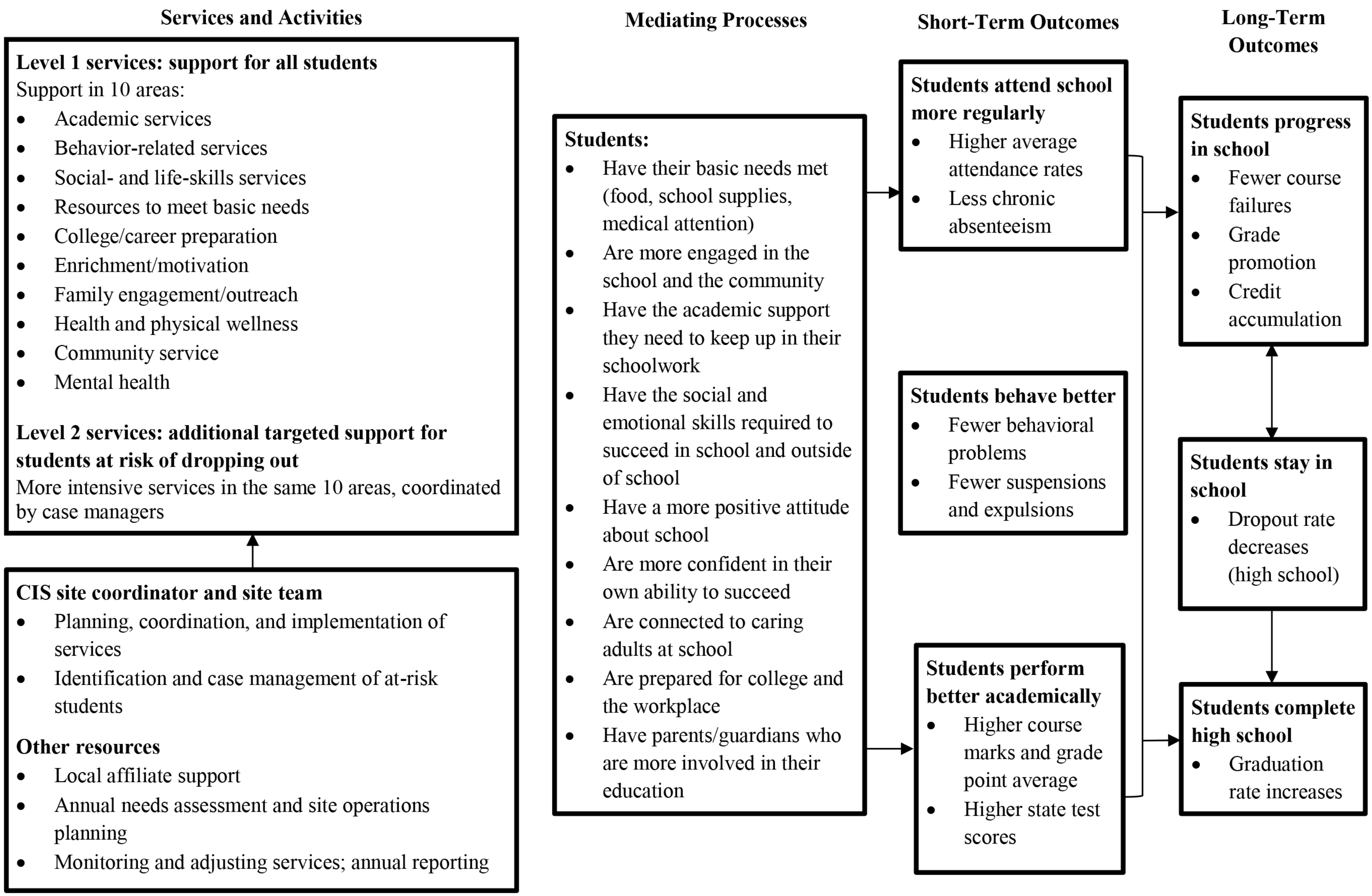
- **Level 2 targeted services:** In contrast, Level 2 services are intensive, often long-term, and targeted forms of support that are delivered to students who are displaying one or more significant risk factors for dropping out, such as poor academic performance, a high absentee rate, or behavioral problems. Level 2 services include individual and group counseling, tutoring, and after-school programs. Level 2 services are typically long term and high intensity,

[13]ICF International (2008); Porowski and Passa (2011).

[14]Sometimes services are not available to all students, but are available to large groups of students with particular characteristics (for example, those in a specific grade) or with specific needs (for example, food backpacks for students whose families are short on food).

Figure 1.1

The CIS Model's Theory of Change



and they are delivered through a case management process that includes individual assessments, goals, and plans. Level 2 services are intended to reach 5 percent to 10 percent of the student population.[15]

Table 1.1 shows examples of the kinds of services that are offered at each level, in each service category. As illustrated by these examples, the CIS model does *not* attempt to restructure schools or to improve the quality of instruction provided during regular school hours. Instead, the CIS model provides students with support services that address the challenges they may be facing in different aspects of their lives.[16] The specific types of Level 1 and Level 2 services offered can vary from school to school depending on those schools' priorities.[17]

The site coordinator is the linchpin of the CIS model. This staff person supervises and monitors the planning, coordination, and implementation of Level 1 and Level 2 services at a school. CIS site coordinators also identify at-risk students, assess their needs, develop case plans for them, connect them with support in school and in the community based on those plans, and monitor their progress to ensure that their needs are met.

The implementation of the CIS model is also reinforced by staff members and resources at the affiliate and national levels. For example, CIS National provides guidelines to schools for developing an annual plan that lays out what services will be provided to students, and the local affiliate provides support to schools to implement this plan. CIS National also sets standards meant to strengthen the implementation of the model.

Outcomes and Goals of the Model

The CIS model's theory of change posits that the tiered, integrated services offered by CIS will provide students with the skills and resources they need to succeed (see Figure 1.1). Students will have greater confidence in themselves; be more engaged in their schools and communities; build trusting relationships with caring adults at school; have their basic needs

---

[15]CIS's internal standards call for Level 2 services to reach 5 percent of a school with more than 1,000 students and 10 percent of a school with less than 1,000 students.

[16]Some schools implement only either Level 1 or Level 2 services. In this report, the term "CIS model" is used to refer to the simultaneous implementation of both levels of services.

[17]CIS is now moving toward a three-tiered model of support. Tier I will be provided to the whole school (the current Level 1); Tier II will target groups of students with a common need; and Tier III will consist of the most intensive, individually tailored services provided one-on-one. If a student is receiving Tier II or III services, that student will also receive case management. However, the core elements of the CIS model have not changed — the model still provides integrated services using a tiered approach — so the findings from this study are still relevant to CIS and to the field of dropout prevention.

Table 1.1
CIS Level 1 and 2 Services

Service Category	Examples of Level 1 (All Students, Low Intensity)	Examples of Level 2 (Targeted Students, High Intensity)
Academic services	After-school academic assistance	Long-term tutoring
Behavior-related services	Guest speaker on gangs	Counseling related to anger-management issues
Social- and life-skills services	Guest speaker on attire/personal appearance	Personal financial management workshops
Resources to meet basic needs	Clothing closets open to all students	Weekly book bags with food for the weekends
College/career preparation	School-wide career fair	A series of workshops on résumé writing and interviewing skills
Enrichment/motivation	Class awards for community service	Enrichment activities focused on the arts in an after-school program
Family engagement/outreach	School-wide family night	A "Focus on the Arts" program for students and families
Health and physical wellness	School assembly on physical fitness or nutrition	A physical exercise program or nutrition class
Community service	School-wide adoption of a neighborhood to keep clean	Weekly volunteering at a retirement home
Mental health	Guest speaker brought in to discuss a school tragedy	Professional counseling

SOURCE: Communities In Schools.

met; and develop their social and emotional skills. These skills and resources will then ultimately lead to improvements in students' outcomes. In the short term, students will attend school more regularly; behave better; and perform better academically. These improvements, in turn, will increase students' likelihood of staying in, progressing through, and completing high school.[18] The relationship between the short-term and longer-term outcomes in the CIS model's

[18]The three longer-term outcomes in the CIS model's theory of change are also the three domains that the What Works Clearinghouse at the Institute of Education Sciences uses to rate the effectiveness of dropout prevention programs: (1) staying in school (measured by school dropout rates), (2) progressing in school (measured by grade promotion, credit accumulation, or both), and (3) high school completion (measured by rates of earning diplomas or equivalents).

theory of change is based on several studies showing that students' attendance, behavior, and academic performance predict whether or not they will stay in school and graduate.[19]

The goals of the CIS model differ from elementary school to middle school to high school. In all three, the CIS model provides support in each of the 10 service categories in Figure 1.1. However, the emphasis in each case is calibrated to the model's core goals for those grades. In elementary schools, the model focuses on improving attendance rates by reaching out to and engaging parents. In middle schools, the model begins to emphasize helping students improve their behavior. In high schools, the model focuses on services specifically intended to prevent students from dropping out, help them progress through school, and make sure they graduate.

The Context of This Study

This report is part of a two-study evaluation of the CIS model being conducted by MDRC, a nonprofit, nonpartisan education and social policy research organization. The evaluation is supported by the federal Social Innovation Fund (SIF). The SIF is a program of the Corporation for National and Community Service, which combines public and private resources to expand the impact of innovative, community-based solutions that have compelling evidence of improving the lives of people in low-income communities throughout the United States. (See Box 1.1 for further information.)

One study being conducted by MDRC — the one that is the subject of this report — evaluates the effect of the entire CIS model using a quasi-experimental research design. The study examines whether school outcomes improve when CIS places a site coordinator in a school and that person oversees the provision of Level 1 preventive services to all students and Level 2 intensive services for the students most at risk of dropping out. The schools in this study are located in Texas and North Carolina, the two states where CIS serves the most students.

MDRC's quasi-experimental evaluation of the CIS model examines the following primary research question: What is the CIS model's effect on the rates at which students stay in high school and complete high school? These are the longer-term outcomes in the CIS theory of change, and improving them is the CIS model's core goal in high schools. This study also examines the CIS model's effect on attendance rates, the main outcome the model aims to affect in elementary schools. Unfortunately, this study cannot determine whether the CIS model had an effect on students' behavioral outcomes, its core objective in middle schools. (As will be

[19]For example, see Balfanz, Herzog, and Mac Iver (2007).

Box 1.1

The Edna McConnell Clark Foundation (EMCF) Social Innovation Fund

The Social Innovation Fund (SIF) — an initiative enacted under the Edward M. Kennedy Serve America Act — directs millions of dollars in public-private funds to expand effective solutions in three issue areas: economic opportunity, healthy futures, and youth development and school support. This work seeks to create a catalog of proven approaches that can be replicated in communities across the country. The SIF generates a 3:1 private-public match, sets a high standard for evidence, empowers communities to identify solutions to social problems, and creates an incentive for grant-making organizations to channel funding more effectively to promising programs. Administered by the federal Corporation for National and Community Service (CNCS), the SIF is part of the government's broader agenda to redefine how evidence, innovation, service, and public-private cooperation can be used to tackle urgent social challenges.

The Edna McConnell Clark Foundation, in collaboration with MDRC and The Bridgespan Group, is leading a SIF project that aims to expand the pool of organizations with proven programs that can help low-income young people make the transition to productive adulthood. The project focuses particularly on young people who are at greatest risk of failing or dropping out of school or of not finding work; who are involved or likely to become involved in the foster care or juvenile justice system; or who are engaging in risky behavior, such as criminal activity or teenage pregnancy.

EMCF, with its partners MDRC and Bridgespan, selected an initial group of nine programs and a second group of three programs to receive SIF grants: BELL (Building Educated Leaders for Life), the Center for Employment Opportunities, Children's Aid Society-Carrera Adolescent Pregnancy Prevention Program, Children's Home Society of North Carolina, Communities In Schools, Gateway to College Network, PACE Center for Girls, Reading Partners, The SEED Foundation, WINGS for Kids, Youth Guidance, and Children's Institute, Inc. These organizations were selected through a competitive selection process based on evidence of impacts on economically disadvantaged young people, a track record of serving young people in communities of need, strong leadership and a potential for growth, and the financial and operational capabilities necessary to expand to a large scale.

The EMCF Social Innovation Fund initiative is called the "True North Fund" and includes support from CNCS and 15 private coinvestors: The Edna McConnell Clark Foundation, The Annie E. Casey Foundation, The Duke Endowment, The William and Flora Hewlett Foundation, The JPB Foundation, George Kaiser Family Foundation, The Kresge Foundation, Open Society Foundations, Penzance Foundation, The Samberg Family Foundation, The Charles and Lynn Schusterman Family Foundation, The Starr Foundation, Tipping Point Community, The Wallace Foundation, and Weingart Foundation.

discussed later, this study relies on publicly available school-level data from state websites, and states do not consistently make public data on behavioral outcomes such as disciplinary infractions.) However, the study does examine the CIS model's effect on students' scores on state English/language arts and math tests (a measure of academic performance), and these scores represent a short-term outcome in the model's theory of change.

MDRC's study considers the CIS model's effect during the first three years of implementation. The third year of implementation is of the greatest interest, because all else being equal, one would expect effects to grow over time, for two reasons. First, by the third year, some students have received CIS services for multiple consecutive school years (and have therefore been exposed to the CIS model more). Second, by the third year, the CIS services in the schools are more "mature" and their implementation potentially stronger.[20]

The current study is not the first time the CIS model has been evaluated. A previous evaluation of the model was conducted by ICF International.[21] That study examined the CIS model's effect on students' progress in school, attendance rates, and proficiency on state tests.[22] Its sample included 602 CIS schools in seven states (321 elementary schools, 158 middle schools, and 123 high schools) that started implementing the model between 1999 and 2002, and the same number of matched comparison schools. The study used a difference-in-differences research design to evaluate the effect of the CIS model: Effects were estimated by determining whether the outcomes of schools that implemented the CIS model improved more than the outcomes of the comparison schools. Though the estimated effects of the CIS model in

---

[20]Whole-school interventions like CIS are complex. They offer many different types of services to students across the entire school, so their implementation is by definition challenging. District and school staff members must invest in the program. School and program staff members must be trained, and their efforts must be coordinated. The site coordinator must arrange many kinds of services for students. CIS and the school must track a wide variety of student outcomes. In a tiered model such as CIS, there is also the added complexity of identifying which students need more intensive, targeted support. Given these challenges, it can take several years for a whole-school intervention to be implemented with the intended level of quality, so one might expect the effect of whole-school interventions to become larger over time. This expectation is borne out by previous evaluations: In a meta-analysis of comprehensive school-reform models, Borman, Hewes, Overman, and Brown (2003) find that models implemented for between one and four years had effects on academic achievement ranging from 0.13 to 0.17 standard deviations, while models implemented for five years or more had effects of 0.23 to 0.50 standard deviations.

[21]ICF International (2008); Porowski and Passa (2011).

[22]To conduct this analysis, ICF identified comparison schools not implementing CIS with very similar baseline characteristics to the schools in their sample that were implementing CIS. ICF looked for comparison schools of the same grades (elementary, middle, or high schools) in the same localities that had similar baseline characteristics, including their attendance rates, their numbers of students receiving free or reduced-price lunches, their total numbers of students, their numbers of students with special needs, their percentages passing the state math and English/language arts tests, and their racial compositions. For high schools, ICF also looked for comparison schools with similar dropout rates.

the ICF study are not statistically significant, there is suggestive evidence that the outcomes of CIS schools may have improved by more than the outcomes of the comparison schools. For example, the ICF study found that CIS had an estimated effect of 2.0 percentage points on ninth-grade students' predicted probability of making it to twelfth grade, and an estimated effect of 1.7 percentage points on ninth-grade students' predicted probability of graduating from high school.[23] However, these estimated effects could actually be due to preexisting differences in the trends of the two groups of schools. The outcomes of CIS schools may have improved more because the CIS schools were already improving faster than the comparison schools in the years before the intervention was launched.

A second study of the CIS model is an (as yet) unpublished school-level random assignment evaluation in 47 Chicago K-8 schools (20 CIS schools and 27 control schools).[24] The CIS schools started implementing the model in the 2012-2013 school year and effects on student outcomes were examined for the first two years of implementation. The study finds statistically significant reductions of 3 percentage points on truancy rates in the first year of implementation (though these effects disappear in the second year).[25] It also finds positive effects of 5.6 percentage points on students' math proficiency and of 4.3 percentage points on their reading proficiency in the second year of implementation. Though the Chicago study uses the strongest possible research design (a randomized experiment), it only includes K-8 schools and therefore it does not provide information on the CIS model's effect on graduation and dropout rates, the ultimate outcomes the model aims to influence.

MDRC's evaluation aims to produce additional evidence on the CIS model's effect in elementary, middle, and high schools using a strong quasi-experimental design — a comparative interrupted time series (CITS) design. In a CITS design, the effect of an intervention is evaluated by determining whether the outcomes of schools that implemented the intervention deviated from their baseline trend by a greater amount than similar comparison schools.[26] As will be explained in this report, the CITS design is more rigorous than most other quasi-experimental designs because it can account for preexisting differences in trends between CIS

---

[23]The probability of persisting to twelfth grade is based on a school's "promoting power," the ratio of twelfth-grade students in a given school year to the number of ninth-grade students three years earlier. The probability of graduating from high school is based on the cumulative promotion index (CPI). CPI is calculated using a series of ratios that compare enrollments in consecutive school years. As an example, for the 2008-2009 school year:

$$\text{CPI} = \frac{10th-graders, fall\ 2009}{9th-graders, fall\ 2008}\, x\, \frac{11th-graders, fall\ 2009}{10th-graders, fall\ 2008}\, x\, \frac{12th-graders, fall\ 2009}{11th-graders, fall\ 2008}\, x\, \frac{Diploma\ recipients, spring\ 2009}{12th-graders, fall\ 2008}$$

[24]Figlio (2015).

[25]A student is considered truant in this study if he or she had an attendance rate of less than 95 percent.

[26]Bloom (2003); Cook, Campbell, and Day (1979); Shadish, Cook, and Campbell (2002).

and comparison schools. Several studies have found that when well implemented, the CITS design is able to reproduce the results of a randomized experiment.[27]

In addition to evaluating the effect of CIS' whole-school model, MDRC is also evaluating the effect of one component of the CIS model — Level 2 case-managed services — using a student-level random assignment research design. That study is being conducted in 28 secondary schools in North Carolina and Texas. The CIS schools included in the study each had more eligible at-risk students than CIS site coordinators could serve, so students were randomly assigned either to receive case management from site coordinators (the case-managed group) or to continue with business as usual at their schools, with access to whatever other student services were available (the non-case-managed group).

The first report for this random assignment study described the effect of CIS Level 2 services on students' behavioral and academic outcomes after one year.[28] The study found that CIS Level 2 services had positive and statistically significant effects on students' reports of having caring, supportive relationships with adults outside of home and school; the quality of their peer relationships; and their belief that education has positive value for their lives (the mediating processes in the theory of change for CIS case management). However, MDRC found no evidence that CIS Level 2 services had improved students' attendance, course performance, or discipline after one year.[29] It is too early to make any definitive conclusions, because one year of case management may not be sufficient to improve the outcomes of these students, all of whom face serious academic and personal challenges. Accordingly, the next (and final) report will examine the effect of Level 2 case management after two years of services.

The present report — which focuses on the effect of the full CIS model — is structured as follows. Chapter 2 discusses various elements of the evaluation: the research design used to evaluate the effect of the CIS model, the study's data sources and outcomes, and the sample of CIS and comparison schools included in the study. Chapter 3 presents the study's main findings. Chapter 4 discusses these findings in the context of other studies of the CIS model, and reflects on the lessons that can be drawn from this study's results.

---

[27]For example, see St. Clair, Cook, and Hallberg (2014); Fretheim et al. (2013).

[28]Corrin et al. (2015).

[29]The first report also contained findings on the implementation of the CIS Level 2 services. The study found that in all affiliates, Level 2 services primarily focused on academic assistance, behavior, and social-skill development. However, the affiliates differed in how the Level 2 services were administered, which was to be expected given the affiliates' autonomy within the CIS national network. In addition, though students received varying amounts of Level 2 services, high-risk students (defined as those who had failed a course, been chronically absent, or been suspended in the previous year) did not receive more Level 2 services than moderate-risk students.

Chapter 2

Study Design

In this study, the effect of the Communities In Schools (CIS) model is evaluated using a comparative interrupted time series (CITS) design, a quasi-experimental design that lends itself well to exploring the effect of school-wide interventions. The study examines the CIS model's effect on the dropout rates, graduation rates, attendance rates, and state test performance of 53 CIS schools in Texas and North Carolina, during their first three years of CIS implementation. As is the case at all schools served by CIS nationally, the majority of students in the study schools are from low-income families, and most are racial minorities.

This chapter provides further information about the analytical approach this study uses to evaluate the CIS model's effect. The first section provides an overview of the comparative interrupted time series design. The second section discusses the data used in the study and the outcomes of interest. The third section describes the sample of CIS schools and comparison schools included in the study, as well as the sensitivity analyses that were conducted to gauge whether the comparison schools provide a credible reference point for what would have happened to the CIS schools had they not implemented the CIS model.

The CITS Design

Comparative interrupted time series designs have been used for decades to evaluate interventions in areas such as epidemiology, political participation, substance abuse, advertising, and employment.[1] In education, CITS designs have been used to evaluate federal policies like No Child Left Behind and the effect of whole-school reform models.[2] Studies have shown that a well-implemented CITS design can, in some circumstances, reproduce the results of a randomized experiment.[3]

In a CITS design, program impacts are evaluated by determining whether schools that implemented an intervention (in this case, the CIS model) "deviated" from their baseline trends by a greater amount than a group of similar comparison schools. Figure 2.1 illustrates how the

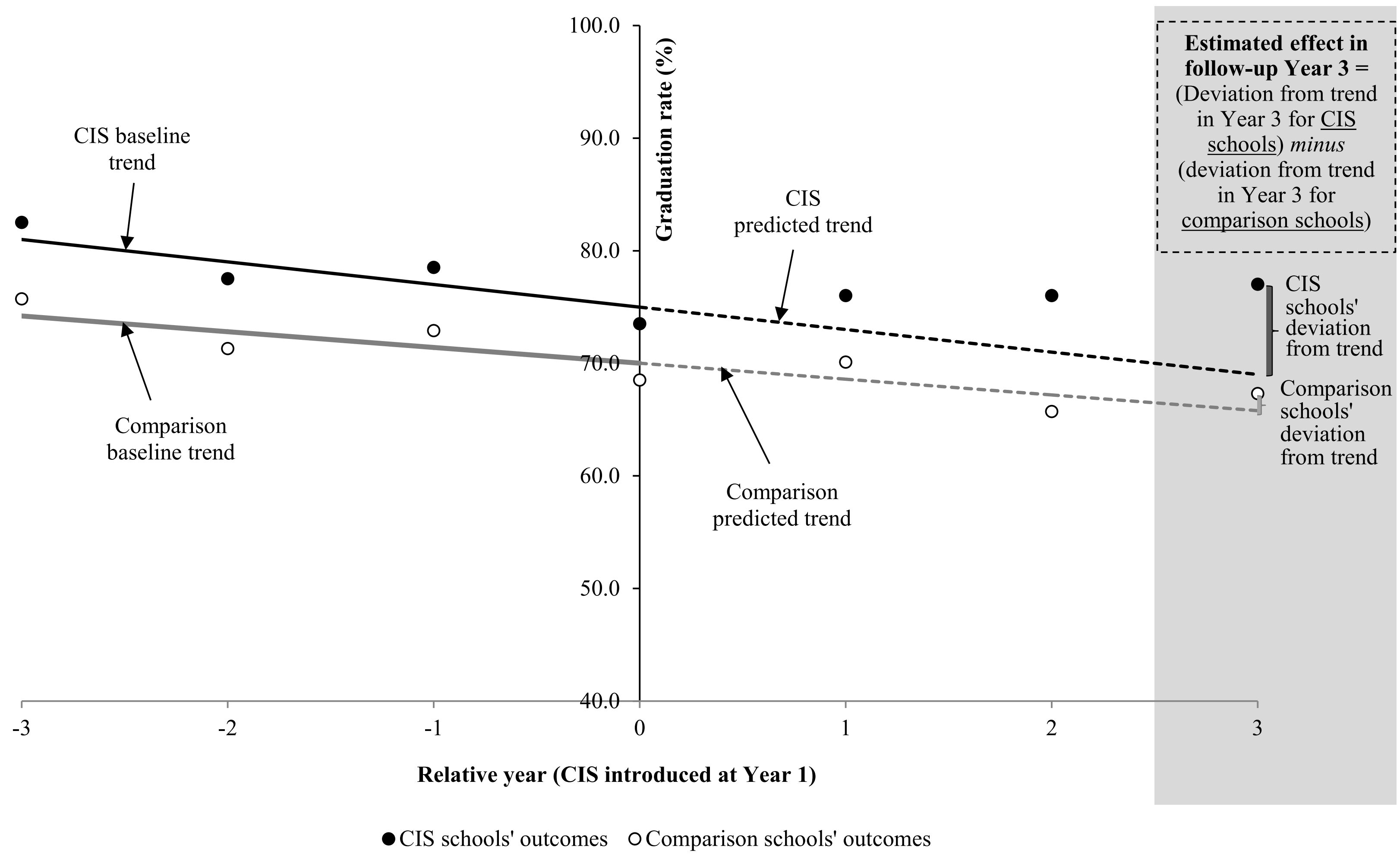
[1]Bloom and Riccio (2005); Ballart and Riba (1995); Campbell and Ross (1968); Mulford, Ledolter, and Fitzgerald (1992). For a discussion and history of CITS designs, see Shadish, Cook, and Campbell (2002). For a discussion of these designs in the context of education research, see Bloom (2003).

[2]Dee and Jacob (2011); Wong, Cook, and Steiner (2009); Kemple, Herlihy, and Smith (2005).

[3]St. Clair, Cook, and Hallberg (2014); Somers, Zhu, Jacob, and Bloom (2013); Fretheim et al. (2013).

Figure 2.1

Using a Comparative Interrupted Time Series Design to Estimate Effects on Graduation Rates: A Hypothetical Example



CITS design is applied in this study, using hypothetical data on graduation rates. The steps in the CITS design are as follows:

- **Mean outcomes:** First, school-level data are used to create a time series of the graduation rates for the CIS schools (black dots) and the comparison schools (white dots). The school years before the CIS model was implemented are referred to as the "baseline years" and the school years during which the CIS model was being implemented are the "follow-up years."

- **Baseline trends:** The next step is to estimate the pre-CIS trend in school outcomes during the baseline years. Baseline trends are estimated separately for the CIS schools in the study (the solid black line) and the comparison schools (the solid gray line). These trends are then projected into the follow-up period, in order to predict what graduation rates would be if these baseline trends were maintained (the dotted lines).

- **Estimated deviations from the baseline trend:** The next step is to compare *actual* graduation rates (the dots in the follow-up period) with the graduation rates *predicted* using baseline trends (the dotted lines). These "deviations from baseline trend" are estimated for the CIS schools and the comparison schools for each follow-up year.

- **Estimated effect:** Finally, the effect of the CIS model is estimated as the difference between the average deviation from baseline trend for the CIS schools and the average deviation from trend for the comparison schools. If the CIS model is more effective than other programs or reforms available (those used by the comparison schools), then CIS schools should experience positive deviations from their baseline trend that exceed the deviations found in the comparison schools. The effect of the CIS model is estimated for each of the first three years of CIS implementation.[4]

The CITS design is more rigorous than most other quasi-experimental designs because it combines time-series data with a matched comparison group. Together, these two design elements can eliminate most plausible alternative explanations (other than the effect of CIS) that could explain why the CIS schools deviated from their baseline trend more than the comparison schools.[5]

---

[4]Appendix A provides further information on the statistical model used to estimate program effects in a CITS design.

[5]Corrin and Cook (1998).

By using several years of preintervention (baseline) data, the CITS design eliminates the possibility that estimated effects are confounded with differences in preexisting trends between the CIS schools and the comparison schools.[6] For example, in Figure 2.1, CIS schools' graduation rates were decreasing somewhat faster than the graduation rates of the comparison schools during the school years before CIS was launched. The CITS design accounts for this difference in preexisting trends, because the effects it estimates are the difference in the *deviations from baseline trend* for each group of schools. If data were only available for one or two school years before the CIS model was launched — meaning that baseline trends could not be estimated — then a difference-in-differences design would have to be used instead. With this type of design, the effect of CIS would be estimated by determining whether the *change* in dropout rates over the follow-up period is greater for CIS schools than for the comparison schools. This design was used in ICF International's earlier study of the CIS model.[7] As explained in Chapter 1, one important problem with the difference-in-differences approach is that the estimated effect of an intervention may be confounded with differences in the two groups of schools' preexisting trends.

The comparison schools play a central role in a CITS design. Their trends reveal what happened to similar schools that did not implement the CIS model, and thus are intended to show how much CIS schools' outcomes would have improved had they not implemented the intervention. In short, comparison schools' deviations from their baseline trends provide a "counterfactual" against which to measure the improvement in CIS schools' performance. For example, had they not launched the CIS model, the CIS schools in this study may have chosen to implement some other whole-school reform to improve their students' outcomes, in which case students' outcomes may have improved even without the CIS model. Or, had they not implemented the CIS model, students' outcomes may have improved due to a state or district-wide educational initiative. The deviations from trend for the comparison schools are intended to capture the effects of alternative reforms that the CIS schools could have implemented (instead of the CIS model), as well as the effect of any district-wide or statewide policy changes that may also have affected the CIS schools' outcomes. In this way, the CITS design can identify the effect of the CIS model over and above the effect of these alternative reforms and system-wide policy changes (that is, the *net* effect of the CIS model).

---

[6]This possibility is also referred to as "maturation bias." See Shadish, Cook, and Campbell (2002).

[7]This design is also called a "pre-post design with comparison group." See Shadish, Cook, and Campbell (2002). For findings from the earlier study of the CIS model, see ICF International (2008).

Data Sources and Outcomes

The data for this study were obtained from existing, publicly available school-level historical databases. These databases include the Common Core of Data and data sets maintained by the Texas Education Agency, the North Carolina Education Research Data Center at Duke University, and the North Carolina State Department of Education. These data sets include data for multiple school years on high school dropout rates, graduation rates, attendance rates, and performance on state English/language arts (ELA) and math tests (see Table 2.1).

Using these data, it is possible to assess whether the CIS model achieved two of its longer-term goals in high schools: to help students *complete school* (as measured by graduation rates) and to help students *stay in school* (as measured by high school dropout rates):

- **High school graduation rate:** This study focuses on the four-year (on-time) longitudinal graduation rate. This measure follows a group of students entering the ninth grade at the same time, and measures the percentage of these students who graduated from high school four years later. For example, in this study, longitudinal graduation rates were about 78 percent for the study schools in 2004-2005, which means that of the students who entered ninth grade four years earlier, 78 percent graduated from high school on time.

- **High school dropout rate:** The measure used here is the annual dropout rate, defined as the number of students in ninth to twelfth grade who dropped out of school in a given school year, as a percentage of the total number of students in those grades.

It is important to note that the graduation rate and the dropout rate do not add up to 100 percent (for example, a 4 percent dropout rate does not correspond to a 96 percent graduation rate), because the two measures are substantively different. The dropout rate measures whether students have left school, or conversely, whether students are *still enrolled in school*. The graduation rate takes this a step further and measures whether students have *completed school* and graduated.[8]

[8]Mathematically, the denominators for these two measures are different because they focus on different target populations. The denominator for the dropout rate is all students at the school in a given school year, while the denominator for the graduation rate is the entering ninth-grade class from four years earlier. The two measures also have different numerators because their follow-up periods are defined differently. The numerator for the dropout rate is all students at the school who dropped out in a given year, whereas the numerator for the graduation rate is first-time twelfth-graders who graduated.

Table 2.1

Study Outcome Measures

Outcome Measure	School Type	Scale	Years of Data	
			Texas	North Carolina
Four-year (on-time) graduation rates	High schools	Percentage of entering 9th-grade students in school year *t*-4 who graduate in year *t*.	2002-2003 to 2010-2011	--[a]
Dropout rates	High schools	Number of students in 9th through 12th grade who dropped out, as a proportion of the number of students enrolled	2001-2002 to 2010-2011	2001-2002 to 2010-2011
Attendance rates	Elementary, middle, and high schools	Daily average of the percentage of students present	2001-2002 to 2010-2011	2003-2004 to 2010-2011
State test scores	Elementary, middle, and high schools	Amount by which a school's average score falls below the state mean, scaled as an effect size (z-score)[b]	2002-2003 to 2010-2011	2003-2004 to 2010-2011[a]

NOTES: [a]Data on high schools are available for some school years, but not for the range of years necessary for a CITS analysis.
[b]Effect sizes are based on the estimated student-level standard deviation in test scores for the state.

Each of these two measures has its strengths and limitations. The four-year graduation rate is stronger in some ways because it follows the same students over time and because it represents the ultimate objective of the CIS model. However, this measure is only available for the study high schools in Texas. Conversely, the annual dropout rate is a weaker measure: Dropout rates tend to be very low, so there is a "floor" to how much an intervention such as CIS can affect them. In this study, for example, only 1 percent to 4 percent of high school students are reported as having dropped out during the 2004-2005 school year. Annual dropout rates are low for two reasons: (1) the dropout rate is measured relative to *all* students at a school (including the lower grades, where dropout is less prevalent), and (2) the number of students in a school who drop out may be underestimated, because it is difficult to confirm which students actually dropped out of school (as opposed to transferring to another school). However, the annual dropout rate is available for *all* study schools.

Given these trade-offs, this study examines the CIS model's effect on both of these measures. Effects on graduation rates will be discussed first — as this is the preferred measure — but effects on dropout rates will also be examined to make sure there are consistent and similar results across both measures.

This study also examines whether the CIS model had an effect on two short-term outcomes in its theory of change:

- **Attendance rates:** This study focuses on the average daily attendance rate, defined as the average percentage of students present in school on a given day. As explained in Chapter 1, improving attendance rates is a core goal of the CIS model in elementary schools.

- **State test scores:** The study uses state test scores to measure students' academic performance. In order to pool state test scores across Texas and North Carolina and across grades, raw scores were converted to "z-scores," which describe performance relative to the distribution of test scores in the state as a whole. That is, the average score for a school represents the amount by which students at the school scored *below the state average*, as a proportion of the total variation in test scores among students in the state. Appendix D provides additional information on this rescaling.[9]

One of the limitations of this study is that it is not possible, given the data available, to evaluate the CIS model's effect on several of the domains it aims to improve. In particular, it is not possible to examine whether CIS has an effect on helping students *progress in school* — as measured by course failures, grade promotion, and credit accumulation. Nor is the study able to examine whether, in the shorter term, the CIS model improved students' behavior and reduced their likelihood of being suspended or expelled, which are important precursors of high school dropout and graduation as well as the core focus of the CIS model in middle schools. Thus, this study is not able to speak definitively about the CIS model's effect on its primary middle school goal.

Another limitation of this study is that data are not available on what kinds of services and other forms of support are offered to students in the comparison schools, and how these services differ from those included in the CIS model. As noted earlier, some of the comparison schools may have decided to improve their outcomes by adopting some other reform strategy. This possibility makes it more challenging to interpret the findings, because it is not clear which other types of initiatives and services the CIS model is being compared with.

[9]The research team also examined the percentage of students whose scores on state tests would lead them to be defined as "proficient" by their states. This measure is relevant for accountability purposes. The pattern of estimated effects on proficiency rates is similar to the effects on scaled scores, so these additional analyses are not presented.

The Study Sample

CIS Schools in the Study

In collaboration with CIS National, the research team decided to focus on the effect of the CIS model in North Carolina and Texas, for two reasons. First, CIS has a strong presence and history in both states: More than half of CIS schools (56 percent) are located in these two states.[10] Second, historical data on school-level outcomes are publicly available in both North Carolina and Texas, which makes it possible to use a CITS design to evaluate the effect of the model. In North Carolina, all CIS schools were considered for inclusion in the study sample, while in Texas, schools in three CIS affiliates were considered for inclusion.[11]

Among CIS schools in these locations, 53 were selected for the study (14 high schools, 15 middle schools, and 24 elementary schools). These schools were chosen because they offered both Level 1 and Level 2 services within the first three years of launching the CIS model, and because they had enough years of data to be included in a CITS analysis (at least four years of baseline data and three years of follow-up data). These CIS schools started implementing the CIS model in the mid-2000s (from 2005 to 2008; see Table 2.2). Appendix B provides further information on the criteria used to select the CIS schools for the study.

Relative to all regular public schools in Texas and North Carolina, the 53 CIS schools in the study serve a higher proportion of low-income and minority students, and their students perform worse academically (see Table 2.3). For example, in 2004-2005, 50 percent of students in the study's high schools were eligible for free or reduced-price lunches (compared with an average of 40 percent across the two states) and almost 70 percent of students were black or Hispanic (compared with 47 percent across the two states). The average dropout rate in the study's high schools was 3.2 percent (compared with an average of 1.2 percent across the two states) and the average graduation rate was 78 percent (compared with 89 percent across the two states). These results confirm that CIS primarily serves underperforming schools.

However, as explained earlier, the schools in the study are a small and purposefully chosen sample selected because they meet the data requirements for a CITS design. One important question, therefore, is whether the CIS schools in this study are representative of the schools served by CIS nationally. On the one hand, the study schools serve a similar proportion

---

[10]In 2013, 33 percent of CIS schools were in Texas, and 20 percent were in North Carolina. See Communities In Schools (2013).

[11]These three affiliates were chosen because they serve a large number of schools. Affiliates are independent CIS branches that serve one or more school districts in their area.

Table 2.2

CIS Schools in the Study

Characteristic	North Carolina	Texas	Total
School type			
Elementary	5	19	24
Middle	2	13	15
High	1	13	14
Start year			
2005-2006	0	14	14
2006-2007	0	11	11
2007-2008	5	7	12
2008-2009	3	13	16
Number of school districts	4	12	16
Number of CIS affiliates	4	3	7

SOURCE: MDRC calculations based on program information provided by Communities In Schools.

of minority students as CIS does nationally, on average: in 2004-2005, 68 percent to 84 percent of students in the CIS study schools were racial minorities, compared with 82 percent in CIS schools nationally. On the other hand, the number of schools in the study is small (especially the number of high schools), and although the study schools are located in the two states with the most CIS schools (Texas and North Carolina), 44 percent of CIS schools nationally are *not* in these two states. Moreover, all of the study schools started implementing the CIS model a decade ago (between 2005 and 2008), yet the CIS model and the education-policy context around it have evolved since then.[12] Therefore, the findings from this study might not represent the situations of CIS schools in other states or associated with other affiliates, and might not represent the CIS model's effect in schools that have implemented it more recently.

[12]The CIS schools in this study also serve a somewhat lower proportion of economically disadvantaged students than CIS schools nationally: 50 to 78 percent of students in the study schools were eligible for free or reduced-price lunches, compared with 94 percent of students in CIS schools nationally. See Communities In Schools (2015).

Table 2.3

Characteristics of CIS Schools in the Study and All Schools in Their States (NC and TX), 2004-2005 School Year

	Elementary Schools		Middle Schools		High Schools	
Characteristic	CIS Schools	All Schools in States	CIS Schools	All Schools in States	CIS Schools	All Schools in States
Graduation rate (%)	NA	NA	NA	NA	78.0	88.6
Dropout rate (%)	NA	NA	NA	NA	3.2	1.2
Attendance rate (%)	96.2	96.5	95.6	95.8	93.4	94.7
Proficiency on state tests (%)						
ELA	81.4	85.8	81.5	83.1	69.6	79.8
Math	80.9	84.0	63.2	67.3	53.8	64.5
Free or reduced-price lunch (%)	78.1	54.2	55.6	48.6	49.9	39.9
Racial/ethnic composition (%)						
Black	26.9	18.1	26.1	15.9	30.8	13.9
Hispanic	53.5	38.6	37.8	35.7	39.3	33.4
White	15.7	40.3	32.0	45.8	23.8	50.4
Other	4.0	3.0	4.1	2.6	6.0	2.3
Enrollment	603.7	510.1	795.1	630.1	2,235.6	1,077.8
Pupil/teacher ratio	14.7	15.1	14.3	14.4	16.1	13.7
Title I status[a] (%)	100.0	79.9	46.7	60.4	35.7	42.4
Location (%)						
City/town	100.0	75.6	93.3	68.5	78.6	60.2
Rural	0.0	24.4	6.7	31.5	21.4	39.8
Number of schools	24	4,966	15	1,690	14	1,350

SOURCES: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: Values in the "CIS Schools" columns are means for CIS schools in the study. Values in the "All Schools in States" columns are means for all regular public schools in North Carolina and Texas; the two states are weighted in proportion to the number of schools each contributed to the CIS sample. The values in this table are based on data from the 2004-2005 school year, the most recent school year when none of the study schools were implementing the CIS model.

NA = not applicable.

[a]This percentage represents the Title I status of schools in the 2004-2005 school year, before any of the CIS schools had implemented the intervention. By the 2010-2011 school year, all CIS schools were eligible for Title I funds.

Comparison Schools in the Study

The study sample includes 78 comparison schools (18 high schools, 24 middle schools, and 36 elementary schools). These comparison schools were selected from the pool of schools not implementing the CIS model ("non-CIS schools") in counties in Texas and North Carolina where CIS is operating.

As noted earlier, the purpose of the comparison schools is to capture the extent to which the program schools' outcomes would have improved had they not implemented the CIS model (the "counterfactual"). Two strategies were used to maximize the likelihood that the comparison schools' deviations from their trend provide the right counterfactual. First, matching methods were used to select comparison schools whose baseline characteristics and trends in school outcomes were as similar as possible to those of the CIS schools. Doing so helps to ensure that the two groups of schools had similar incentives to improve their students' outcomes. Second, the comparison schools are all non-CIS schools in Texas and North Carolina counties where there was a CIS presence, so the comparison schools are located in the kinds of districts where CIS typically operates.

Ideally, it would have been preferable to match CIS schools with non-CIS schools in the same districts, to ensure that both groups of schools were subject to same district-wide educational initiatives. However, it was not possible to do so because there were generally too few similar comparison schools in the same district from which to choose. However, some studies have found that, even when comparison schools are not located in the same district as program schools, a CITS design can in some circumstances replicate the results of a randomized experiment.[13] Another recent study has found that matching schools using baseline measures (focal matching) can actually come closer to replicating the findings of a randomized experiment than matching within districts (local matching).[14]

In this study, each of the 53 CIS schools was matched to the two non-CIS schools in the matching pool with the most similar trends in the school years before CIS implementation, as well as the most similar structural and demographic characteristics in the year immediately before the CIS model launched.[15] In situations where it was not possible to find good matches in

---

[13]St. Clair, Cook, and Hallberg (2014); Somers, Zhu, Jacob, and Bloom (2013); Schneeweiss et al. (2004); Fretheim et al. (2013).

[14]Hallberg, Wong, and Cook (2016).

[15]Specifically, the study looked for comparison schools that had similar baseline trends in dropout rates, attendance rates, and state test scores, and the following school characteristics in the last baseline year: the number of students enrolled at the school, the demographic characteristics of those students (their racial or ethnic composition and the percentage eligible for free or reduced-price lunches), the school's Title I status, and whether the school was rural or urban. (Title I is the federal funding stream designated for schools serving

(continued)

both baseline trends *and* structural and demographic characteristics, the study made a priority of baseline trends. The matching process was conducted separately in each state and separately for elementary schools, middle schools, and high schools.[16] Note that the total number of comparison schools (78 schools) is less than twice the number of CIS schools because some comparison schools are matched to more than one CIS school.[17] Appendix C provides further information about the selection of the comparison schools.

Baseline Equivalence and Sensitivity Analyses

As explained earlier, the CITS design assumes that the comparison schools' deviation from their baseline trend represents what the deviations from trend for the CIS schools would have been had they not implemented the CIS model. This assumption is more credible if the CIS schools and the comparison schools were similar before the CIS model was launched.

To gauge whether or not they were similar, the *magnitude* of the measurable differences between CIS schools and comparison schools is more important than the statistical significance of those differences. A useful rule of thumb is that differences should not exceed an effect size of 0.25 standard deviations.[18] An effect size is the difference in values for an outcome measure or characteristic expressed as a proportion of the standard deviation for that outcome measure or characteristic.[19]

Tables 2.4 to 2.6 examine the characteristics and outcome measures of the CIS schools and their comparison schools in the school year immediately before the start of CIS implementation. Table 2.4 focuses on high schools, Table 2.5 on middle schools, and Table 2.6 on elementary schools.

---

low-income students.) In North Carolina, it was also possible for the study to look for comparison schools that had similar numbers of suspensions per 1,000 students and crimes per 1,000 students.

[16]Even though the study did not aim to match schools within districts, 14 percent of comparison schools ended up being matched to a CIS school in the same district.

[17]Standard errors and hypothesis tests are adjusted to account for the fact that some comparison schools are included in the analysis multiple times. See Appendix E for details on the analysis.

[18]This benchmark is based on studies showing that quasi-experimental designs are more likely to produce biased results when baseline differences exceed this threshold. See Ho, Imai, King, and Stuart (2007). This is the criterion used by the What Works Clearinghouse, which is the clearinghouse for education research hosted by the Institute for Education Sciences in the U.S. Department of Education. See What Works Clearinghouse (2014a).

[19]For student outcome measures and characteristics (such as dropout rates or free or reduced-price lunch status), ideally the student-level standard deviation should be used to calculate effect sizes; for school characteristics (such as enrollment), the school-level standard deviation is most appropriate.

Table 2.4

Characteristics and Outcome Measures of CIS Schools and Comparison Schools in the Last Baseline Year, High Schools

Baseline Measure	CIS Schools	Comparison Schools	Estimated Difference		Effect Size	P-Value
Structural and demographic characteristics						
Free or reduced-price lunch (%)	52.9	42.9	9.9	***	0.20	0.001
Racial/ethnic composition (%)						
Black	32.7	29.5	3.2		0.07	0.588
Hispanic	40.0	34.7	5.3		0.11	0.297
White	21.6	31.2	-9.6	**	-0.22	0.043
Other	5.8	4.7	1.1		0.05	0.366
Enrollment	2,295.4	2,159.7	135.7		0.12	0.460
Pupil/teacher ratio	15.7	15.6	0.1		0.05	0.851
Title I status[a] (%)	35.7	35.7	0.0		0.00	NA
Location (%)						
City/town	78.6	78.6	0.0		0.00	NA
Rural	21.4	21.4	0.0		0.00	NA
Student outcome measures						
Graduation rate (%)	68.1	76.9	-8.9	*	-0.20	0.061
Dropout rate (%)	5.1	3.5	1.5	***	0.08	0.004
Attendance rate (%)	92.6	92.9	-0.4		-0.26	0.193
Proficient on state tests (%)						
ELA	78.2	81.7	-3.5	*	-0.09	0.089
Math	51.7	54.6	-2.9		-0.06	0.423
State test scores (z-scores)[b]						
ELA	-0.19	-0.09	-0.09		-0.09	0.143
Math	-0.18	-0.11	-0.07		-0.07	0.289
Number of schools	14	18				

(continued)

Table 2.4 (continued)

SOURCES: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the column labeled "CIS Schools" are the observed means in the last baseline year for schools that implemented CIS. The "Comparison Schools" values in the next column are the observed means in the last baseline year for matched comparison schools. The values in the "Estimated Difference" column are the differences between CIS and comparison schools in the last baseline year. Values in the "Effect Size" column are the estimated difference divided by the standard deviation of the full study sample in the last baseline year. Rounding may cause slight discrepancies in calculating sums and differences. The percentages of Title I and urban schools are exactly the same among the CIS and comparison schools because CIS and comparison schools were matched exactly with respect to their Title I statuses and their locations (rural/urban).

A two-tailed test was applied to estimated differences between CIS schools and comparison schools. The statistical significance of estimated differences is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

NA = not applicable.

[a]This percentage represents the Title I status of study schools in the last baseline year. By the third follow-up year, the percentage of Title I-eligible schools had risen. Among high schools, 79 percent of CIS schools and 71 percent of comparison schools were eligible for Title I funds by the third follow-up year.

[b]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

With respect to their structural and demographic characteristics, the CIS and comparison schools in this study are alike in many respects. For example, the CIS and comparison schools have identical Title I statuses and locations (city/town versus rural).[20] Although the CIS schools in the study serve a higher proportion of economically disadvantaged students than the comparison schools, these differences are less than 0.25 standard deviations. The only school characteristics where the baseline differences exceed 0.25 standard deviations are school enrollment and pupil/teacher ratio.[21]

[20]The percentage of Title I and urban schools is exactly the same among the CIS and comparison schools, because comparison schools were selected that exactly matched CIS schools' Title I statuses and urban character.

[21]These two characteristics are not used by the What Works Clearinghouse to rate baseline equivalence in studies of dropout-prevention programs. See What Works Clearinghouse (2014b). Appendix D provides standard deviations that can be used to convert baseline differences to effect sizes.

Table 2.5

Characteristics and Outcome Measures of CIS Schools and Comparison Schools in the Last Baseline Year, Middle Schools

Baseline Measure	CIS Schools	Comparison Schools	Estimated Difference	Effect Size	P-Value
Structural and demographic characteristics					
Free or reduced-price lunch (%)	56.2	50.0	6.2 *	0.12	0.060
Racial/ethnic composition (%)					
Black	26.5	28.0	-1.6	-0.03	0.681
Hispanic	38.9	34.2	4.7	0.10	0.317
White	30.2	33.2	-3.0	-0.07	0.465
Other	4.4	4.5	-0.1	-0.01	0.901
Enrollment	776.1	943.7	-167.6 *	-0.58	0.085
Pupil/teacher ratio	14.0	15.8	-1.8 ***	-0.87	0.007
Title I status[a] (%)	46.7	46.7	0.0	0.00	NA
Location (%)					
City/town	93.3	93.3	0.0	0.00	NA
Rural	6.7	6.7	0.0	0.00	NA
Student outcome measures					
Attendance rate (%)	95.4	95.7	-0.2	-0.27	0.163
Proficient on state tests (%)					
ELA	78.5	78.2	0.3	0.01	0.828
Math	67.3	66.6	0.7	0.01	0.739
State test scores (z-scores)[b]					
ELA	-0.20	-0.16	-0.04	-0.04	0.402
Math	-0.10	-0.12	0.02	0.02	0.739
Number of schools	15	24			

(continued)

Table 2.5 (continued)

SOURCES: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the column labeled "CIS Schools" are the observed means in the last baseline year for schools that implemented CIS. The "Comparison Schools" values in the next column are the observed means in the last baseline year for matched comparison schools. The values in the "Estimated Difference" column are the differences between CIS and comparison schools in the last baseline year. Values in the "Effect Size" column are the estimated difference divided by the standard deviation of the full study sample in the last baseline year. Rounding may cause slight discrepancies in calculating sums and differences. The percentages of Title I and urban schools are exactly the same among the CIS and comparison schools because CIS and comparison schools were matched exactly with respect to their Title I statuses and their locations (rural/urban).

A two-tailed test was applied to estimated differences between CIS schools and comparison schools. The statistical significance of estimated differences is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

NA = not applicable.

[a]This percentage represents the Title I status of study schools in the last baseline year. By the third follow-up year, the percentage of Title I-eligible schools had risen. Among middle schools, 79 percent of CIS schools and 67 percent of comparison schools were eligible for Title I funds by the third follow-up year.

[b]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

With respect to their baseline values for outcome measures, the two groups of schools are similar in some respects but different in others. The CIS and comparison elementary, middle, and high schools all had similar percentages of students who were proficient on state tests. However, in the school year before CIS was launched, the CIS high schools performed worse on two primary outcome measures: graduation rates and dropout rates. The dropout rate was 5.1 percentage points in the CIS high schools, compared with 3.5 percent in the comparison schools; the graduation rate in CIS high schools was 68 percent, compared with 77 percent in the comparison schools.

These differences in baseline graduation and dropout rates are less than 0.25 standard deviations, and therefore they satisfy the criterion for baseline equivalence.[22] However, in practical terms, the differences are large enough that they could give schools different incentives to turn themselves around. The CIS schools, faced with low and rapidly declining graduation

[22]The effect size is 0.20 standard deviations for graduation rates and 0.08 for dropout rates.

Table 2.6

Characteristics and Outcome Measures of CIS Schools and Comparison Schools in the Last Baseline Year, Elementary Schools

Baseline Measure	CIS Schools	Comparison Schools	Estimated Difference		Effect Size	P-Value
Structural and demographic characteristics						
Free or reduced-price lunch (%)	78.6	74.0	4.6	***	0.11	0.007
Racial/ethnic composition (%)						
Black	25.1	26.7	-1.5		-0.04	0.317
Hispanic	55.5	53.1	2.4		0.05	0.370
White	15.1	15.5	-0.4		-0.01	0.871
Other	4.2	4.7	-0.5		-0.03	0.594
Enrollment	616.8	617.2	-0.4		0.00	0.993
Pupil/teacher ratio	12.1	11.4	0.6	**	0.11	0.042
Title I status (%)	100.0	100.0	0.0		0.00	NA
Location (%)						
City/town	100.0	100.0	0.0		0.00	NA
Rural	0.0	0.0	0.0		0.00	NA
Student outcome measures						
Attendance rate (%)	95.9	96.2	-0.3	**	-0.34	0.013
Proficient on state tests (%)						
ELA	78.3	79.8	-1.6	**	-0.04	0.033
Math	74.2	74.6	-0.4		-0.01	0.749
State test scores (z-scores)[a]						
ELA	-0.33	-0.30	-0.02		-0.02	0.388
Math	-0.18	-0.21	0.03		0.03	0.411
Number of schools	24	36				

(continued)

Table 2.6 (continued)

SOURCES: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the column labeled "CIS Schools" are the observed means in the last baseline year for schools that implemented CIS. The "Comparison Schools" values in the next column are the observed means in the last baseline year for matched comparison schools. The values in the "Estimated Difference" column are the differences between CIS and comparison schools in the last baseline year. Values in the "Effect Size" column are the estimated difference divided by the standard deviation of the full study sample in the last baseline year. Rounding may cause slight discrepancies in calculating sums and differences. The percentages of Title I and urban schools are exactly the same among the CIS and comparison schools because CIS and comparison schools were matched exactly with respect to their Title I statuses and their locations (rural/urban).

A two-tailed test was applied to estimated differences between CIS schools and comparison schools. The statistical significance of estimated differences is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

NA = not applicable.

[a]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

rates, may have been very motivated to improve their students' outcomes (which is why they decided to adopt a whole-school reform model such as CIS). Their incentives to improve their graduation rates would have been especially high during the study period, due to the sanctions imposed on failing schools under No Child Left Behind.[23] Therefore, had the CIS schools not implemented the CIS model, they presumably would have implemented some other intensive reform strategy.

In contrast, the comparison schools may have had less incentive to turn themselves around, perhaps leading them to adopt less intensive strategies to improve their graduation rates. In that case, the comparison schools' deviations from trend would not provide the right information about what would have happened to the CIS schools had they not implemented the model. Any improvements in the comparison schools' graduation rates would understate the amount CIS schools' own rates would have improved without the CIS model; by extension, the

[23]Under No Child Left Behind, states were required to test students in reading and math in grades 3 to 8 and once in high school, and they were required to report the results for all students and for subgroups of students, including English language learners, students in special education, racial minorities, and children from low-income families. States were also required to achieve "proficiency" for all students by the 2013-2014 school year. (States decided what test to use to measure proficiency.) Schools had to make "adequate yearly progress" toward this goal; schools missing their annual targets for two or more years were subject to sanctions, such as having to allow students to transfer to better-performing schools, having to offer tutoring, being closed, or being restructured.

findings from this study would overstate the true effect of the CIS model. Unfortunately, because this study is retrospective, information is not available on the programs and strategies used by the comparison schools. It is not possible to determine whether dropout-prevention programs in the comparison schools were similar to or less intensive than the CIS model.

Various sensitivity tests were conducted to investigate whether the comparison schools provide the right reference point for the CIS schools, not only for graduation and dropout rates but also for the other outcomes in this study. These tests are described below. The results of these analyses, which will be described alongside the main findings, were used to assess the degree of confidence one can have that the findings in this study can be attributed to the effect of the CIS model.

General Sensitivity Tests

A series of sensitivity tests were conducted for elementary schools, middle schools, and high schools to explore several possible reasons why the comparison schools might not provide the right counterfactual for the CIS schools.

The first sensitivity test examines whether the CIS schools and the comparison schools — which are located in different school districts — were affected by different district-level policy changes. If so, then the comparison schools' deviations from their baseline trend would not provide the right counterfactual. This question was examined by comparing the trends in school outcomes in CIS schools' districts with the trends in the comparison schools' districts, to try to discern whether the two groups of districts experienced different changes in their trends.

The second sensitivity test validates the method used to choose the comparison schools by estimating the "effect" of the CIS model in the school year *before* it launched. If the test does find an "effect," it would cast doubt on the choice of comparison schools, because the CIS model was not yet operating. If the study design and the comparison-school selection approach are sound, then the CIS model's "effect" in the year before it launched should be close to zero in magnitude and not statistically significant, and estimated effects in subsequent follow-up years should be similar to the estimated effects from the main analysis (which will be presented in Chapter 3). In practice, this sensitivity test pretended that the school year before CIS implementation (the last baseline year) was the first follow-up year in the CITS design. New comparison schools were then selected using the newly defined baseline period (which now excluded the last baseline year), and the CIS model's effects on school outcomes were reestimated.

The third sensitivity test considers whether the CIS schools and the comparison schools experienced different shifts in their demographic compositions after the CIS model launched. If so, then the comparison schools may not provide the right counterfactual for the CIS schools.[24] To investigate this possibility, a CITS design was used to estimate the CIS model's "effect" on the demographic compositions of the study schools. If the estimated effect is small and not statistically significant, then the two groups of schools did not experience different demographic shifts in the follow-up period. As an extension of this analysis, the effect of the CIS model was also reestimated controlling for the average characteristics of students in the study schools in each year of the study, and for the schools' structural characteristics.[25] If the CIS model's estimated effects in this sensitivity analysis are similar to the estimated effects from the main analysis, it would further indicate that the findings from this study are not confounded with differential demographic shifts.

The fourth sensitivity test analyzes whether the overall CITS results change if the "similarity" between CIS schools and non-CIS schools is measured differently when choosing the comparison schools. If the results do change, then the comparison schools would be a less credible counterfactual. In this study, the similarity between a CIS school and its potential comparison schools is based on the Euclidian distance, a metric that captures the differences between two schools across several matching variables (that is, school baseline characteristics and values on outcome measures). For this sensitivity test, comparison schools were rechosen using propensity score matching instead (another method of measuring the "similarity" between schools), and the impact of CIS was reestimated using these new comparison schools. If the estimated effects from this analysis are similar to the main analysis findings, then the comparison schools in the main analysis may provide the right counterfactual.

Outcome-Specific Sensitivity Tests

As discussed earlier, large baseline differences in schools' outcome measures may give them different incentives to turn themselves around. Therefore, additional sensitivity tests were conducted for the outcome measures where baseline differences are large either from a practical perspective, or based on commonly used thresholds for baseline equivalence (greater than 0.25 standard deviations). Three outcome measures qualify as having "large" differences using these criteria: high school graduation and dropout rates (where baseline differences are substantively large) and elementary school attendance rates (where baseline differences are not substantively large, but are larger than 0.25 standard deviations). Additional sensitivity tests specific to these three outcome measures were conducted to explore whether the comparison schools may have

[24]This problem is sometimes referred to as "selection bias."

[25]Specifically, the analysis controlled for the following characteristics for each school year in the study: the percentage of a school's students in each racial or ethnic group, the percentage eligible for free or reduced-price lunches, and the percentage of male students, as well as a school's enrollment and its pupil/teacher ratio.

had less incentive to improve their students' outcomes.[26] Details of these outcome-measure-specific analyses are discussed in the next chapter alongside the main study findings for these outcomes.

[26]These extra analyses were not conducted for the other outcomes in the study, because baseline differences in these outcomes are substantively small and are less than 0.25 standard deviations. The baseline difference in attendance rates for middle schools is 0.26 standard deviations and for high schools it is 0.27 standard deviations, but these effect sizes are calculated using the school-level standard deviation in attendance rates (and not the student-level standard deviation), which artificially inflates the effect size. If the student-level standard deviation were known and available, then these effect sizes would probably be lower than 0.25. For elementary schools the baseline difference is 0.34 standard deviations. That effect size is also artificially inflated, but in this case it is not clear that using the student-level standard deviation would lower it below 0.25.

Chapter 3

The Effects of the Communities In Schools Model

This chapter examines whether the Communities In Schools (CIS) model was successful at increasing on-time high school graduation rates and reducing dropout rates. The chapter also examines whether the CIS model improved average attendance rates in the study's high schools, middle schools, and elementary schools, and whether it improved students' academic performance on state tests.

The central finding of this study is that on-time graduation rates increased — and dropout rates decreased — in the study schools after the CIS model was launched. However, it is not clear whether these improvements were caused, in whole or in part, by the CIS model. On the one hand, graduation and dropout rates improved more in the CIS high schools than in the comparison high schools, which suggests that the CIS model may have improved these schools' graduation and dropout rates more than they would have improved otherwise. On the other hand, the comparison schools had substantially higher baseline graduation rates (and lower dropout rates) than the CIS schools, so they may not truly represent what would have happened to the CIS schools had they not implemented the CIS model. Some sensitivity tests suggest that the comparison schools do provide a credible reference point, while others do not. It is therefore not possible to determine whether the CIS model was *more* effective than the strategies used by the comparison schools. The study's findings do suggest that the CIS model may be at least as effective as these other approaches.

For younger students, the results are quite different. In elementary schools, attendance rates (a central outcome measure the CIS model aims to affect in the elementary grades) improved in CIS schools more than they did in comparison schools. This finding suggests that the CIS model may have improved elementary school attendance more than it would have improved otherwise. It was not possible to evaluate the effect of the model on its core objective in middle schools: to improve students' behavioral outcomes. State test scores did not improve in middle schools implementing the CIS model, whereas they did improve in comparison middle schools, which suggests that the CIS model was less effective at improving middle school state test scores than were other available strategies or reforms.

The remainder of this chapter discusses these findings in greater detail. The first section describes the framework used to assess the degree of confidence one can have that the findings in this study can be attributed to the effect of the CIS model. The following sections then describe the CIS model's effects in elementary schools, middle schools, and high schools during the first three years of CIS implementation.[1]

Presenting and Interpreting the Findings from This Study

In this chapter, the comparative interrupted time series (CITS) design is used to answer two questions:

- *After launching the CIS model, did the CIS schools have better outcomes than predicted by their baseline trend?* The answer to this question is found in CIS schools' deviations from their baseline trends. For all outcome measures except the dropout rate, positive deviations from trend indicate that schools performed better during the follow-up period than their baseline trends predicted, while negative deviations indicate that schools performed worse than predicted. For dropout rates the pattern is reversed: A negative deviation indicates that schools performed better than their baseline trend predicted.
- *Did the CIS schools' outcomes improve more than they would have improved had they not implemented the CIS model?* The answer to this question is found in the *differences* between CIS schools' deviations and comparison schools' deviations from their respective baseline trends. These differences represent the estimated net effect of the CIS model. For all outcome measures except the dropout rate, positive estimated effects suggest that by implementing the CIS model, CIS schools' outcomes improved more than they would have otherwise. For dropout rates, negative estimated effects suggest that CIS schools' dropout rates decreased more than they would have otherwise.

In this chapter, two criteria are used to assess whether the CIS model appears to have had an effect on school outcomes. First, the CIS model's estimated effect on a given outcome measure should be *statistically significant*. Statistical significance is described using the p-value, which indicates the probability of finding the estimated effect (or a larger one) if there were in fact no difference between the CIS schools and the comparison schools. In this study, an estimated effect that has a p-value of 10 percent or less is considered statistically significant.

[1]Appendix F also presents supplemental findings on the CIS model's effect on study schools in Texas, where the sample of schools is large enough to consider effects only for that state.

Second, the CIS model's estimated effect on an outcome measure must be *credible*. To be credible, estimated effects should be consistent in all three follow-up years. As noted earlier, if the model is effective, then one would expect its effect to grow over the first few years of implementation, because students are receiving cumulatively more services and because the model's implementation may be getting stronger. Thus, the CIS model's effect should remain stable — or become larger — over time. A pattern of effects of this kind not only makes the findings more credible, but also indicates whether the CIS model's effects are sustained over time (an important practical consideration for practitioners and policymakers).

As discussed in the last chapter, the credibility of the results was also assessed by conducting a suite of general and outcome-measure-specific sensitivity tests, whose purpose was to examine the extent to which the comparison schools provide a credible counterfactual for the CIS schools. These tests investigate several plausible factors that could be confounded with program effects. If some of these factors did seem to be playing a role in the results, then one would have less confidence that the estimated effects from this study could be attributed to the CIS model. The results of the sensitivity tests are discussed alongside the main results.

Because the CITS design is often better understood visually, this chapter's findings are presented in figures as well as tables. See Box 3.1 for an explanation of how to read the tables in this chapter. Bullets are used throughout the chapter to summarize findings related to the two research questions outlined above.

Findings for High Schools

Graduation and Dropout Rates

- **After three years, on-time graduation rates and dropout rates improved by statistically significant amounts in the CIS high schools, relative to what would have been expected given their baseline trends.**

- **It is not possible to determine the extent to which these improvements can be attributed to the CIS model. Graduation and dropout rates improved more in the CIS schools than in the comparison high schools, but this pattern of results depends on the choice of comparison schools.**

As explained in Chapter 2, this study estimates the CIS model's effects on both graduation rates and dropout rates, because neither measure is perfect. Graduation rates are the preferred measure, but they are only available for high schools in Texas; conversely, dropout rates are available for all study schools, but they are less useful because of how they are calculated.

Box 3.1

How to Read the Impact Tables in This Report

The tables in this chapter present the CIS model's estimated effects in each follow-up year (Year 1, 2, and 3 of CIS implementation). To put the pattern of effects in context, the tables also show the extent to which CIS and comparison schools' actual outcomes deviated from the outcomes predicted by their baseline trends, since these "deviations from trend" are used to calculate the effect of the CIS model.

The columns in the tables present the following information:

- **Estimated deviations from baseline trend:** The first two columns of results are the estimated deviations from their baseline trends of the CIS schools and their matched comparison schools. For all outcomes except dropout rates, positive numbers indicate that schools in the CIS or comparison group did better in the follow-up period than their baseline trends predicted, while negative numbers indicate they did worse than expected. For dropout rates, the reverse is true. The statistical significance of these estimated deviations from trends is indicated for levels of 10 percent (†), 5 percent (††), and 1 percent (†††).

- **Estimated effect:** The third column of results is the CIS model's estimated effect. The estimated effect is the difference between CIS schools' and comparison schools' deviations from their baseline trends (column 1 minus column 2).

- **Confidence intervals:** The next two columns in the table are the lower and upper limits of the 90 percent confidence interval for the CIS model's estimated effect. Generally speaking, the true impact of the CIS model has a 90 percent probability of falling somewhere within this interval.

- **P-value for the estimated effect:** The last column of the table shows the p-value for each estimated effect. The p-value indicates the probability of finding the estimated effect (or a larger one) if there were in fact no difference between the CIS schools and the comparison schools with respect to their deviations from their baseline trends. In this study, an estimated effect with a p-value of 10 percent or less is considered "statistically significant" because it is unlikely that this estimated effect would be observed if the CIS schools and comparison schools did not have different deviations from their trends. The asterisks beside the estimated effect column indicate whether the estimated effect is statistically significant at the 10 percent (*), 5 percent (**), or 1 percent (***) level.

The rows in the tables represent outcomes for each implementation year (Year 1, Year 2, and Year 3 of CIS implementation).

Figure 3.1 shows the trend in graduation rates for the CIS and comparison high schools in the school years before and after the implementation of the CIS model.[2] The baseline trends in the figure show that before CIS was implemented in the study schools, graduation rates in both the CIS and comparison high schools had been decreasing steadily — by about 2 percentage points to 3 percentage points per school year. In the follow-up period, the trend in graduation rates changed direction and rates started increasing in both CIS and comparison schools. However, graduation rates started increasing sooner — and more — in the CIS high schools.[3]

Figure 3.2 compares the estimated deviations from baseline trend for the two groups of schools in each follow-up year. As summarized in this bar chart, estimated deviations from trend are positive for CIS schools in all three follow-up years, indicating that graduation rates were higher than their baseline trends predicted. The comparison schools also had higher graduation rates than predicted, but their deviations from their baseline trend are smaller and do not become statistically significant until the third follow-up year.

Table 3.1 presents more detailed results. The CIS model's estimated effect on graduation rates is positive in all three follow-up years and it is increasing over time. By the third year of implementation, the estimated effect on graduation rates is 7.5 percentage points, which is statistically significant at the 10 percent level (p-value = 0.088). Before CIS began operating in these high schools, their average graduation rate was 68 percent, so an effect of 7.5 percentage points represents a one-tenth (11 percent) increase in the graduation rate relative to baseline levels. An effect of this size means that in the third year of implementation, 55 additional students in each study school (two extra classrooms of students) would have graduated from high school on time.[4]

One noteworthy result from this study is that graduation rates also increased in the comparison schools in the follow-up period, especially in the third follow-up year. There are two plausible explanations for this improvement. First, graduation rates may have increased due to the Great Recession, which began at the end of 2007. Recessions typically bring about an increase in graduation rates, because there are fewer opportunities in the labor market, and

---

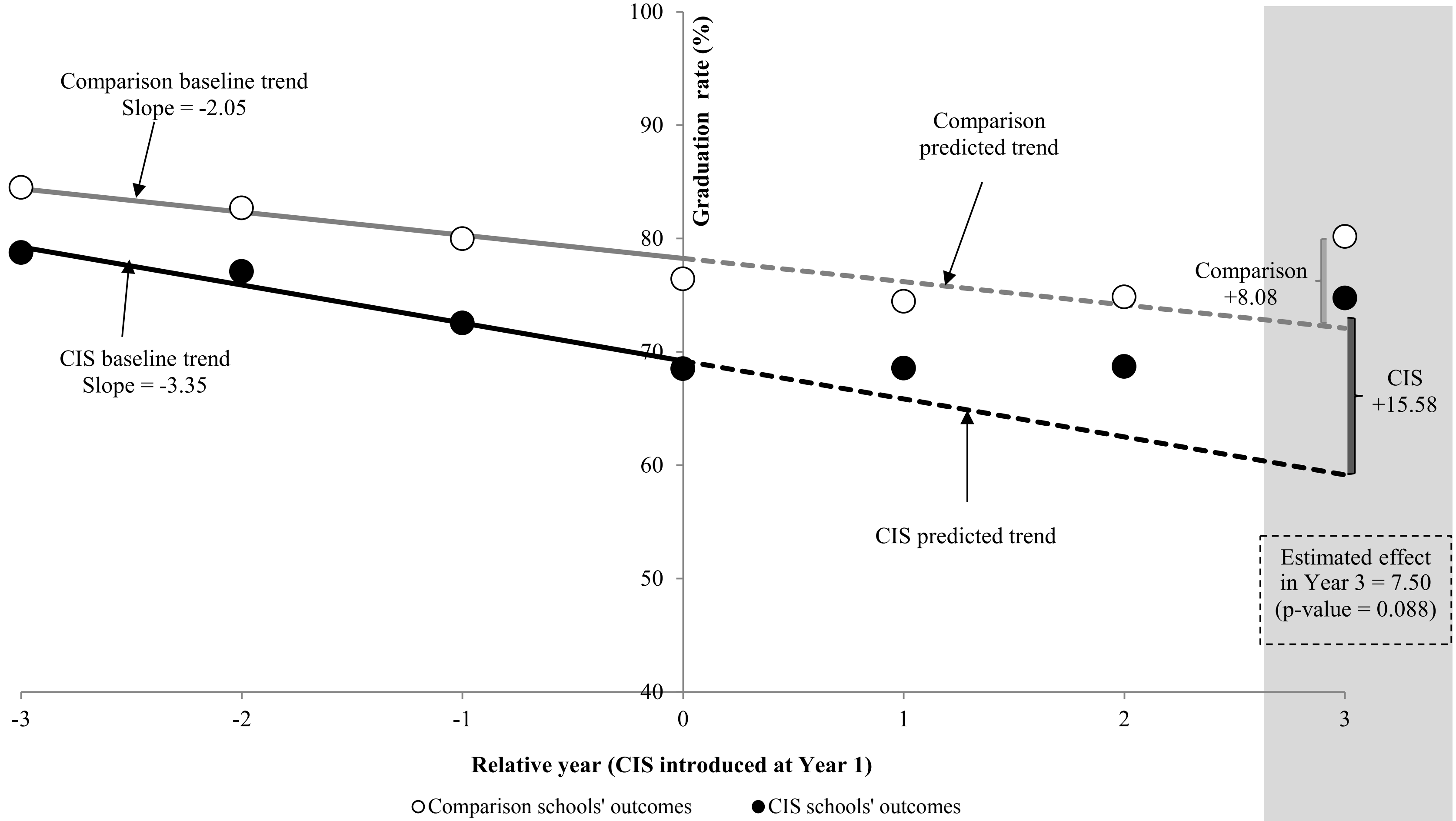
[2]Information on the estimated baseline slopes and intercepts used to construct these figures can be found in Appendix A.

[3]It is important to note here that the baseline trend in graduation rates is precisely estimated: The graduation rates of schools in the baseline period fluctuate tightly around the baseline trend. As a result it is easier to see that the deviations from trend in the follow-up period are real departures and not simply due to noise (an observation confirmed by the fact that the estimated deviations from trend are statistically significant in some follow-up years, as shown in Table 3.1).

[4]Based on a ninth-grade class size of 735 students, the average ninth-grade enrollment in study high schools before CIS was launched.

Figure 3.1

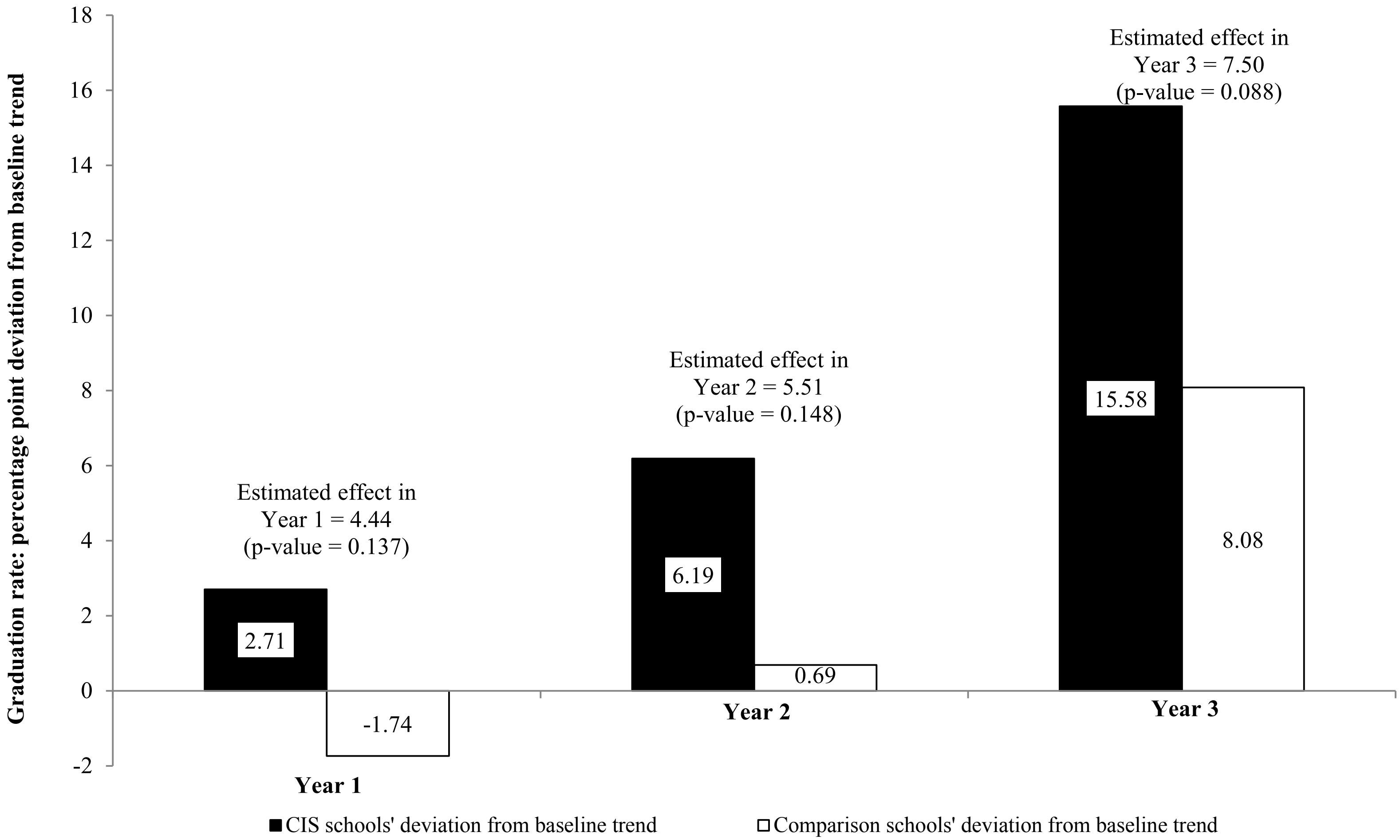
Baseline Trends and Deviations from Trends in High School Graduation Rates



SOURCES: MDRC calculations based on school-level data on student outcomes from the Texas Education Agency.

Figure 3.2

Deviations from Trends and Estimated Effects on High School Graduation Rates



SOURCES: MDRC calculations based on school-level data on student outcomes from the Texas Education Agency.

Table 3.1

Estimated Effects on Dropout and Graduation Rates, High Schools

	Deviation from Baseline Trend								
Outcome and Follow-Up Year	CIS Schools		Comparison Schools		Estimated Effect		Lower 90% CI	Upper 90% CI	P-Value for Estimated Effect
Graduation rate (%)									
Year 1	2.71	†	-1.74		4.44		-0.48	9.37	0.137
Year 2	6.19	†††	0.69		5.51		-0.77	11.78	0.148
Year 3	15.58	†††	8.08	††	7.50	*	0.27	14.73	0.088
Number of schools	8		11						
Dropout rate (%)									
Year 1	-0.75		0.30		-1.06		-2.40	0.29	0.197
Year 2	-1.85	††	-0.65		-1.19		-2.75	0.36	0.207
Year 3	-3.77	†††	-2.29	†††	-1.49		-3.44	0.47	0.211
Number of schools	14		18						

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated deviations from baseline trend for each group of schools. The values in the "Estimated Effect" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. The values in the "Lower 90% CI" and "Upper 90% CI" columns are the 90 percent confidence intervals for the estimated effects. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed test was applied to estimated deviations and estimated differences between CIS schools and comparison schools. The statistical significance of estimated deviations is indicated as: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

students therefore have less incentive to leave school.[5] By the third follow-up year, all of the study schools' follow-up periods overlap with Great Recession, which may explain why the comparison schools' graduation rates were higher than expected especially in that year.[6]

[5]Graduation rates started increasing beginning in 2007-2008 (after the start of the Great Recession), in schools as a whole — not only those in the study — in both Texas and North Carolina.

Second, the comparison schools may have adopted measures to improve their graduation rates. As Figure 3.1 clearly shows, graduation rates had been decreasing in the baseline period, which may have pushed schools to adopt various strategies to help students stay in school and graduate. The CIS schools in the study did just that by deciding to implement the CIS model. Similarly, the comparison schools in the study may have chosen some other set of strategies with the same goal in mind.[7] However, as noted earlier, the comparison schools' graduation rates improved by a smaller amount than the CIS schools' graduation rates.

The CIS model's estimated effects on dropout rates follow a pattern that is the inverse of the pattern of effects on graduation rates. Figure 3.3 shows the trend in dropout rates for CIS and comparison high schools during the school years before and after the CIS model was introduced. In the period before the CIS model was implemented in the study's high schools, dropout rates were increasing. In the follow-up period, dropout rates started declining in both CIS and comparison schools, but they started decreasing sooner — and more — in the CIS high schools.

Figure 3.4 compares the estimated deviations from baseline trend for the two groups of schools in each follow-up year. As shown in this chart, estimated deviations from trend are negative for CIS schools in all three follow-up years, indicating that dropout rates were lower than their baseline trends predicted. The comparison schools also had lower dropout rates than predicted beginning in the second follow-up year, but their deviations from trend are smaller in size and do not become statistically significant until the third follow-up year. As noted earlier, comparison schools' dropout rates may have improved because of the Great Recession, or because comparison schools adopted alternative strategies or initiatives to help students stay in school.

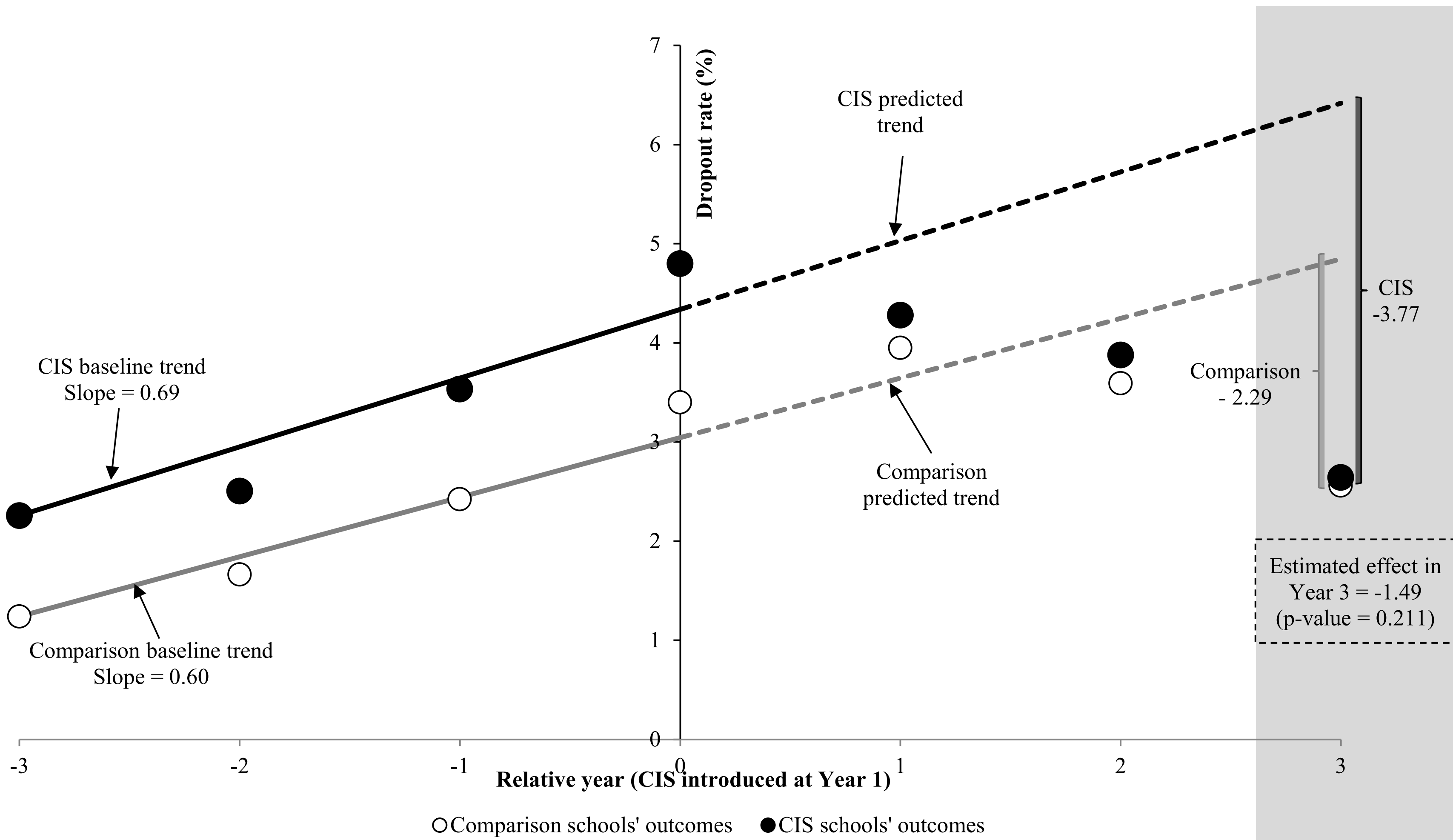
Although the CIS model's estimated effects on dropout rates are not statistically significant, they are consistently negative in the first three follow-up years and their size increases over time (Figure 3.4 and Table 3.1). By the third year of CIS implementation, the estimated effect on dropout rates is -1.5 percentage points (p-value = 0.211). Because dropout rates were

[6]High schools in this study started implementing the CIS model in different school years, from 2005-2006 to 2008-2009, so some schools launched their CIS programs before the Great Recession. Therefore, in the first and second follow-up years, not all schools would have yet been affected by the recession.

[7]As noted in Chapter 2, in order for the CITS design to correctly estimate program effects, the comparison schools' deviations from their baseline trend must represent what the CIS schools would have experienced had they not implemented the CIS model. By extension, it must be assumed here that had CIS schools *not* chosen to implement the CIS model, they would instead have chosen another reform model (or another set of strategies) similar to those used by comparison schools.

Figure 3.3

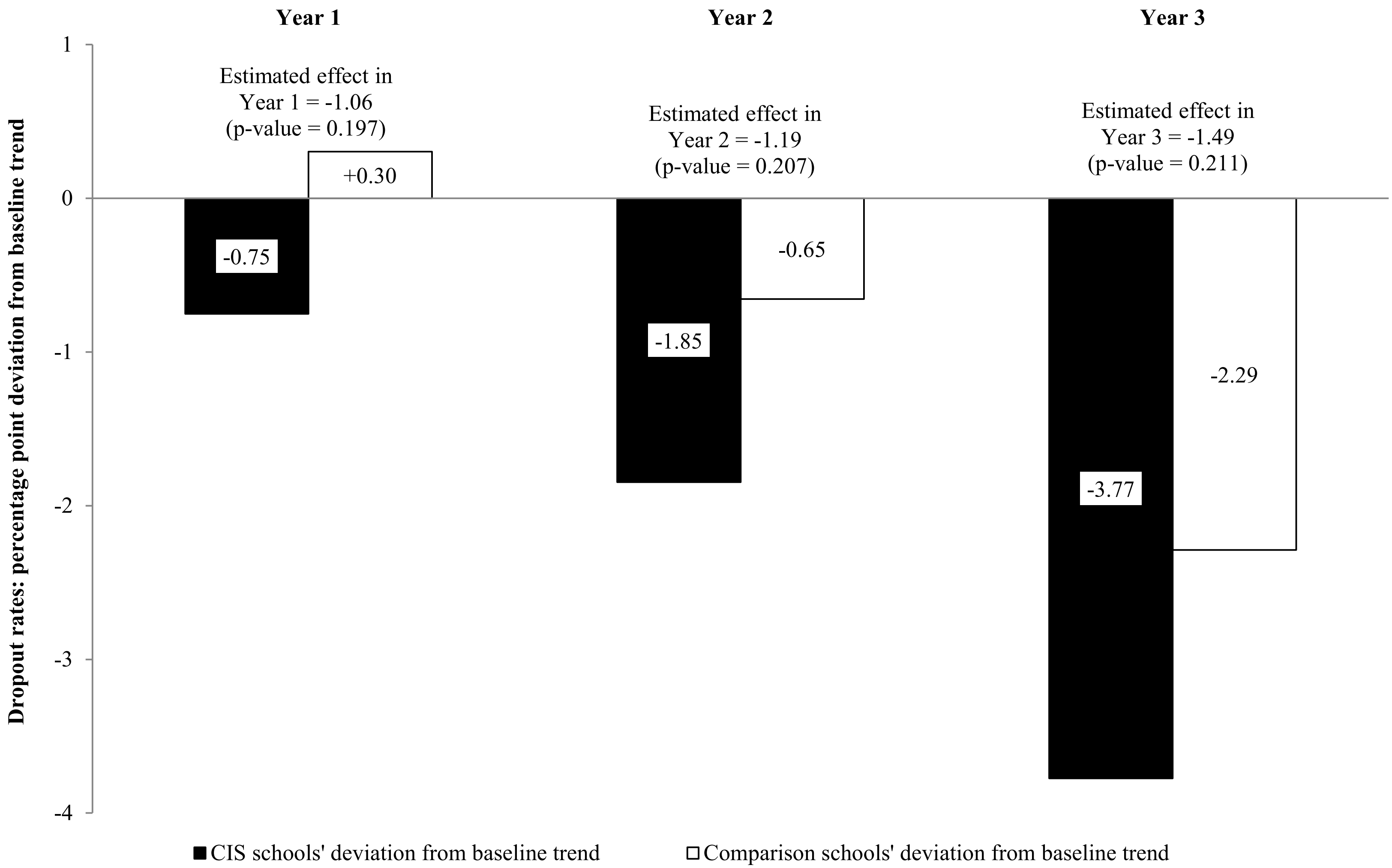
Baseline Trends and Deviations from Trends in High School Dropout Rates



SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

Figure 3.4

Deviations from Trends and Estimated Effects on High School Dropout Rates



SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

already low in CIS schools before CIS was implemented (5.1 percent), an estimated decrease of 1.5 percentage points represents almost a one-third (29 percent) reduction in dropout rates.[8] On average, an effect of this size would have prevented about 35 students in each study school (more than one extra classroom of students) from dropping out in the third year of implementation.[9]

As one might expect, the estimated effects on graduation rates and dropout rates are similar in magnitude. It appears that the CIS model helped 55 extra students graduate on time and prevented a similar number, about 35 students, from dropping out. Also, an effect of 7.5 percentage points on graduation rates translates into a 23 percent reduction in "nongraduation" rates relative to baseline levels, which is similar in magnitude to the estimated effect on dropout rates (a 29 percent reduction).[10]

As explained in Chapter 2, four general sensitivity tests were conducted to explore whether the comparison schools' deviations from trend provide the right counterfactual for the CIS schools. The first sensitivity test examines whether the CIS schools and the comparison schools' outcomes were affected by different district-level policy changes. The second test validates the method used to choose the comparison schools by estimating the CIS model's effect in the school year before it was launched (which should be zero). The third sensitivity test considers whether the CIS schools and the comparison schools experienced different shifts in their demographic compositions after the CIS model launched. The fourth test analyzes whether the overall CITS results change if the "similarity" between CIS schools and non-CIS schools is measured differently when choosing the comparison schools.

---

[8]Given the number of high schools in the study, the CIS model's estimated effect would have to be as large as 2 percentage points to 3 percentage points to be statistically significant — difficult to achieve given baseline dropout rates of 5 percent. See Appendix G for further discussion of statistical power and minimum detectable effects.

[9]This number assumes an average enrollment of 2,300 students.

[10]The average on-time graduation rate in the CIS schools before CIS was launched was 68 percent, or equivalently, about 32 percent of students did *not* graduate in four years. Thus, an estimated increase in graduation rates of 7.5 percentage points (to 75.5 percent) represents a 23 percent reduction relative to the baseline nongraduation rate (a reduction from 32 percent to 24.5 percent).

As noted earlier, data on graduation rates are not available for all schools. Therefore, as a supplemental analysis, the CIS model's estimated effect on dropout rates was examined only for the subset of schools with data on graduation rates. The purpose of this analysis is to examine the extent to which the overall estimated effects on graduation rates might be larger or smaller were data on this outcome available for all of the study schools. Overall, the CIS model's estimated effect on the dropout rates of this subset of schools is about 0.5 percentage points smaller than the effect on the full study sample: -1.5 percentage points in Year 1 (p-value = 0.205), -1.8 percentage points in Year 2 (p-value = 0.156), and -2.0 percentage points in Year 3 (p-value = 0.191). These findings suggest that if data on graduation rates were available for all of the study schools, the estimated effects on graduation rates would be slightly smaller than the effects presented in this section.

Based on the results of these four tests, the comparison schools do appear to provide a credible counterfactual. In the first test, district-level trends in graduation rates are similar for the CIS districts and the comparison districts, which means that the results are not confounded with changes in district-level policies. In the second, the CIS model's estimated effects on graduation and dropout rates in the last baseline year are close to zero, while estimated effects in the subsequent follow-up years are very similar in size to the main analysis findings.[11] Third, the estimated effects on graduation and dropout rates are similar when the analysis controls for student composition, which suggests that the findings are not confounded with changes in student composition.[12] Finally, the estimated effect on graduation rates is similar when another popular matching method (propensity score matching) is used to choose the comparison schools.[13] Appendix E provides further information on the sensitivity analysis results.

Substantively, however, there is a large difference in baseline graduation and dropout rates between the CIS schools and the comparison schools. The CIS high schools had lower and more rapidly declining graduation rates than the comparison schools (see Figure 3.1) and higher dropout rates (see Figure 3.3). Although these differences are small enough to satisfy commonly used criteria for baseline equivalence, they are large enough that the comparison schools could have felt less pressure or need to turn around their graduation rates. This difference places an additional burden on the study to demonstrate that despite their higher graduation rates, the comparison schools can still provide the right counterfactual for the CIS schools.

The difference in the baseline graduation and dropout rates of the CIS high schools and their comparison schools mostly reflects a subset of the study schools. Specifically, before they started implementing the CIS model, three CIS high schools had especially low graduation rates (50 percent on average) and high dropout rates (8.5 percent on average), and their graduation rates were rapidly decreasing. None of the non-CIS high schools in the comparison pool were performing as poorly as these three CIS schools, so these three were matched to comparison schools that were performing better already (with graduation rates of 80 percent and dropout rates of 4.1 percent on average). Conversely, the remaining CIS high schools in the study were very similar to their comparison schools. The average baseline graduation rate was 77 percent for these CIS schools and their comparison schools, and the average dropout rate was 3 percent.[14]

---

[11]For example, the estimated effect on graduation rates in the last baseline year is -1.6 percentage points and the estimated effect in Year 3 is 6.7 percentage points.

[12]The estimated effect on graduation rates is 6 percentage points in Year 3 in this sensitivity analysis.

[13]The estimated effect on graduation rates is 4.5 percentage points in Year 3 in this sensitivity analysis.

[14]See Appendix E for detailed findings.

To further explore how these subgroup differences in baseline graduation and dropout rates might be affecting the results, the CIS model's effect on these outcomes was separately estimated for the CIS high schools that were well matched to their comparison schools and those that are not. This analysis is the equivalent of conducting a subgroup analysis for CIS high schools that were performing better during the baseline period versus those that were performing worse.

The top panel of Table 3.2 shows the CIS model's estimated effects on graduation and dropout rates for the schools that are well matched and performing better. These schools provide the most convincing evidence of the effect of the CIS model because they and their comparison schools had similar baseline graduation and dropout rates. After three years of implementation, these CIS schools' graduation and dropout rates had improved, but their comparison schools' rates had improved as well, and by slightly larger amounts. Therefore, the CIS model's estimated effects on the graduation and dropout rates of this group of schools are in the wrong direction, though the effects are not statistically significant.

The bottom panel of Table 3.2 shows the results for the three lowest-performing CIS schools. Three years after launching the CIS model, these schools' average graduation rate had deviated from their baseline trend by 25 percentage points, whereas their comparison schools had continued on their downward trajectory (they did not deviate from their trend by a statistically significant amount). The CIS model's estimated effect on the graduation rate of this subgroup of schools is therefore large and statistically significant in all follow-up years. By the third year of implementation, the estimated effect is 27.5 percentage points (p-value < 0.001).[15] Similarly, the estimated effect on dropout rates is also large and statistically significant in all three follow-up years. By the third year, the CIS model is estimated to have decreased these schools' average dropout rate by 7 percentage points (p-value = 0.003).

There are two plausible (and possibly overlapping) explanations for the pattern of subgroup results in Table 3.2. The first is that the CIS model may have been more effective for the lowest-performing CIS high schools, because these schools may have benefited more from the structure and support provided by CIS. Alternatively, the estimated effect on these CIS schools may be overstated because their comparison schools may not provide the right counterfactual. As noted above, the graduation and dropout rates of this subgroup of comparison schools did not improve during the follow-up period. (Their deviations from trend are small and not statistically significant.) Therefore, the implicit assumption is that if the three lowest-performing CIS schools had *not* implemented the CIS model, their graduation and dropout rates would not have improved either. Is this a credible assumption?

[15]The estimated effect on these three high schools is different from the effect on the other five high schools by a statistically significant amount.

Table 3.2

Estimated Effects on High School Graduation and Dropout Rates, by Match Quality and Baseline Characteristics

	Deviation from Baseline Trend			P-Value for
Outcome and Follow-Up Year	CIS Schools	Comparison Schools	Estimated Effect	Estimated Effect
Schools that are well matched and higher performing				
Graduation rate (%)				
Year 1	-0.10	2.01	-2.11	0.419
Year 2	3.13	6.75 †††	-3.62	0.286
Year 3	10.42 †††	14.20 †††	-3.78	0.337
Number of schools	5	9		
Dropout rate (%)				
Year 1	-0.19	-0.03	-0.17	0.833
Year 2	-0.56	-0.80	0.24	0.792
Year 3	-2.19 ††	-2.48 †††	0.28	0.808
Number of schools	11	17		
Schools that are not well matched and are lower performing				
Graduation rate (%)				
Year 1	7.82 †††	-8.02	15.84 **	0.012
Year 2	11.94 †††	-9.40	21.34 ***	0.003
Year 3	25.03 †††	-2.43	27.46 ***	<0.001
Number of schools	3	5		
Dropout rate (%)				
Year 1	-2.45	1.48	-3.94 *	0.060
Year 2	-6.06 †††	-0.17	-5.89 ***	0.004
Year 3	-8.90 †††	-1.68	-7.22 ***	0.003
Number of schools	3	5		

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated deviations from baseline trend for each group of schools. The values in the "Estimated Effect" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed test was applied to estimated deviations and estimated differences between CIS schools and comparison schools. The statistical significance of estimated deviations is indicated as: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

Two additional analyses were conducted to explore this question. First, data from the state of Texas were used to examine the extent to which high schools with low and declining graduation rates are able to turn themselves around.[16] High schools located in school districts where CIS operates were excluded.[17] The analysis found that non-CIS high schools with low and declining graduation rates during the study period had higher graduation rates three years later, and that their graduation rates were 20 percentage points higher on average than their earlier trends would have predicted.[18] Similarly, all schools with high and rising dropout rates had lower dropout rates three years later, with an average improvement of 7 percentage points.[19]

Second, an additional sensitivity analysis was conducted to explore how the estimated effect of the CIS model would be affected if the difference in baseline graduation rates between the CIS schools and the comparison schools were smaller. To do so, a new set of comparison schools was chosen for the three lowest-performing CIS schools, this time by matching their baseline graduation rates only (as opposed to selecting matches using multiple outcome measures and school characteristics). After re-matching, there remains a large difference in baseline graduation rates between the three CIS schools and their new comparison schools, due to limitations in the pool of non-CIS schools. However, the new comparison schools have slightly more similar baseline graduation rates, and their baseline graduation rates were declining at the same rate on average as the CIS schools.[20]

When these new comparison schools are used in the analysis, the pattern of findings is quite different. Most notably, the new comparison schools *do* turn themselves around: After three years, these schools' average graduation rate had improved on their baseline trend by 12.7 percentage points. Therefore, when these new comparison schools are used as the counterfactu-

---

[16]The analysis focused on high schools in Texas that meet the following criteria: (1) between 2003 and 2008, their graduation rates were declining by amounts that were statistically significant at the 10 percent level; and (2) they had graduation rates of 65 percent or less at the end of this period. A cut-off of 65 percent was chosen because it is the highest baseline graduation rate among the three lowest-performing CIS schools.

[17]In the three CIS affiliates where the CIS high schools in this study are located (Houston, Central Texas, and Greater Central Texas), information about which specific schools implemented the CIS model is known, so only schools that implemented the CIS model were dropped from the sample. In all other CIS affiliates, all high schools in the associated school districts were excluded, because it was not possible to determine which specific schools had implemented the CIS model during the study period.

[18]For 83 percent of these schools, the improvement was statistically significant. For the others, graduation rates improved but not by statistically significant amounts.

[19]In this analysis, schools with "high" dropout rates are defined as those with dropout rates exceeding 5 percent, which is the lowest baseline dropout rate among the three lowest-performing CIS schools. For 77 percent of these schools, the improvement they experienced was statistically significant. For the others, dropout rates decreased but not by statistically significant amounts.

[20]The difference in graduation rates between CIS schools and the new comparison schools is 27 percent, down from 30 percent. The *slopes* of the baseline trends of the CIS schools and the new comparison schools are not different by a statistically significant amount.

al, the CIS model's estimated effect on the three lowest-performing CIS schools is still large, but half of its original size (an estimated effect of 12.3 percentage points, compared with 27.5 using the original comparison schools). This sensitivity analysis was also conducted for dropout rates and the pattern of findings is similar. Using the new comparison schools (chosen by matching dropout rates only), the estimated effect on dropout rates for the three lowest performing CIS schools is -2 percentage points (compared with -7 percentage points using the original comparison schools).[21]

Together, these analyses suggest that struggling high schools do typically turn themselves around, and that the graduation and dropout rates of the three lowest-performing CIS high schools would probably have improved at least somewhat had they not implemented the CIS model. Therefore, the comparison schools for the three lowest-performing CIS high schools — the comparison schools whose graduation and dropout rates did not improve — may not provide the right counterfactual. The estimated effect of 27.5 percentage points on graduation rates for the lowest-performing CIS schools probably overstates the CIS model's true effect on these schools.

Unfortunately, it is not possible to determine the right counterfactual for the lowest-performing CIS schools, and by extension, it is not possible to determine whether the CIS model improved their graduation and dropout rates more than they might have improved otherwise. As reported earlier, failing high schools in Texas turned their graduation rates around by 20 percentage points on average during the study period. However, this improvement does not necessarily represent the right counterfactual for the three lowest-performing CIS schools, because this collection of non-CIS failing schools across Texas may differ from the CIS schools in various ways (their locations, student compositions, resources, etc.). To cite one important difference, these failing high schools are all located in districts where CIS does not operate, so their local policy environments and resources may be very different from those of the CIS schools (which is why these schools were not included in the comparison pool for this study in the first place).

Overall, what the study does show is that the average graduation and dropout rates of the CIS high schools improved more than would have been expected given their prior trends. The study's findings also suggest that the CIS model may be at least as effective as the other approaches used by the comparison schools (though it cannot be determined whether they are *more* effective).

[21]See Appendix E for these results.

Attendance Rates

- **After three years, the average attendance rate improved by a statistically significant amount in the CIS high schools, relative to what would have been expected given their baseline trend.**
- **The comparison schools' average attendance rate also increased, by a similar amount. Therefore, it does not appear as though CIS schools' attendance rates improved more than they would have otherwise.**

Table 3.3 presents the CIS model's estimated effect on attendance rates. After they started implementing the CIS model, CIS high schools' attendance rates were higher than their baseline trend predicted, but the comparison schools' attendance rates were also higher than predicted, by a similar amount. Therefore, the estimated effect on this outcome is small in all follow-up years (the equivalent of less than a quarter of a day of school at most, assuming a school year of 180 days) and is not statistically significant. This pattern of results remains consistent in the four general sensitivity analyses that were conducted to examine the credibility of the comparison schools.

State Test Scores

- **After three years, English/language arts (ELA) state test scores improved by a statistically significant amount in the CIS high schools, relative to what would have been expected given their baseline trend.**
- **It is not possible to determine the extent to which these improvements can be attributed to the CIS model. ELA state test scores improved more in the CIS schools than in comparison schools, but this pattern of results depends on the choice of comparison schools.**

After they started implementing the CIS model, CIS high schools' ELA scores were higher than their baseline trend predicted, whereas the comparison schools' ELA scores were not higher than predicted (they did not deviate from their baseline trend by statistically significant amounts). Therefore, the CIS model's estimated effect on students' performance on ELA state tests is positive and statistically significant in all three follow-up years (see Table 3.3). However, these results do not remain consistent in one of the four general sensitivity analyses.[22]

[22]In the sensitivity analysis that examines the "effect" of CIS in the last baseline year, the estimated effect in the last baseline year is small and not statistically significant (as expected), but it also continues to be small in subsequent years when the CIS model was actually being implemented. See Appendix E.

Table 3.3

Estimated Effects on Attendance Rates and State Test Scores, High Schools

	Deviation from Baseline Trend							
Outcome and Follow-Up Year	CIS Schools		Comparison Schools		Estimated Effect	Lower 90% CI	Upper 90% CI	P-Value for Estimated Effect
Attendance rate (%)								
Year 1	-0.15		-0.19		0.04	-0.63	0.71	0.924
Year 2	0.31		0.30		0.00	-0.72	0.73	0.993
Year 3	0.88	††	0.76	††	0.12	-0.71	0.94	0.814
Number of schools	14		18					
ELA state test scores (z-scores)[a]								
Year 1	0.06		-0.03		0.08 **	0.02	0.15	0.046
Year 2	0.09	††	0.01		0.09 *	0.01	0.17	0.072
Year 3	0.15	†††	0.04		0.10 *	0.00	0.20	0.085
Number of schools	9		12					
Math state test scores (z-scores)[a]								
Year 1	0.02		0.01		0.01	-0.08	0.10	0.841
Year 2	0.10	††	0.05		0.05	-0.03	0.14	0.307
Year 3	0.16	†††	0.09	††	0.07	-0.04	0.17	0.322
Number of schools	9		12					

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated deviations from baseline trend for each group of schools. The values in the "Estimated Effect" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. The values in the "Lower 90% CI" and "Upper 90% CI" columns are the 90 percent confidence intervals for the estimated effects. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed test was applied to estimated deviations and estimated differences between CIS schools and comparison schools. The statistical significance of estimated deviations is indicated as: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

- **After three years, math state test scores improved by a statistically significant amount in the CIS high schools, relative to what would have been expected given their baseline trend.**
- **The comparison schools' math state test scores also increased, by a similar amount. Therefore, it does not appear as though CIS schools' math state test scores improved more than they would have otherwise.**

After they started implementing the CIS model, CIS high schools' math scores were higher than their baseline trend predicted, but the comparison schools' math scores were also higher than predicted in the third follow-up year, by a similar amount. Therefore, the CIS model's estimated effect on math scores is positive but not statistically significant (Table 3.3). This pattern of results remains consistent in the four general sensitivity analyses that were conducted to examine the credibility of the comparison schools.

Findings for Middle Schools

Attendance Rates

- **After three years, the average attendance rate of the CIS middle schools was higher than would have been expected given their baseline trend, though not by a statistically significant amount.**
- **Their comparison schools' average attendance rate was also higher, by a similar amount. Therefore, it does not appear as though CIS schools' attendance rates improved more than they would have otherwise.**

Table 3.4 presents the CIS model's estimated effect on attendance rates in the study's middle schools. CIS middle schools' average attendance rate in the follow-up period was higher than predicted, but not by a statistically significant amount. The comparison schools had an average attendance rate that was higher than predicted by a similar amount. Therefore, the estimated effects on attendance rates are small in size (the equivalent of less than half a day of school at most, assuming a school year of 180 days) and are not statistically significant. This pattern of results remains consistent in the four general sensitivity analyses that were conducted to examine the credibility of the comparison schools.

State Test Scores

- **After three years, state test scores did not improve in the CIS middle schools, relative to what would have been expected given their baseline trend.**

Table 3.4

Estimated Effects on Attendance Rates and State Test Scores, Middle Schools

	Deviation from Baseline Trend						
Outcome and Follow-Up Year	CIS Schools	Comparison Schools		Estimated Effect	Lower 90% CI	Upper 90% CI	P-Value for Estimated Effect
Attendance rate (%)							
Year 1	0.24	0.04		0.21	-0.15	0.56	0.345
Year 2	0.16	0.17		-0.01	-0.44	0.42	0.977
Year 3	0.41	0.35	††	0.06	-0.44	0.56	0.853
Number of schools	15	24					
ELA state test scores (z-scores)[a]							
Year 1	0.01	0.05	†	-0.05	-0.11	0.02	0.252
Year 2	0.02	0.10	†††	-0.08	-0.16	0.00	0.110
Year 3	0.00	0.11	†††	-0.11 *	-0.20	-0.01	0.061
Number of schools	8	15					
Math state test scores (z-scores)[a]							
Year 1	-0.01	0.07	†	-0.08	-0.16	0.01	0.125
Year 2	0.01	0.08	††	-0.07	-0.17	0.03	0.243
Year 3	-0.01	0.09	†	-0.10	-0.22	0.03	0.194
Number of schools	8	15					

SOURCE: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated deviations from baseline trend for each group of schools. The values in the "Estimated Effect" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. The values in the "Lower 90% CI" and "Upper 90% CI" columns are the 90 percent confidence intervals for the estimated effects. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed test was applied to estimated deviations and estimated differences between CIS schools and comparison schools. The statistical significance of estimated deviations is indicated as: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

- **During the same time period, state test scores did improve by statistically significant amounts in the comparison schools. The CIS middle schools' test scores appear to be lower than they would have been otherwise.**

Figure 3.5 plots the trends in ELA state test scores for all CIS middle schools and their comparison schools, in the school years before and after the CIS model was implemented. Before CIS was launched, test scores in both the CIS schools and the comparison middle schools were steadily decreasing. In the follow-up period, the comparison schools' scores took a turn for the better: These schools' scores started to increase, and they performed better than predicted on ELA state tests. The comparison schools, faced with declining test scores, may have adopted strategies to improve students' performance on the state tests. However, CIS schools did *not* improve. Their ELA scores continued to decrease as predicted by their baseline trend.

As a consequence, the CIS model's estimated effect on ELA state test scores is negative in all three follow-up years and becomes more negative over time (Table 3.4). By the third implementation year, the statistically significant estimated effect of -0.11 standard deviations is equivalent to about 14 weeks of learning.[23] The CIS model's estimated effects on math state test scores follow a similar pattern. Though they are not statistically significant, the estimated effects are consistently negative, and become more negative over time. By the third implementation year, the estimated effect size is -0.10 standard deviations. An effect of this magnitude is equivalent to about 10 weeks of learning.[24] This pattern of results remains consistent in the four general sensitivity analyses that were conducted to examine the credibility of the comparison schools.

Overall, these findings suggest that the strategies used by the comparison schools (whatever they may have been) were more successful at improving students' test scores than the CIS model. However, it is important to note that the CIS model does *not* attempt to improve state test scores or the quality of instruction provided during regular school hours.

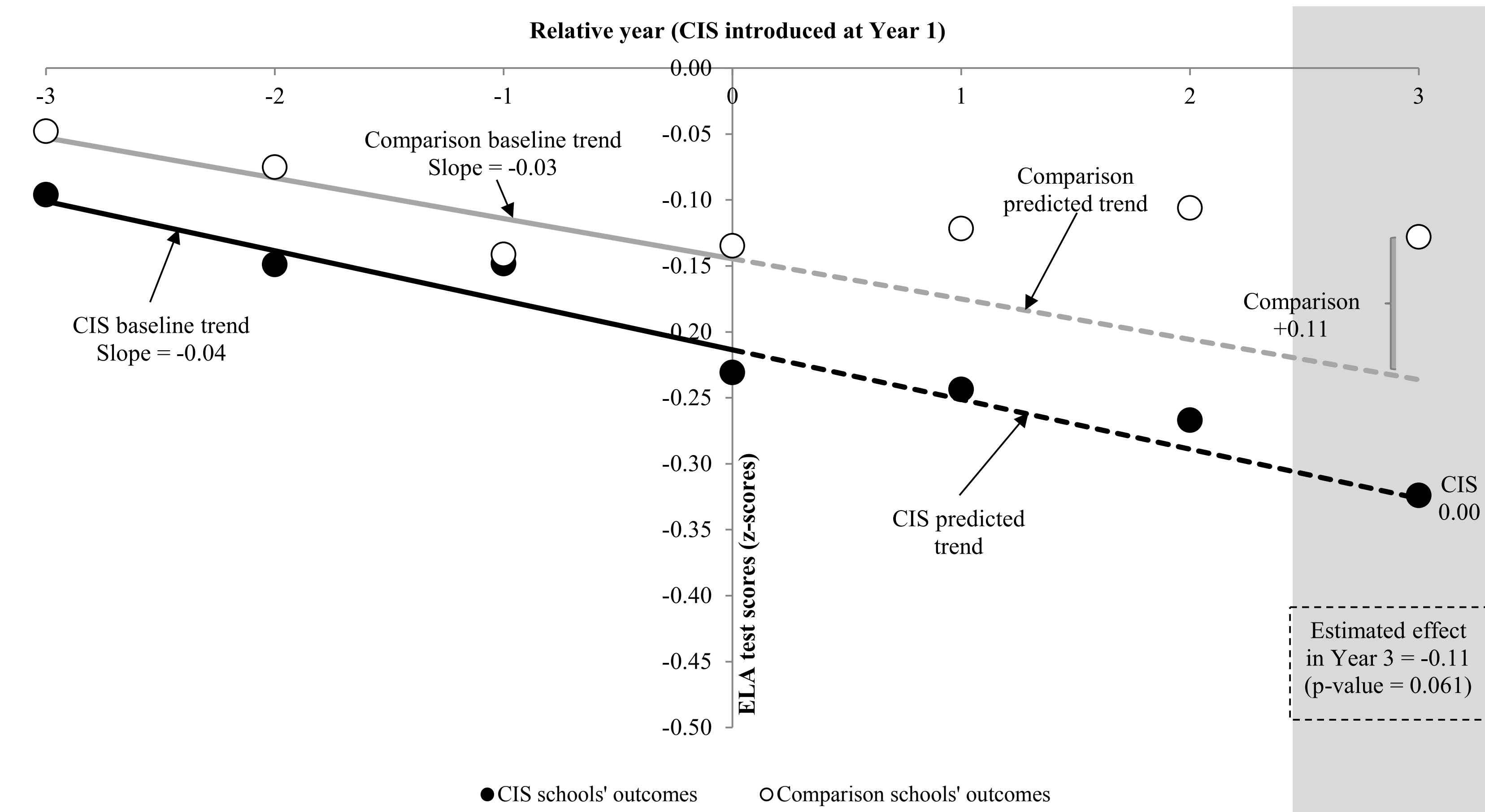
---

[23]The conversion from effect size to weeks of learning is based on data in Hill, Bloom, Black, and Lipsey (2008). Middle school students make gains of about 0.008 standard deviations per week in reading. Therefore, an effect size of 0.11 is equivalent to 14 weeks of learning (0.11 ÷ 0.008).

[24]The conversion from effect size to weeks of learning is based on data in Hill, Bloom, Black, and Lipsey (2008). Middle school students make gains of about 0.010 standard deviations per week in math. Therefore, an effect size of 0.10 is equivalent to 10 weeks of learning (0.10 ÷ 0.010).

Figure 3.5

Baseline Trends and Deviations from Trends in Middle School ELA State Test Scores



SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTE: State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

Findings for Elementary Schools

Attendance Rates

- **After three years, the average attendance rate increased by a statistically significant amount in the CIS elementary schools, relative to what would have been expected given their baseline trend.**
- **The improvement in these schools' average attendance rate was greater than the comparison schools' improvement by a statistically significant amount. The CIS model appears to have improved CIS schools' attendance rates more than they would have improved otherwise.**

Figure 3.6 plots the trends in attendance rates for all CIS elementary schools and their comparison schools, in the school years before and after the CIS model was implemented. Before CIS was launched, attendance rates in both CIS and comparison schools had been decreasing by about 0.09 percentage points per school year (though they were still relatively high, at 96 percent). In the follow-up period, attendance rates in the CIS elementary schools were higher than their baseline trend predicted, and they continued to improve over time. In the comparison schools, attendance rates were also higher than predicted in the first follow-up year, perhaps because these schools were using other strategies to improve their attendance rates.

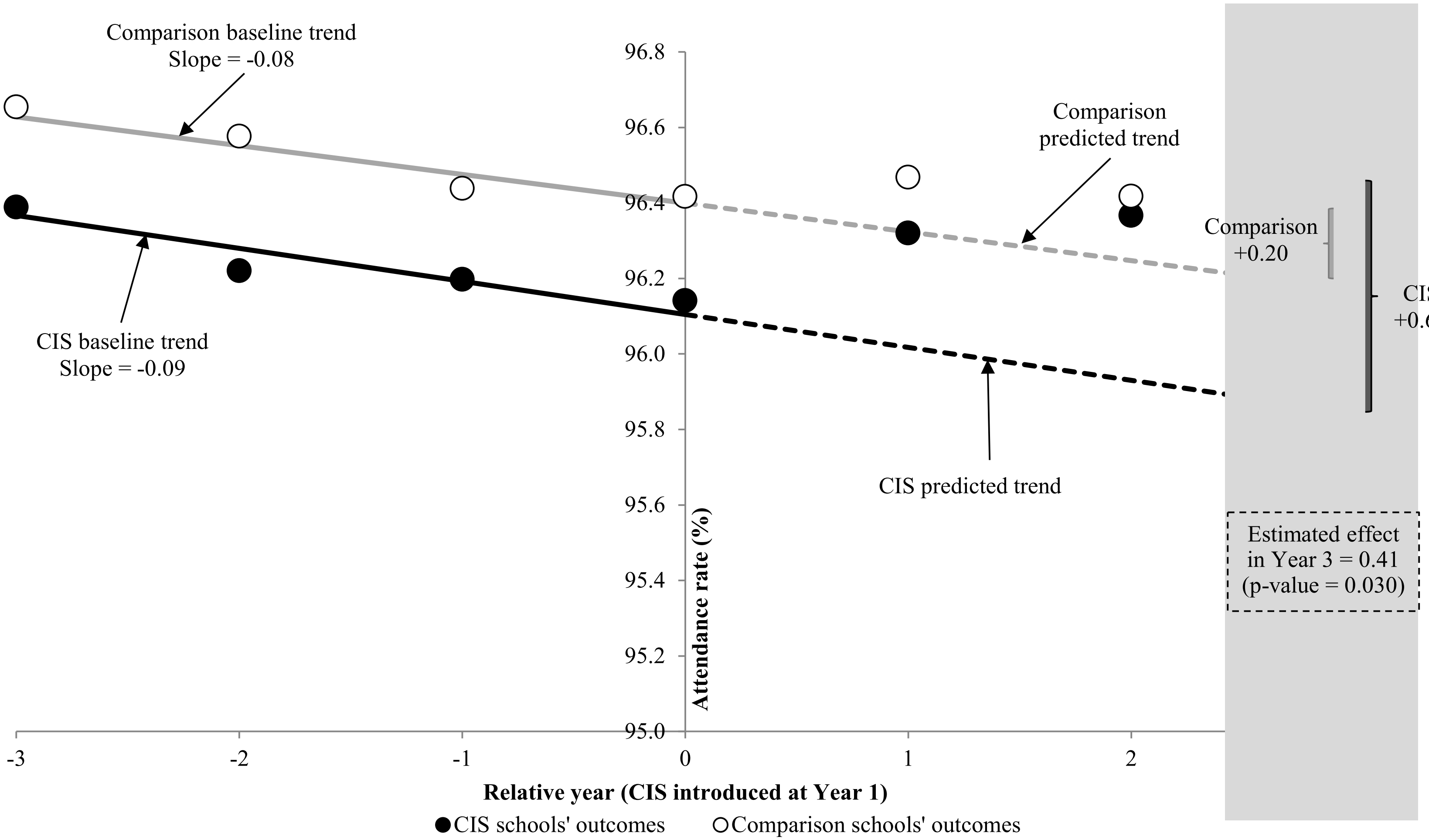
However, as shown in Table 3.5, the CIS schools' deviations from trend were consistently larger than those of the comparison schools. The estimated effects on attendance rates are positive in all three years of CIS implementation, and by the third year the estimated effect is statistically significant, at 0.4 percentage points.

This pattern of results remains consistent in the four general sensitivity analyses that were conducted to explore the credibility of the comparison schools. Because the baseline difference in attendance rates between the CIS schools and the comparison schools is large by conventional standards (greater than 0.25 standard deviations), an additional outcome-specific sensitivity test was conducted. A new set of comparison schools was chosen for the CIS elementary schools, this time matched using baseline attendance rates only (as opposed to multiple outcome measures and school characteristics).[25] After re-matching, there is no baseline difference in attendance rates between the CIS schools and the new comparison schools, nor is there any difference in the rate at which attendance rates were declining in the baseline period.

[25]The baseline difference in attendance rates for elementary schools does not appear to be caused by any particular group of schools. Therefore, the re-matching was conducted for all elementary schools.

Figure 3.6

Baseline Trends and Deviation from Trends in Elementary School Attendance Rates



SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

Table 3.5

Estimated Effects on Attendance Rates and State Test Scores, Elementary Schools

	Deviation from Baseline Trend								
Outcome and Follow-Up Year	CIS Schools		Comparison Schools		Estimated Effect		Lower 90% CI	Upper 90% CI	P-Value for Estimated Effect
Attendance rate (%)									
Year 1	0.30	†††	0.14	†	0.16		-0.07	0.39	0.257
Year 2	0.44	†††	0.17		0.27		-0.01	0.54	0.111
Year 3	0.61	†††	0.20	†	0.41	**	0.10	0.72	0.030
Number of schools	24		36						
ELA state test scores (z-scores)[a]									
Year 1	0.02		0.00		0.02		-0.06	0.09	0.686
Year 2	0.05		0.00		0.05		-0.04	0.14	0.344
Year 3	0.10	†	0.06		0.04		-0.08	0.15	0.591
Number of schools	21		32						
Math state test scores (z-scores)[a]									
Year 1	-0.03		-0.02		-0.01		-0.12	0.10	0.858
Year 2	-0.05		0.03		-0.08		-0.21	0.05	0.317
Year 3	-0.05		0.03		-0.08		-0.23	0.08	0.412
Number of schools	21		32						

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated deviations from baseline trend for each group of schools. The values in the "Estimated Effect" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. The values in the "Lower 90% CI" and "Upper 90% CI" columns are the 90 percent confidence intervals for the estimated effects. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed test was applied to estimated deviations and estimated differences between CIS schools and comparison schools. The statistical significance of estimated deviations is indicated as: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

The CIS model's estimated effects on attendance rates based on this new sample of comparison schools are consistent with the main findings.[26]

On its face, the CIS model's estimated effect on attendance rates may seem small: An effect of 0.4 percentage points means that students attended school only an extra 0.7 days on average as a result of the CIS model (assuming a school year of 180 days). However, it is important to note that attendance rates were already high before the CIS model was launched (96.1 percent), so an effect of 0.4 percentage points gets schools 10 percent closer to perfect attendance.[27] In addition, high average daily attendance rates can mask high rates of chronic absenteeism: 11 percent of elementary school students are chronically absent (meaning they miss 15 or more days during the school year).[28] Therefore, a small effect on attendance rates may represent a larger effect on chronic absenteeism.

State Test Scores

- **After three years, ELA state test scores improved by a statistically significant amount in the CIS elementary schools, relative to what would have been expected given their baseline trend.**
- **The comparison schools' ELA state test scores also increased, and by a similar amount. Therefore, it does not appear as though CIS schools' ELA state test scores improved more than they would have otherwise.**

As shown in Table 3.5, in the follow-up period CIS schools and comparison schools both performed better on ELA tests than their baseline trends predicted, and by similar amounts. The CIS model's estimated effects on state ELA tests are therefore consistently positive, but small (0.02 to 0.05 standard deviations, or one to two weeks of learning) and not statistically significant.[29] This pattern of results is consistent across the four general sensitivity analyses.

- **After three years, math state test scores did not improve in the CIS elementary schools, relative to what would have been expected given their baseline trend.**

---

[26]See Appendix E.

[27]The gap to perfect attendance is 3.9 percentage points (100 - 96.1). An effect of 0.4 percentage points is 10 percent of that gap (0.4 ÷ 3.9).

[28]Office of Civil Rights (2016).

[29]The conversion from effect size to weeks of learning is based on data in Hill, Bloom, Black, and Lipsey (2008). Elementary school students make gains of about 0.021 standard deviations per week in reading. Therefore, an effect size of 0.05 is equivalent to about two weeks of learning (0.05 ÷ 0.021).

- **The comparison schools' math state test scores did not improve either. Therefore, it does not appear as though the CIS model had an effect on schools' math test scores.**

Students in CIS schools performed slightly worse than predicted on math state tests in the follow-up period, though not by a statistically significant amount, whereas comparison schools performed slightly better than predicted. The CIS model's estimated effects are therefore consistently negative, but small in magnitude (-0.01 to -0.08 standard deviations, or one to three weeks of learning) and not statistically significant.[30] This pattern of results remains consistent in the four general sensitivity analyses that were conducted to examine the credibility of the comparison schools.

[30]The conversion from effect size to weeks of learning is based on data in Hill, Bloom, Black, and Lipsey (2008). Elementary school students make gains of about 0.023 standard deviations per week in math. Therefore, an effect size of 0.08 is equivalent to about three weeks of learning (0.08 ÷ 0.023).

Chapter 4

Discussion

The main finding of this study is that graduation and dropout rates improved in the Communities In Schools (CIS) high schools in this study after they started implementing the CIS model. Graduation and dropout rates improved in the comparison schools as well, which makes it challenging to determine whether the CIS model was *more* effective than the strategies used by the comparison schools. However, the study's findings suggest that the CIS model may be at least as effective as these other approaches.

In elementary schools, attendance rates — one of the core outcome measures the CIS model targets in those schools — appear to have improved more than they would have otherwise, which suggests that the CIS model may be more effective at improving attendance than the strategies used by the comparison schools. In middle schools, the CIS schools' state test scores did not improve after they implemented the model while the comparison schools' test scores did, which suggests that the CIS model may have been less successful at improving students' test scores than the strategies and reforms used by the comparison schools. It is worth noting that the CIS model does not specifically target test scores. Unfortunately, it was not possible to evaluate the model's effect on its main objective in middle schools: improving students' behavior.

In interpreting these results, it is important to remember that because this study's findings are based on a small number of purposefully selected schools in only two states, they may not represent the effect of the CIS model nationally, as it exists today. The schools in this study were chosen because data on their students' outcomes are available for several years before and after the CIS model was launched. These are therefore all schools that started implementing the model in the mid- to late 2000s — between 2005 and 2008 — and the last school year of the follow-up period is 2010-2011, which is several years ago.

Given the small number of study schools and their implementation timeline, the results from this study should be considered alongside the results of the other two evaluations of the CIS model: a national quasi-experimental study of the CIS model in 602 schools conducted by ICF International, and a recent randomized experiment of the CIS model in 47 K-8 schools in Chicago.[1]

[1]ICF International (2008); Figlio (2015).

Where there is consistency in the findings across studies, one can be more confident that the CIS model is effective and that it will generally have effects on a broad group of schools. It is notable, therefore, that all three studies of the CIS model conducted thus far have found positive effects on attendance rates for younger students. MDRC's study finds a statistically significant effect of 0.4 percentage points on attendance rates. The ICF International study also finds a positive though not statistically significant effect on this outcome (of 0.1 percentage points). The Chicago study finds a statistically significant reduction of 3 percentage points in truancy rates in the first year of implementation (though these effects disappear in the second year).[2]

Conversely, if different studies find different results, it may indicate that the CIS model's effect is tied to the context in which it is implemented. In this study, for example, the CIS model appears to be less effective at improving state test scores in middle schools than the strategies being used by the comparison schools. In contrast, in the second year of implementation the Chicago study finds a positive effect of 5.6 percentage points on the proportion of students who score as "proficient" in math, and of 4.3 percentage points on the proportion who score as proficient in reading.[3] This result suggests that in some contexts, the CIS model can have positive effects on students' test scores and be more effective than the other strategies available to schools.

The findings in high schools are promising but more difficult to interpret. The ICF study finds a positive effect of 1.7 percentage points on ninth-grade students' probability of graduating, but this effect is not statistically significant. In the present study, CIS schools' graduation rates improved, but it is unclear how much of this improvement can be attributed to the CIS model.

Therefore, at the high school level especially, it would be a useful next step for CIS to undertake a rigorous and large-scale study of its model that was based on a randomized experiment and supplemented by a cost study. One of the great challenges of evaluating an educational intervention is that there is no true "no-services" control group: There are many different programs and interventions available to schools, so the effectiveness of a particular intervention (like the CIS model) must be evaluated against the effectiveness of others. At the same time, some interventions are less costly and resource-intensive than others. Therefore, even if an intervention is not more effective than others, it may be more *cost*-effective. A school-level randomized experiment would provide the best evidence of the CIS model's effect on school outcomes relative to other programs and strategies. Moreover, by collecting cost data on the CIS model's implementation — as well as cost data on the strategies and interventions used by the

---

[2]In that study, a student was considered "truant" if he or she had an attendance rate of less than 95 percent.

[3]Effects are not reported separately for middle school grades in the Chicago study.

control schools — one could also determine the CIS model's relative cost-effectiveness, which would ultimately provide the most useful decision-making information for school districts. Future studies of the CIS model should also attempt to collect data on the full range of student outcomes the model aims to influence — especially behavioral outcomes such as disciplinary referrals — so that the cost-effectiveness of the model can be examined in many areas.

Next Steps for the Evaluation

As noted earlier in this report, MDRC is also evaluating the effect of one component of the CIS model — Level 2 case management services — using a student-level random assignment research design. The first report from that study described the effect of Level 2 case management on at-risk students' chronic absenteeism, credit accumulation, and disciplinary referrals after one year of services.[4] At the end of the first year, case-managed students did not have better academic or behavioral outcomes than non-case-managed students. However, one year is too early to make any definitive conclusions, because a year of case management may not be sufficient to improve the outcomes of these students, all of whom face serious academic and personal challenges. Accordingly, the next report will examine the effect of Level 2 case management after two years of services. It will also attempt to distill the results of these two studies (the comparative interrupted time series study described in this report and the Level 2 case management random assignment study) into some overall lessons about the CIS model.

[4]Corrin et al. (2015).

Appendix A

The Statistical Model and Baseline Trend Estimates

This appendix provides information on the statistical model used to estimate the Communities In Schools (CIS) model's effect, as well as estimates of the baseline trend (slope and intercept) for the CIS schools and the comparison schools from this model.

In this study, the CIS model's effect is estimated using a comparative interrupted time series (CITS) design. In practice, the effect of CIS was estimated by fitting the following model to school-level panel data for the CIS schools and their comparison schools:[1]

$$Y_{jt} = \sum_j \gamma_j SCHOOL_{jt} + \sum_j \phi_j SCHOOL_{jt} * RELYEAR_t + \alpha_1 FY1_t + \beta_1 CIS_j * FY1_t + \alpha_2 FY2_t + \beta_2 CIS_j * FY2_t + \alpha_3 FY3_t + \beta_3 CIS_j * FY3_t + \varepsilon_{jt}$$

where *j* denotes schools and *t* denotes time. In a CITS design, time is measured relative to the follow-up period: *t* is equal to 1, 2, and 3 for the first three follow-up years, and is 0 or negative for the baseline years.

The variables in the model are defined as follows:

Y_{jt}	School-level outcome for school *j* in year *t*
$SCHOOL_j$	A set of j dichotomous indicators for school (equal to 1 for school j and 0 otherwise).
$RELYEAR_t$	Continuous variable for school year centered at the last baseline year for a school, such that RELYEAR = 0 in the last baseline year[2]
CIS_j	Dichotomous indicator for whether school *j* is a CIS school (equal to 1 if school is a CIS school; equal to 0 if it is a matched comparison school)
$FY1_t$	Dichotomous indicator for whether an outcome is for the first year of CIS implementation (equal to 1 if yes, 0 otherwise)
$FY2_t$	Dichotomous indicator for whether an outcome is for the second year of CIS implementation (equal to 1 if yes, 0 otherwise)

[1]The baseline and follow-up school years for a comparison school are the same as the CIS school to which it is matched (that is, the definitions of *RELYEAR*, *FY1*, *FY2*, and *FY3* are the same for a CIS school and its matched comparison schools).

[2]Some schools have five years of baseline data, while others have four years of baseline data. The statistical model can accommodate this variability.

$FY3_t$	Dichotomous indicator for whether an outcome is for the third year of CIS implementation (equal to 1 if yes, 0 otherwise)
ε_{jt}	Random variation in the outcome of interest over time within schools

The model provides estimates of the following quantities of interest related to the CITS study design:

γ_j	Predicted mean (intercept) for each CIS and comparison school in the study in the last baseline year
ϕ_j	Baseline slope for each school in the study
α_1	Deviation from baseline trend for comparison schools in the first follow-up year
$\alpha_1 + \beta_1$	Deviation from baseline trend for CIS schools in the first follow-up year
α_2	Deviation from baseline trend for comparison schools in the second follow-up year
$\alpha_2 + \beta_2$	Deviation from baseline trend for CIS schools in the second follow-up year
α_3	Deviation from baseline trend for comparison schools in the third follow-up year
$\alpha_3 + \beta_3$	Deviation from baseline trend for CIS schools in the third follow-up year

In this model, β_1 represents the estimated effect of the CIS model in the first follow-up year — that is, the deviation from baseline trend for CIS schools minus the deviation from trend for comparison schools. Similarly, β_2 represents the estimated effect of CIS in the second follow-up year, and β_3 represents the estimated effect in the third follow-up year. Based on recommendations from the field related to using school-level data, the analysis does not weight each observation by the number of students in each school.[3]

As explained in Chapter 2 and Appendix C, some comparison schools are the "match" for more than one CIS school, and are therefore in the data set more than once. To account for

[3]Jacob, Goddard, and Kim (2014).

these duplicate observations, cluster-robust standard errors are used for hypothesis testing, where clusters are defined as repeated instances of a school's average outcomes for a particular school year.

Also note that this model allows each school to have its own baseline intercept and slope. This strategy is used to account for the longitudinal nature of the data, whereby school years are clustered within schools. In order to estimate the *average* intercept and slope of all schools in the CIS group and the comparison group — for the purposes of plotting the average baseline trends for CIS and comparison schools — the school-level slopes (ϕ_j) and intercepts (γ_j) from the regression model are weighted by the precision of the estimated slope for each school.[4]

Appendix Table A.1 presents the estimated baseline intercepts for the CIS schools and the comparison schools. These intercepts represent the *predicted* mean outcomes for CIS and comparison schools in the last baseline year (in Figure 2.1, for example, the predicted value of the outcome in Year 0). The pattern of differences in the predicted mean outcomes (intercepts) of the CIS and comparison schools closely mirrors the differences in their observed outcomes in the last baseline year (Tables 2.4 to 2.6).[5]

Appendix Table A.2 presents the slopes of the baseline trends of the CIS schools and the comparison schools: the average annual change in each outcome measure during the baseline period. Positive slopes indicate that the outcome measure's value was increasing during the baseline period, while negative slopes indicate that the outcome measure's value was decreasing. For all outcomes, the sign of the slopes is the same for the CIS and the comparison schools (both are positive or both are negative).

[4]The precision of a school-level slope is the squared inverse of the standard error of the slope.

[5]More of the differences in Appendix Table A.1 are statistically significant than is the case in Tables 2.4 to 2.6, because predicted means are more precisely estimated than observed means. ELA stands for English/language arts.

Appendix Table A.1

Predicted Mean Outcome Measures of CIS Schools and Comparison Schools in the Last Baseline Year (Baseline Intercepts)

Outcome Measure	CIS Schools	Comparison Schools	Estimated Difference	P-Value
High schools				
Graduation rate (%)	69.2	78.2	-9.0 ***	<0.001
Dropout rate (%)	4.3	3.0	1.3 ***	<0.001
Attendance rate (%)	92.6	93.2	-0.6 ***	<0.001
State test scores (z-scores)[a]				
ELA	-0.19	-0.10	-0.09 ***	<0.001
Math	-0.18	-0.06	-0.11 ***	<0.001
Number of schools	14	18		
Middle schools				
Attendance rate (%)	95.5	95.7	-0.2 ***	<0.001
State test scores (z-scores)[a]				
ELA	-0.21	-0.14	-0.07 ***	<0.001
Math	0.01	-0.10	0.12 ***	<0.001
Number of schools	15	24		
Elementary schools				
Attendance rate (%)	96.1	96.4	-0.3 ***	<0.001
State test scores (z-scores)[a]				
ELA	-0.33	-0.32	-0.01	0.221
Math	-0.23	-0.23	0.01	0.627
Number of schools	24	36		

SOURCES: MDRC calculations based on school-level data on student outcome measures from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the predicted mean outcome measures for CIS schools and comparison schools, in the school year immediately before CIS was implemented (the intercept of the baseline trend). The values in the "Estimated Difference" column are the differences between CIS schools and comparison schools with respect to their predicted baseline means. Rounding may cause slight discrepancies in calculating differences.

A two-tailed test was applied to estimated differences between CIS schools and comparison schools. The statistical significance of estimated differences is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

Appendix Table A.2

Slopes of Baseline Trends in Outcome Measures for CIS Schools and Comparison Schools

Outcome Measure	CIS Schools' Slope	Comparison Schools' Slope	Estimated Difference		P-Value
High schools					
Graduation rate (%)	-3.4	-2.0	-1.3	***	<0.001
Dropout rate (%)	0.7	0.6	0.1		0.201
Attendance rate (%)	-0.3	-0.3	-0.1	*	0.083
State test scores (z-scores)[a]					
ELA	-0.02	-0.01	-0.01		0.173
Math	-0.02	-0.03	0.00		0.560
Number of schools	14	18			
Middle schools					
Attendance rate (%)	-0.1	-0.1	-0.1	***	0.002
State test scores (z-scores)[a]					
ELA	-0.04	-0.03	-0.01	*	0.093
Math	-0.04	-0.02	-0.01	**	0.028
Number of schools	15	24			
Elementary schools					
Attendance rate (%)	-0.1	-0.1	0.0		0.308
State test scores (z-scores)[a]					
ELA	-0.02	-0.03	0.01		0.199
Math	-0.01	-0.01	-0.01		0.360
Number of schools	24	36			

(continued)

Appendix Table A.2 (continued)

SOURCES: MDRC calculations based on school-level data on student outcome measures from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated slopes in the baseline trends for CIS schools and comparison schools, based on four or five years of preintervention data on the outcome measure of interest. The values in the "Estimated Difference" column are the differences between CIS schools and comparison schools with respect to the slopes of their baseline trends. Rounding may cause slight discrepancies in calculating differences.

A two-tailed test was applied to estimated differences between CIS schools and comparison schools. The statistical significance of estimated differences is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

Appendix B

The Selection of Schools Implementing the Communities In Schools Model

MDRC and Communities In Schools (CIS) decided that the study should include CIS schools in two states: Texas and North Carolina. These states were chosen because (1) CIS has a strong presence and history in both states and (2) historical data on school-level outcomes are publicly available in both states, making it possible to use a comparative interrupted time series (CITS) design to evaluate the effect of the CIS model. In Texas, it was further decided that the study should focus on three CIS affiliates: Greater Central Texas (an area north of Austin), Central Texas (the area around Austin), and Houston.[1] These affiliates were chosen because they serve a large number of schools.

For the purposes of this study, CIS provided MDRC with a list of all CIS schools in North Carolina and the three Texas affiliates. CIS also provided information on the school year that CIS was launched in these schools and the types of CIS services provided by the schools (Level 1, Level 2, or both). To be included in the study, a CIS school had to meet the following criteria related to data availability and CIS implementation.

- **Number of years of baseline and follow-up data:** To be included in the study, a CIS school had to have at least four years of data on the outcome measures of interest before it began implementing the CIS model, and three years of follow-up data.[2] In both Texas and North Carolina, school-level data were available from the school years 2001-2002 through 2010-2011 at the time of the analysis. In order to meet the data requirements for the study, therefore, a CIS school had to have started implementing the CIS model from 2005-2006 to 2008-2009 (see Appendix Table B.1).[3]

- **Strong implementation of the CIS model:** CIS schools sometimes roll out the CIS model in stages. For example, a CIS school may start with Level 1 services and then in a later school year add Level 2 services, or vice versa. To ensure that this study is a fair test of the CIS model, it was decided that a CIS school must have implemented both levels of CIS services (Level 1 and

[1]CIS affiliates are administrative regions where CIS operates, often made up of multiple school districts.

[2]At least four years of data are needed to reliably estimate baseline trends in a CITS design. See Somers, Zhu, Jacob, and Bloom (2013). Three years of follow-up data are required because whole-school models are difficult to implement and their effects tend to grow over time. See Borman, Hewes, Overman, and Brown (2003).

[3]In North Carolina, CIS middle schools and elementary schools had to have started implementing the model in 2007-2008 or 2008-2009 to be included in the study sample, because outcome data on middle and elementary schools are only available from 2003-2004 onwards.

Appendix Table B.1

Baseline and Follow-Up Periods for This Study, by Start Year

	School Year									
Start Year	**2001-2002**	**2002-2003**	**2003-2004**	**2004-2005**	**2005-2006**	**2006-2007**	**2007-2008**	**2008-2009**	**2009-2010**	**2010-2011**
2005-2006	PRE	PRE	PRE	PRE	POST	POST	POST			
2006-2007	PRE	PRE	PRE	PRE	PRE	POST	POST	POST		
2007-2008		PRE	PRE	PRE	PRE	PRE	POST	POST	POST	
2008-2009			PRE	PRE	PRE	PRE	PRE	POST	POST	POST

NOTE: PRE = Baseline school years (not implementing CIS); POST = first three follow-up years (implementing CIS).

Level 2) during at least part of the three-year follow-up period. However, information on Level 1 and 2 offerings was not available for all CIS schools.[4] For schools where that information was lacking, the study team instead used information on whether or not the school had achieved "comprehensive" status, a designation given by CIS National to schools that satisfy certain implementation criteria in its internal monitoring system.[5] Specifically, to be included in the study, a school had to have achieved "comprehensive" status within six years of launching the CIS model (which makes it highly likely that the school offered both Level 1 and 2 services during the first three years of implementation).[6]

[4]Data on Level 1 and 2 service offerings were available for North Carolina and Houston schools, but not for those in Greater Central Texas and Central Texas.

[5]The CIS network has also implemented a Total Quality System that includes standards for effective business practices and a defined program model of integrated student services for schools. One component of the Total Quality System classifies schools as "developing" or "comprehensive." To be designated "comprehensive," a school has to hit specific implementation targets. Some of these targets are related to planning and record keeping. For instance, the school must have a site operations plan; must develop case plans for students; and must track the services provided to students. Other targets are loosely related to program operations. For instance, a school must have at least a half-time site coordinator; must provide at least eight types of Level 1 services during the year that reach at least 75 percent of students; and must provide Level 2 services to at least 10 percent of students if the school has less than 1,000 students, or at least 5 percent if the school has more than 1,000 students.

[6]For some schools in North Carolina and Texas, data are available on their Level 1 and Level 2 services, as well as their "comprehensive" statuses. In these schools, it was possible to see how much time passed from the school year they first offered both Level 1 and 2 services to the year they became "comprehensive." On average, a school achieved "comprehensive" status about three years after it started offering both Level 1 and

(continued)

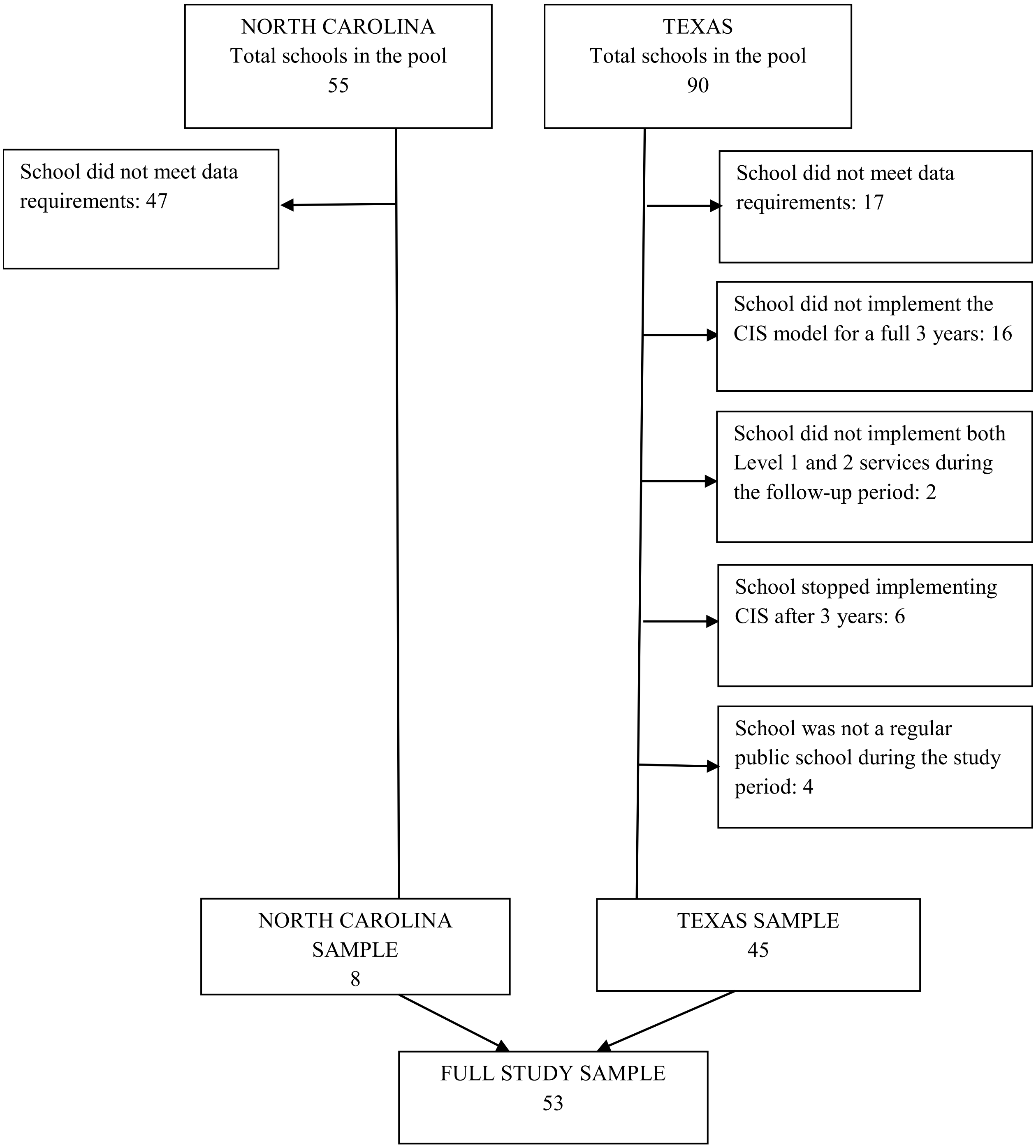
In addition to these core criteria, a CIS school also had to be a regular public school during the entire study period (not a charter school, a magnet school, or another kind of alternative school), and it had to have implemented the CIS model for more than three years. The latter criterion was used to ensure that the study schools were committed to the CIS model.

A total of 53 CIS schools met these criteria and were included in the study: 14 high schools, 15 middle schools, and 24 elementary schools. Appendix Figure B.1 depicts how the final study sample of CIS schools was created.

Level 2 services. If a school became comprehensive within six years of launching the model, then it was very probably offering Level 1 and Level 2 services during the first three follow-up years.

Appendix Figure B.1

Creation of the Study Sample



SOURCE: MDRC calculations based on program data provided by Communities In Schools, as well as school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

Appendix C

The Selection of Comparison Schools

This appendix describes the process that was used to choose the comparison schools in the study. The first section provides describes the pool of schools from which the comparison schools were chosen, and the second section describes how comparison schools were chosen from this pool.

The Comparison Pool

The pool of potential comparison schools consisted of schools not implementing the Communities In Schools (CIS) model ("non-CIS schools") in several counties where CIS is operating:

- In North Carolina, the pool of potential comparison schools includes non-CIS schools in the four counties where the CIS study schools are located, as well as non-CIS schools in five neighboring counties where CIS had also been working (nine counties total).
- In Texas, the pool of potential comparison schools includes non-CIS schools in the six counties where the CIS study schools are located, as well as non-CIS schools in 15 other counties served by CIS in Texas (21 counties total).

To be included in the pool of potential comparison schools, a non-CIS school also had to have data on all of the study's outcome measures from 2001-2002 to 2010-2011, to ensure that it could be matched to any of the CIS schools.

The Matching Process

The comparison schools in the study were chosen by matching each of the 53 CIS schools in the study to the two non-CIS schools in the comparison pool with the most similar outcome measure values in the year before CIS was launched, the most similar baseline trends in the outcome measures of interest, and the most similar demographic and structural characteristics before CIS was launched. (Note that for high schools, *dropout rates* were used as a matching variable rather than graduation rates because dropout rates are available for all schools. See Chapter 2 for more information on data availability.)

In this study, the "similarity" between a CIS school and its potential comparison schools is based on the Euclidian distance, a metric that captures the differences between two schools in several matching variables (in this case, schools' baseline outcome measure values and characteristics). The Euclidian distance across M matching variables is defined as: $D = \sqrt{\sum_M (T_m - C_m)^2}$ where T_m is the value of characteristic *m* for the CIS school and C_m is the value of characteristic *m* for the non-CIS school. Appendix Table C.1 summarizes the matching

Appendix Table C.1

Matching Characteristics Used to Identify Comparison Schools, by State and School Type

Matching Characteristic	Texas: Elementary and Middle Schools	Texas: High Schools	North Carolina: Elementary and Middle Schools	North Carolina: High Schools
Structure/demographics in the last baseline year				
Total enrollment	✓	✓	✗	✓
Free or reduced-price lunch (%)	✓	✓	✓	✓
Black (%)	✓	✓	✓	✓
White (%)	✓	✓	✓	✓
Title I status[a]	✓	✓	✓	✓
School location (urban/rural)[a]	✓	✓	✓	✓
Number of crimes per 1,000 students	No data	No data	✗	✓
Number of short-term suspensions	No data	No data	✗	✓
Outcomes in the last baseline year				
Dropout rate (%)	NA	✓	NA	✓
Attendance rate (%)	✓	✓	✓	✓
ELA z-score	✓	✓	✓	✓
Math z-score	✓	✓	✓	✓
Proficient on state ELA test (%)	✓	✓	✓	✓
Proficient on state math test (%)	✓	✓	✓	✓
Slope of the baseline trend				
Dropout rate baseline trend	NA	✓	NA	✓
Attendance rate baseline trend	✓	✓	✓	✓
ELA test scores baseline trend	✓	✓	✓	✓

(continued)

Appendix Table C.1 (continued)

SOURCES: The Common Core of Data and school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: NA = outcome not applicable for this age group.
No data = data not available on this outcome in Texas.
✓ = variable used in matching.
✗ = including this outcome decreased the quality of the match so it was not used.
ELA = English/language arts.
[a]This variable was not included in the Euclidian distance, because schools were matched exactly with respect to it (it was used as a stratum for the matching).

variables that were used to calculate the Euclidian distance in this study. Note that the matching variables differ between the two states (due to differences in data availability) and among elementary, middle, and high schools (because some outcome measures are only applicable to high schools).

Based on the Euclidian distance, each CIS school was then matched to the two most similar (least distant) comparison schools. This process was conducted separately for the two states; for elementary, middle, and high schools; for urban and rural schools; and for schools that did and did not receive Title I funds (the federal funding stream for schools serving low-income students). For example, based on these criteria, a rural CIS high school in Texas receiving Title I funds would be matched to the two most similar (least distant) non-CIS rural high schools in Texas that were also receiving Title I funds.[1] Comparison schools were chosen with replacement — meaning that a non-CIS school could be chosen as the comparison school for more than one CIS school. Though matching with replacement reduces the number of unique comparison schools in the study, it maximizes the similarity of the comparison schools and the CIS schools with respect to the matching variables.[2]

[1]In other words, comparison schools are matched exactly with respect to state, school type, location, and Title I status.

[2]The baseline and follow-up period used to calculate a comparison school's trends and deviations from trends are the same periods as those used for the CIS school it is matched with. If a comparison school is chosen more than once and the CIS schools to which it is matched launched CIS in different school years, then different school-year ranges of its data are used for the comparisons with those two CIS schools. For this reason, comparison schools that are chosen more than once must be included more than once in the analysis. Standard errors and hypothesis tests are adjusted to account for the fact that some comparison schools are included in the analysis multiple times.

To verify that the matching process is sound (and that it can yield a credible comparison group), the study team estimated the "impact" of the CIS model in the last baseline year. The effect of the CIS model in the baseline period is known (it is zero by definition), so this "effect" provides a useful benchmark. If the CIS model's estimated effect in the last baseline year is *not* zero, then the matching process is not producing a credible comparison group (which could mean that the problem was in the definition of the comparison pool, the matching characteristics, or the Euclidian distance). The impact of CIS in the last baseline year was estimated as follows.

1. **Redefine the baseline/follow-up period:** The follow-up period for the CIS schools was redefined so that the school year immediately before CIS implementation became the first follow-up year in the CITS design.

2. **Select comparison schools:** Next, the Euclidian distance was recalculated using the values of the matching variables in the redefined baseline period (which now excluded the school year before CIS was launched). Two comparison schools were then chosen for each CIS school based on this metric.

3. **Estimate effects:** The CIS model's effect was estimated in the new "first follow-up year" (really the last baseline year) using a CITS design.[3] If the comparison-school-selection process is sound, then the CIS model's "effect" in the last baseline year should be close to zero in magnitude and not statistically significant.

This validation exercise made it possible to "optimize" the set of characteristics used for the matching. The research team conducted the analysis several times, using different combinations of matching characteristics, until the CIS model's estimated effects in the last baseline year were as close to zero as possible.

The matching variables in Appendix Table C.1 are those that produced the smallest estimated effects on the outcome measures of interest in the last baseline year.[4] As shown in Appendix Table C.1, some variables were excluded as matching variables (those marked with an "x"), because including them in the matching process increased the size of the CIS model's estimated effect in the last baseline year. Estimated effects in the last baseline year (based on the variables in Appendix Table C.1) can be found in Appendix E. As discussed in that appendix,

---

[3]Because the study period is shifted back, schools in the study sample now have fewer years of baseline data (three to four years rather than four to five years). Although trends are less reliably estimated when there are only three years of baseline data, this limitation was deemed acceptable for the purposes of this exercise.

[4]This exercise was set up so that the team could not look at the CIS model's effect in the new "second and third follow-up years," when CIS was actually being implemented. Seeing these results could have inadvertently biased the team toward choosing matching variables that led to bigger effects.

the estimated effects in the last baseline year generated by this exercise are small, which lends general credibility to the matching process.

In addition to facilitating the choice of matching variables, this validation exercise also made it possible to examine whether matching *within district* might provide a more credible group of comparison schools. To investigate this possibility, the selection of comparison schools was constrained so that each CIS school was matched to the two most similar comparison schools in the same school district. Imposing this constraint increased the CIS model's estimated effects in the last baseline year, which suggests that matching within district does not improve the credibility of the comparison group. This result may reflect the fact that there are few non-CIS schools in each district, and thus the CIS schools in a given district may actually have very different characteristics and values on outcome measures from the non-CIS schools in the same district. For these reasons, the research team decided not to match within districts explicitly.

Once the final set of optimal matching variables was determined (those in Appendix Table C.1) — and it was decided that matching within district was not necessary — the matching process was conducted one last time based on the values of the matching variables in the actual baseline period of the CIS schools. This final step resulted in 78 unique comparison schools (18 high schools, 24 middle schools, and 36 elementary schools) matched to the 53 CIS schools in the study. Before the impact analysis was conducted, the list of comparison schools was sent to CIS National to confirm that none of the comparison schools had implemented the CIS model during the study period.

Appendix D

Standard Deviations

This appendix reports on the standard deviations for the outcome measures in this study, so that the baseline differences and effects in this study can be converted to effect sizes for future reviews and meta-analyses.

For effect sizes to be comparable across studies, they should be based on the student-level standard deviation in the outcome measure or characteristic of interest, which is larger than the school-level standard deviation for that same outcome measure or characteristic.[1] Typically, when an analysis is based on school-level data (as is the case in the present study), it is not possible to calculate student-level standard deviations. However, in this study, most student outcomes and characteristics are binary (dichotomous), and therefore student-level standard deviations can be derived.

This first section of this appendix explains the calculation of standard deviations for each data type (binary student outcomes and characteristics, school characteristics, and state test scores). The second section presents the standard deviations.

Calculation of Standard Deviations

Binary Outcomes and Characteristics

The student-level standard deviation of a binary variable can be obtained as follows:

Student-level standard deviation = $\sqrt{p * (1 - p)}$

where p is the percentage of students who have a particular characteristic or outcome. (In other words, it is the student-level mean of the binary outcome or characteristic.)

Therefore, for the binary student outcomes and characteristics in this study (graduation, dropping out, proficiency on state tests, free or reduced-price lunch status, etc.), the student-level standard deviation was obtained by calculating the student-level mean for each binary student outcome or characteristic (which is equal to the school-level mean weighted by school enrollment), and then using the formula above to obtain the student-level standard deviation.[2]

Note that the attendance rate is not a binary outcome (even though it is expressed as a percentage), so it is not possible to calculate the student-level standard deviation for this outcome. Thus, in this appendix, the school-level standard deviation is shown. The school-level

---

[1]What Works Clearinghouse (2014a).

[2]The mean outcomes in Tables 2.4 to 2.6 are not used to calculate the standard deviations, because the means in these tables are school-level means (that is, they are not weighted by the number of students in each school).

standard deviation is smaller than the student-level standard deviation, so the effect size for baseline differences in attendance rates in Tables 2.4 to 2.6 are larger than they should be (that is, the effect sizes would be smaller if the student-level standard deviation were available to calculate them).

School Characteristics

For school characteristics that are the features of schools and not students (enrollment, pupil/teacher ratio, Title I status, location), standard deviations are school-level standard deviations by definition.

State Test Scores

In this study, the student-level standard deviation in test scores is not directly available, so it was estimated by dividing the school-level standard deviation by the square root of the intraclass correlation (ICC):[3]

$$\text{Student-level standard deviation} = \frac{School-level\ standard\ deviation}{\sqrt{ICC}}$$

The school-level standard deviation in test scores was calculated for each state by grade and by school year (based on all schools in Texas and North Carolina). These school-level standard deviations were then converted to student-level standard deviations by making an assumption about the ICC in each type of school based on work from other studies.[4]

The resulting student-level standard deviations are estimates of the statewide student-level standard deviations in test scores by state, grade, and school year. These standard deviations were then used to standardize (z-score) the state test scores by state and by grade and by school year. These standard deviations are not presented in this appendix, because estimated differences and effects in the report are already expressed as effect sizes.

Standard Deviations

Appendix Table D.1 presents the standard deviations for student outcomes in the year before Communities In Schools (CIS) was launched (the last baseline year). The last baseline year

---

[3]What Works Clearinghouse (2014a). The ICC is the proportion of the total variation in an outcome measure that occurs between schools (as opposed to within schools).

[4]For high schools, it was assumed that the ICC is 0.14 for English/language arts (ELA) test scores and 0.19 for math; for middle schools, it was assumed that the ICC is 0.21 for ELA test scores and 0.22 for math; for elementary schools, it was assumed that the ICC is 0.18 for ELA test scores and 0.19 for math. These assumptions are from Bloom, Richburg-Hayes, and Black (2007).

provides a useful reference point for the amount of variation in these outcome measures before CIS was launched. Standard deviations are shown for the CIS schools and comparison schools separately, and for both groups pooled together. As noted earlier, the standard deviation for the attendance rate is the school-level standard deviation, but for all other outcome measures the student-level standard deviation is shown.

Appendix Tables D.2 to D.4 present the standard deviations for student and school characteristics drawn from the Common Core of Data (for example, racial/ethnic composition and pupil/teacher ratio) in the last baseline year, for elementary schools, middle schools, and high schools. For student characteristics, the student-level standard deviation is shown; for school characteristics, the school-level standard deviation is shown.

Appendix Table D.1

Standard Deviations of Student Outcome Measures in the Last Baseline Year

	CIS Schools		Comparison Schools		All Schools	
Outcome Measure	N	SD	N	SD	N	SD
High schools						
Graduation rate (%)	8	46.5	11	42.3	19	44.2
Dropout rate (%)	14	21.9	18	18.4	32	19.7
Attendance rate (%)	14	1.6	18	1.4	32	1.4
Proficient on state tests (%)						
ELA	9	40.9	12	37.3	21	38.7
Math	9	50.0	12	49.5	21	49.8
Middle schools						
Attendance rate (%)	15	1.0	24	0.9	39	0.9
Proficient on state tests (%)						
ELA	8	40.9	15	41.0	23	41.0
Math	8	47.5	15	47.3	23	47.3
Elementary schools						
Attendance rate (%)	24	1.0	36	0.8	60	0.9
Proficient on state tests (%)						
ELA	21	41.1	32	39.2	53	39.9
Math	21	43.4	32	42.6	53	42.8

SOURCES: MDRC calculations based on school-level data on student outcome measures from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: N = number of schools; SD = standard deviation; ELA = English/language arts.
Standard deviations for attendance rates are school-level standard deviations. Standard deviations for the other outcome measures are student-level standard deviations.

Appendix Table D.2

Standard Deviations of School Characteristics in the Last Baseline Year, High Schools

	CIS Schools		Comparison Schools		All Schools	
Baseline Characteristic	N	SD	N	SD	N	SD
Free or reduced-price lunch (%)	14	50.0	16	49.1	30	49.7
Racial/ethnic composition (%)						
Black	14	47.4	18	45.0	32	45.9
Hispanic	14	49.1	18	47.8	32	48.3
White	14	38.5	18	46.6	32	44.5
Other	14	25.9	18	20.9	32	22.8
Enrollment	14	1,165.7	18	1,137.4	32	1,134.4
Pupil/teacher ratio	14	2.7	18	1.7	32	2.0
Title I status (%)	14	49.7	18	48.8	32	48.5
Location (%)						
City/town	14	42.6	18	41.8	32	41.5
Rural	14	42.6	18	41.8	32	41.5

SOURCE: MDRC calculations based on data from the Common Core of Data.

NOTES: N = number of schools; SD = standard deviation.
Standard deviations for the student characteristics (free or reduced-price lunch status and racial/ethnic group) are student-level standard deviations. Standard deviations for school characteristics (enrollment, pupil/teacher ratio, Title I status, and location) are school-level standard deviations.

Appendix Table D.3

Standard Deviations of School Characteristics in the Last Baseline Year, Middle Schools

	CIS Schools		Comparison Schools		All Schools	
Baseline Characteristic	N	SD	N	SD	N	SD
Free or reduced-price lunch (%)	15	49.5	23	49.9	38	50.0
Racial/ethnic composition (%)						
Black	15	43.2	24	45.9	39	45.2
Hispanic	15	49.5	24	46.7	39	47.8
White	15	44.8	24	46.9	39	46.4
Other	15	20.1	24	21.4	39	21.1
Enrollment	15	215.8	24	307.5	39	289.0
Pupil/teacher ratio	15	1.3	24	2.2	39	2.1
Title I status (%)	15	51.6	24	50.7	39	50.5
Location (%)						
City/town	15	25.8	24	25.4	39	25.2
Rural	15	25.8	24	25.4	39	25.2

SOURCE: MDRC calculations based on data from the Common Core of Data.

NOTES: N = number of schools; SD = standard deviation.

Standard deviations for the student characteristics (free or reduced-price lunch status and racial/ethnic group) are student-level standard deviations. Standard deviations for school characteristics (enrollment, pupil/teacher ratio, Title I status, and location) are school-level standard deviations.

Appendix Table D.4

Standard Deviations of School Characteristics in the Last Baseline Year, Elementary Schools

	CIS Schools		Comparison Schools		All Schools	
Baseline Characteristic	N	SD	N	SD	N	SD
Free or reduced-price lunch (%)	24	39.9	35	43.9	59	42.7
Racial/ethnic composition (%)						
Black	24	43.1	36	43.2	60	43.2
Hispanic	24	49.3	36	49.7	60	49.6
White	24	32.9	36	36.0	60	35.1
Other	24	20.6	36	20.5	60	20.5
Enrollment	24	186.4	36	160.8	60	168.4
Pupil/teacher ratio	24	6.1	36	5.7	60	5.8
Title I status (%)	24	0.0	36	0.0	60	0.0
Location (%)						
City/town	24	0.0	36	0.0	60	0.0
Rural	24	0.0	36	0.0	60	0.0

SOURCE: MDRC calculations based on data from the Common Core of Data.

NOTES: N = number of schools; SD = standard deviation.

Standard deviations for the student characteristics (free or reduced-price lunch status and racial/ethnic group) are student-level standard deviations. Standard deviations for school characteristics (enrollment, pupil/teacher ratio, Title I status, and location) are school-level standard deviations.

Appendix E

Sensitivity Analyses

This appendix discusses the sensitivity analyses that were conducted to examine whether the comparison schools provide a credible representation of the Communities In Schools (CIS) schools had they not implemented the CIS model. These sensitivity analyses explore various factors that could compromise the credibility of the comparison group.

The discussion in this appendix focuses on whether the comparison schools provide the right counterfactual for estimating effects on graduation rates, dropout rates, elementary school attendance rates, and middle and high school English/language arts (ELA) state test scores, which are the outcomes where there are statistically significant effects in the main analysis.

As will be discussed in this appendix, none of the sensitivity analyses that were conducted contradict the finding that CIS appears to have had a positive effect on elementary school attendance rates in the study schools, and a negative effect on middle school ELA state test scores. Though the *size* of the estimated effect on these outcomes varies from analysis to analysis, the general pattern of estimated effects is consistent and robust, which suggests that the observed effect of the CIS model is not confounded with the effect of other factors.

In contrast, some of the sensitivity analyses suggest that the comparison schools may not provide the right counterfactual for evaluating the CIS model's effect on high school graduation rates, dropout rates, and state test scores.

Hypothesis 1: The Effect of CIS Is Confounded With the Effect of a District Policy Change

The first hypothesis explored in this appendix is that the effect of CIS is confounded with a district-wide educational initiative or policy shock that affected the CIS schools but not the comparison schools. As explained in Chapter 2, the CIS and comparison schools are located in different school districts, because it was not possible to find similar matches for the CIS schools in the same districts.[1] Thus, in theory, the districts where the CIS schools are located may have experienced policy changes or instructional reforms that are confounded with the effects observed in this study.

For example, as discussed in the report, the estimated effect of the CIS model on graduation rates is 7.5 percentage points in the third year of CIS implementation. One possible explanation for this pattern of results — other than the effect of CIS — is that the CIS schools

[1]CIS high schools are located in 9 districts and their matched comparison high schools are located in 17 others, with no overlap. CIS middle schools are located in 5 districts and their comparison schools are located in 14 others, with no overlap. CIS elementary schools are located in 9 districts and their comparison schools are located in 19, with 4 districts in common.

are located in districts where graduation rates responded more to the Great Recession, perhaps because of features of the local economy, than was the case in the comparison schools' districts. Another possibility is that the districts where the CIS schools are located implemented some other reform or initiative at the same time as the CIS model was launched, while the comparison schools did not. If that were the case, it could also explain why attendance rates improved by a greater amount in CIS elementary schools than in comparison elementary schools. The reverse could explain the CIS model's estimated negative effect on state test scores in middle schools: These results could reflect the fact that comparison districts implemented district-wide instructional reforms while the CIS districts did not.

Thus, the threat to the validity of the results is that the CIS districts and the comparison districts may have experienced different changes in their trends, whether because they implemented different district-wide initiatives or because they were affected differently by changes in the economy. If so, then the comparison schools' deviation from their baseline trend may not provide the right counterfactual for what would have happened to the CIS schools had they not implemented the CIS model.

To test this hypothesis, the research team examined (1) the district-level trends in school outcome measures in the districts where the CIS and comparison schools are located, and (2) deviations from these baseline trends at the district level. The idea is to determine whether the two groups of districts have similar deviations from their baseline trends. If they do, it would suggest that the CIS schools and comparison schools in the study are located in districts that experienced similar trends and changes during the study period. This conclusion, in turn, would indicate that the estimated effects in this study are not confounded with the effects of district-wide reforms or initiatives.

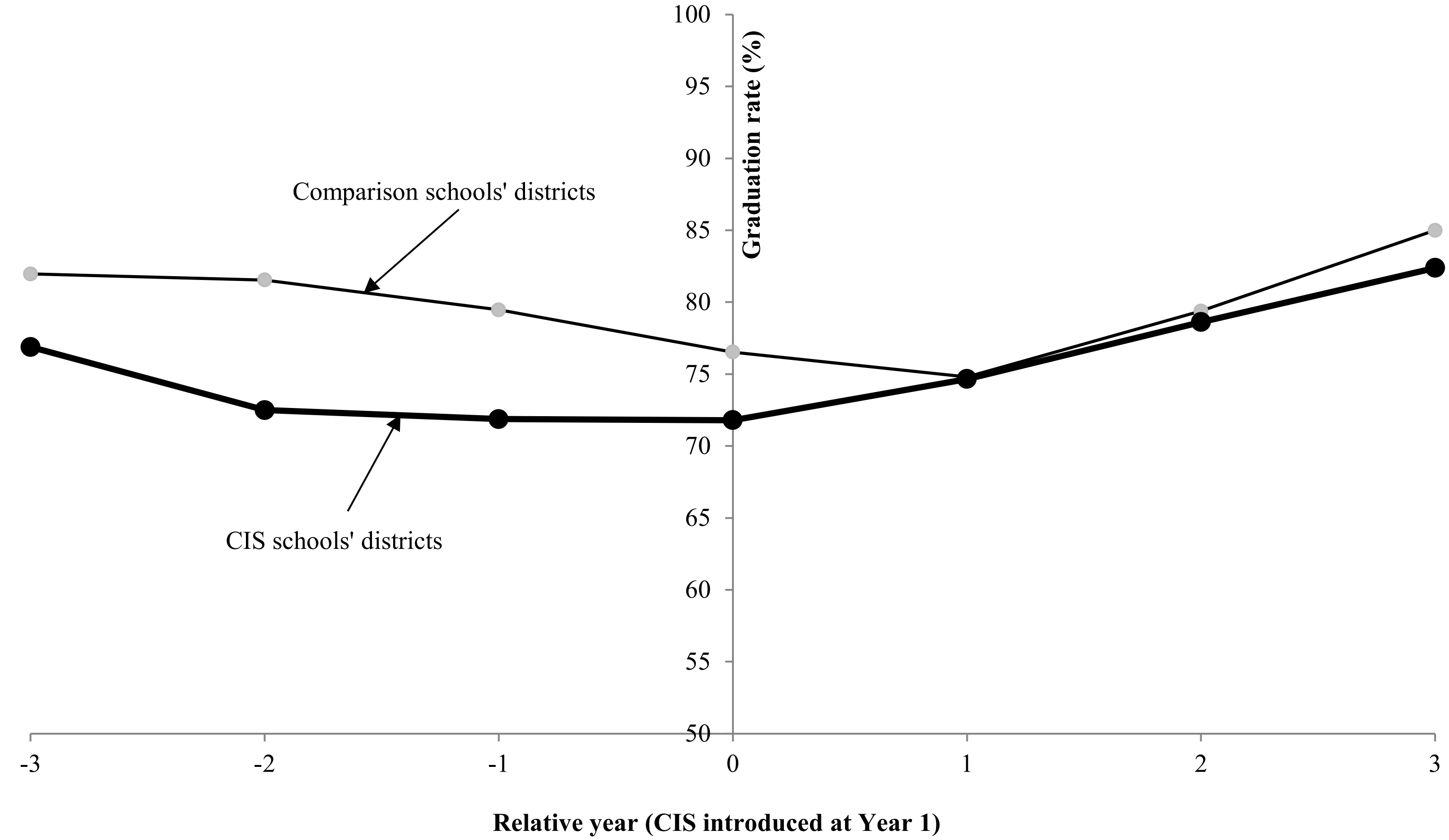
These results are discussed in more detail below. The discussion focuses on graduation rates, dropout rates, elementary school attendance rates, and middle school ELA state test scores, because these are the outcomes where there are statistically significant estimated effects in the main analysis. In general, it does not appear that the estimated effects in this study are confounded with differences in district-level initiatives or reforms.

Appendix Figure E.1 shows the district-level trends in graduation rates for the CIS schools' districts and the comparison schools' districts. (Note that the district-level trends for the CIS districts exclude the CIS schools, in order to also exclude any effect that CIS might be having.) Appendix Figure E.2 shows whether the two groups of districts had higher or lower graduation rates in the follow-up years than their baseline trends predicted.

Overall, it appears that the CIS schools and comparison schools are located in districts that experienced similar trends and changes in their graduation rates over the study period.

Appendix Figure E.1

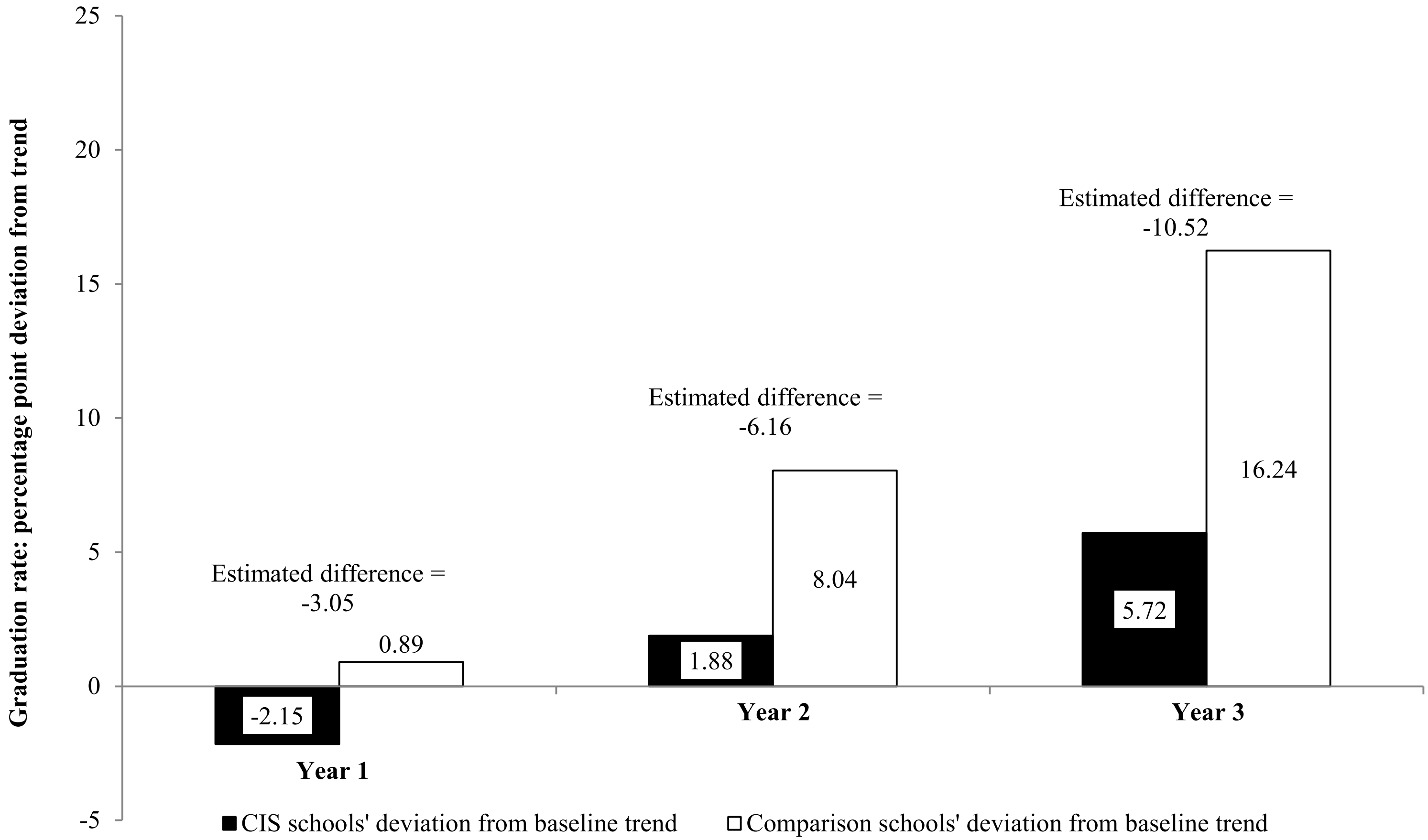
Trends in High School Graduation in Districts Where CIS and Comparison Schools Are Located



SOURCES: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

Appendix Figure E.2

Deviations from Trends in High School Graduation Rates in Districts Where CIS and Comparison Schools Are Located



SOURCES: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTE: No statistical tests were conducted. Rounding may cause slight discrepancies in calculating differences.

Appendix Figure E.1 shows that graduation rates were decreasing in both the CIS districts and the comparison districts in the baseline period. Appendix Figure E.2 shows that districts' deviations from these baseline trends were also similar. In any given follow-up year, the deviations from trend of the CIS districts and the comparison districts are either both very small (as in Year 1) or larger but in the same direction (as in Year 2 and Year 3). The deviations from trend are largest and most positive in the third year. (As noted in Chapter 3, in Year 3 all schools in the study would have been affected by the Great Recession.)

Therefore, the CIS model's estimated effect on graduation rates is not explained by some district-wide initiative that improved graduation rates in the CIS districts more than the comparison districts, nor is it explained by the fact that the Great Recession had a bigger impact on graduation rates in in the CIS districts. Appendix Figure E.2 shows that the CIS districts actually experienced smaller deviations from trend than the comparison districts. For example, in Year 3, graduation rates improved in the CIS districts relative to their baseline trend, but not by as much as they did in the comparison districts. This pattern of findings, if anything, should have made it less likely that there would be estimated effects on the CIS schools' graduation rates.

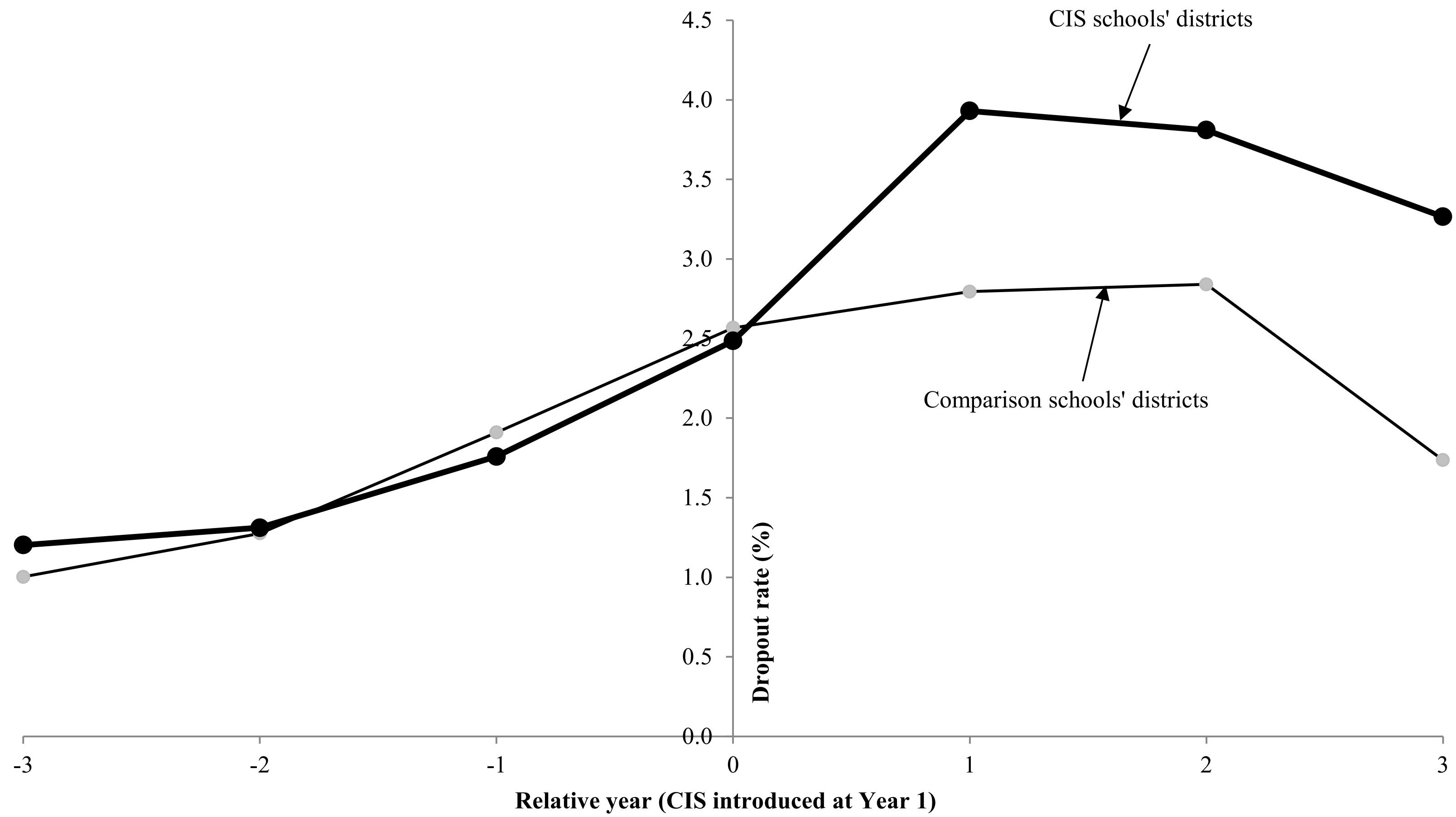
Appendix Figures E.3 and E.4 show the district-level trends and deviations from baseline trends for dropout rates. These figures are the inverse of the results for graduation rates: dropout rates in the CIS districts improved less (relative to baseline trends) than they did in the comparison districts. This pattern of findings suggests that the CIS model's estimated effects on this outcome measure are not confounded with a district-wide initiative that improved dropout rates in the CIS districts relative to the comparison districts.

Appendix Figures E.5 and E.6 show the district-level trends and deviations from baseline trends for middle schools' ELA state test scores. As can be seen in Appendix Figure E.6, the CIS districts experienced only small deviations from trend, much like the CIS study schools themselves. The comparison districts' deviations from baseline trend were positive and larger. The comparison school districts may have initiated some district-wide strategy for improving test scores that improved test scores in the comparison schools by about 0.02 standard deviations more than the CIS schools in Year 3. However, in practice this difference is very small, and it only accounts for a small fraction of the CIS model's negative estimated effect on ELA state test scores in Year 3 (-0.11 standard deviations). This result suggests that the negative effect on state test scores is not simply due to differences in the district contexts of the CIS and comparison schools.

Appendix Figures E.7 and E.8 show the district-level trends and deviations from baseline trends for elementary school attendance rates. Appendix Figure E.8 shows that some sort of

Appendix Figure E.3

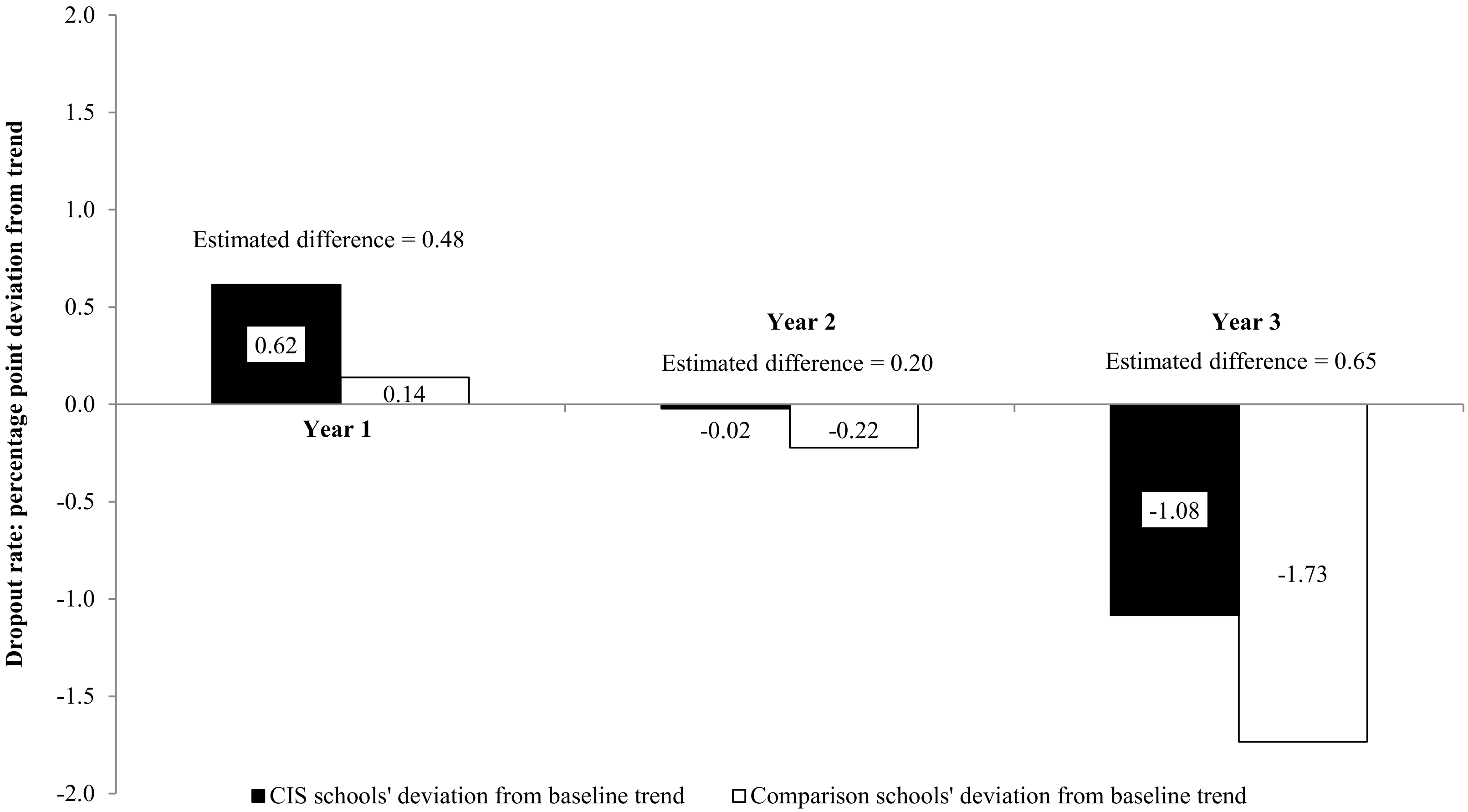
Trends in High School Dropout Rates in Districts Where CIS and Comparison Schools Are Located



SOURCES: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

Appendix Figure E.4

Deviations from Trends in High School Dropout Rates in Districts Where CIS and Comparison Schools Are Located

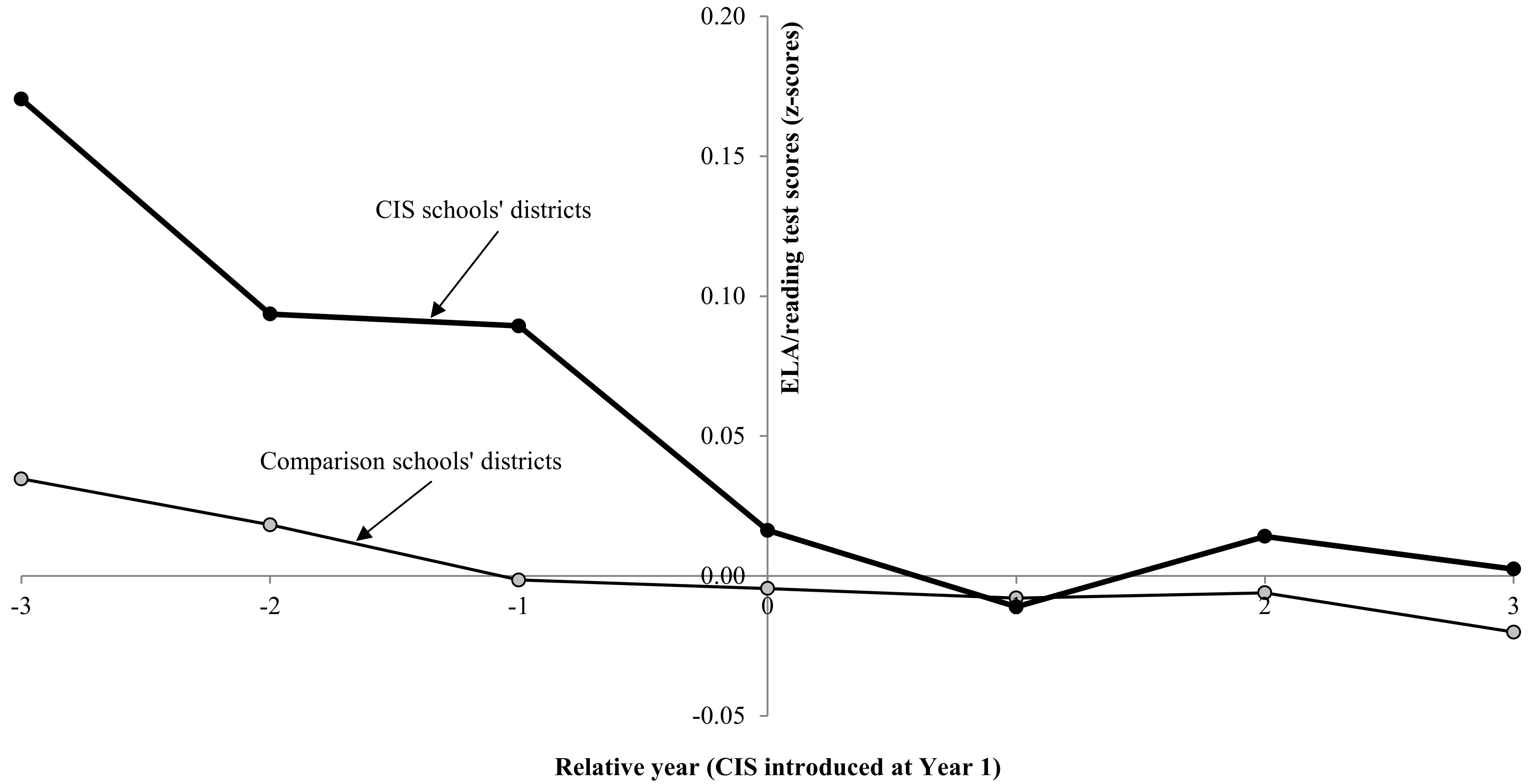


SOURCES: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTE: No statistical tests were conducted. Rounding may cause slight discrepancies in calculating differences.

Appendix Figure E.5

Trends in Middle School ELA State Test Scores in Districts Where CIS and Comparison Schools Are Located

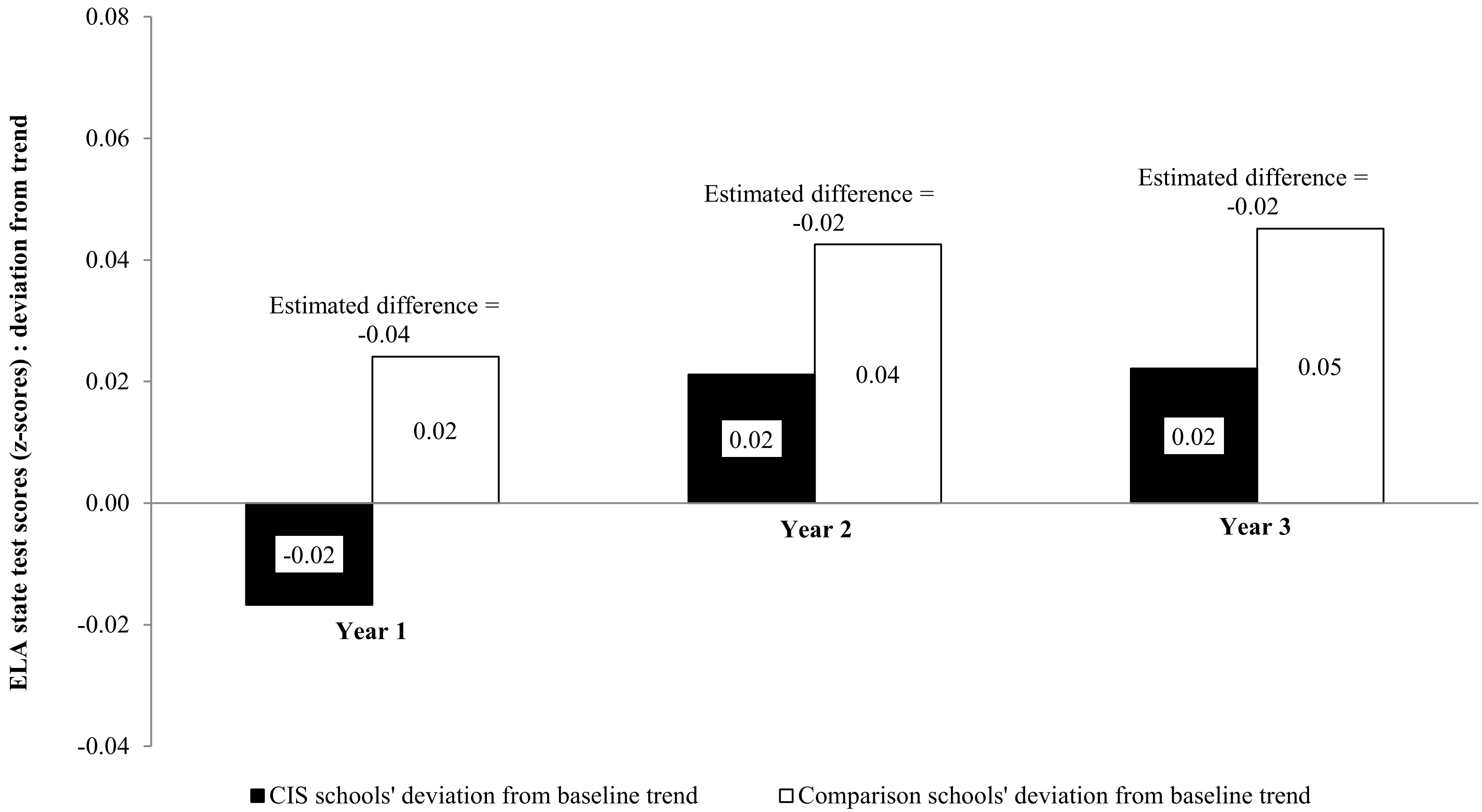


SOURCES: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTE: State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

Appendix Figure E.6

Deviations from Trends in Middle School ELA State Test Scores in Districts Where CIS and Comparison Schools Are Located

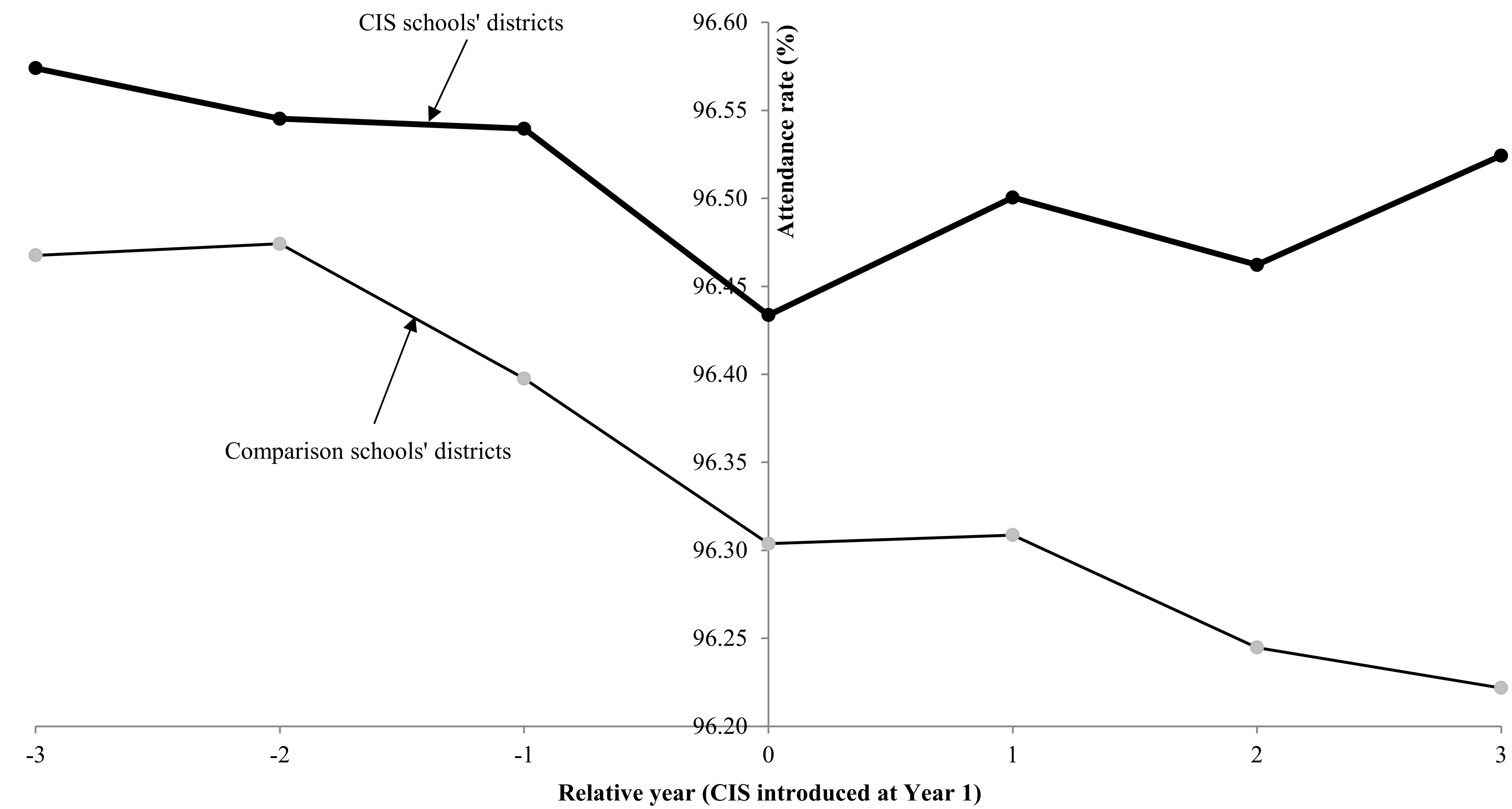


SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTES: No statistical tests were conducted. Rounding may cause slight discrepancies in calculating differences.
State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

Appendix Figure E.7

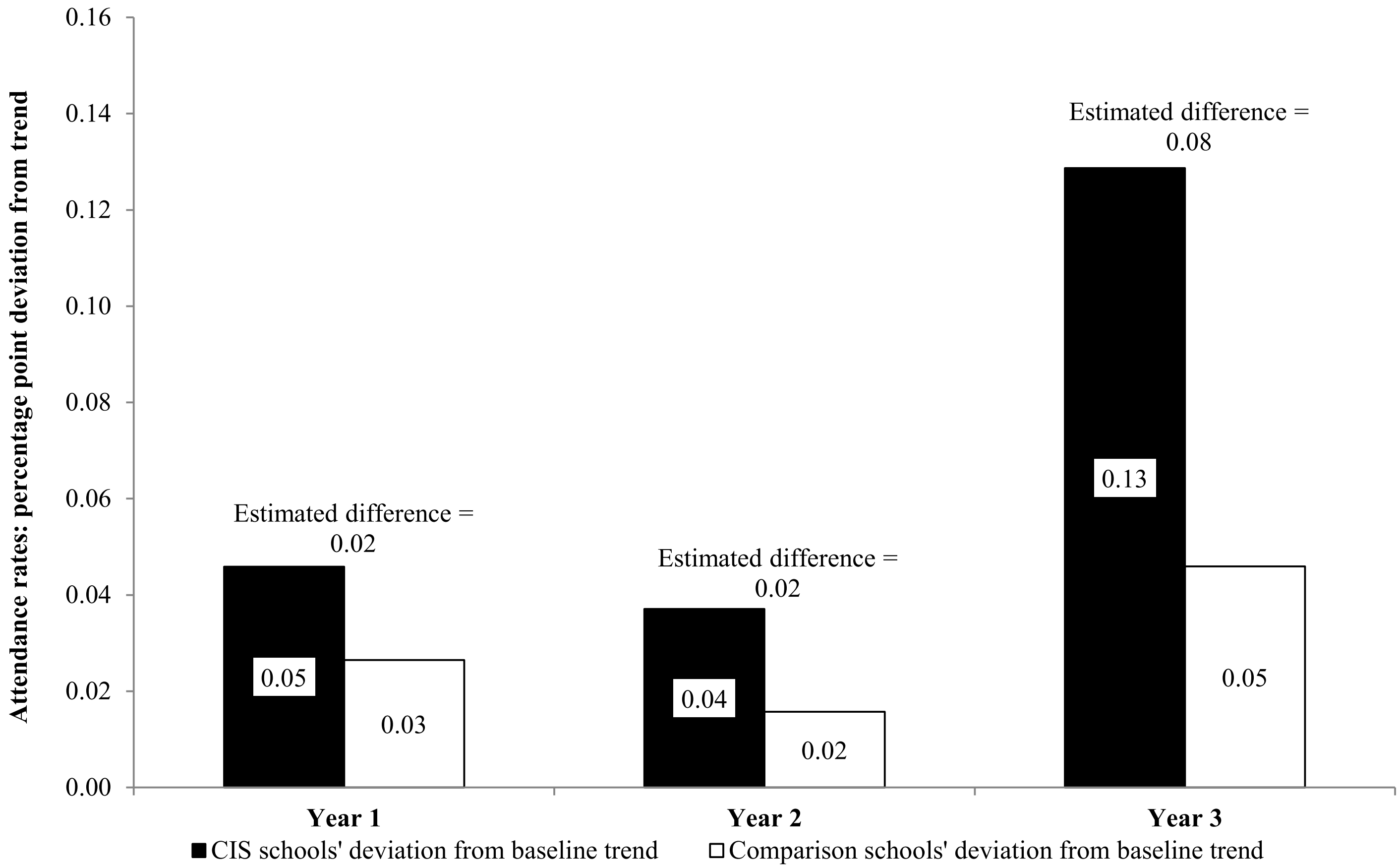
Trends in Elementary School Attendance Rates in Districts Where CIS and Comparison Schools Are Located



SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

Appendix Figure E.8

Deviations from Trends in Elementary School Attendance Rates in Districts Where CIS and Comparison Schools Are Located



SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTE: No statistical tests were conducted. Rounding may cause slight discrepancies in calculating differences.

district-wide initiative or reform may have occurred in Year 3 that improved the CIS districts' attendance rates by about 0.08 percentage points more than the comparison districts improved. However, this difference is still very small, and it only accounts for a small fraction of the CIS model's estimated effect on attendance rates in Year 3 (0.41 percentage points).

Overall, the pattern of effects observed in this study does not appear to be confounded with district-wide policy changes or instructional reforms, or with differences in the CIS or comparison schools' local contexts.

Hypothesis 2: The CITS Design Cannot Provide Credible Results in This Case

As discussed in Chapter 2, the comparison schools are different from the CIS schools in two respects. First, they are located in different school districts than the CIS schools. Second, the comparison schools were performing better than the CIS schools before the model was launched (see Tables 2.4 and 2.5). These differences may compromise the comparison schools' ability to represent accurately what would have happened in the CIS schools had they not implemented the CIS model. By extension, the CITS design may be unable to provide rigorous estimates of the CIS model's effect.

Whether or not this hypothesis is true can be examined by estimating the CIS model's "effect" before it was actually launched. The CIS model's effect in the baseline period is known (it is zero by definition), so it provides a useful benchmark against which to check the choice of comparison schools. Specifically, if the CIS model's estimated "effect" in the last baseline year is not zero, then it would cast doubt on the credibility of the comparison group and of the findings. The sensitivity analysis was conducted as follows.

1. **Redefine the baseline/follow-up period:** The follow-up period in the CITS design is redefined so that the school year immediately before CIS implementation becomes the first follow-up year.

2. **Select comparison schools:** Next, comparison schools are chosen based on the redefined baseline period (which now excludes the school year immediately before CIS implementation). These comparison schools are chosen using the same matching variables as in the main analysis, but using schools' outcome measure values and characteristics in the newly defined baseline period.

3. **Reestimate effects:** Using the comparison schools selected in Step 2, the CIS model's effect is then reestimated for each year of the redefined follow-up period, which

now includes the school year before CIS was implemented, as well as the first three years of CIS implementation.[2]

If the comparison schools provide a credible counterfactual, then the pattern of effects should be as follows:

- The CIS model's "effect" in the new "first follow-up year" (really the last baseline year) should be close to zero in magnitude and not statistically significant.
- The model's estimated effects in the second to fourth follow-up years (the first three years of actual implementation) should be larger than the "effect" in the last baseline year.
- The model's estimated effects in the second to fourth follow-up years should be similar to the effects found in the main impact analysis.

The findings from this analysis are shown in Appendix Tables E.1 to E.3 in the "Design Validation" column, for high schools, middle schools, and elementary schools respectively. As a reference point, the first column of results in each table shows the estimated effects from the main analysis ("Main Sample").

The findings from the design validation analysis are as expected for graduation rates, dropout rates, elementary school attendance rates, and middle school ELA state test scores. Estimated effects on these four outcomes are close to zero in the last baseline year, as they should be. Moreover, estimated effects are larger in the years when CIS was implemented and are similar in size to the effects from the main analysis, even though they are not statistically significant.

On the other hand, the design validation analysis does contradict the finding that the CIS model had a positive effect on high school ELA state test scores. The model's estimated effects on this outcome are small in the last baseline year, as expected (about 0.05 standard deviations), but so are the estimated effects in the years when the CIS model was actually being implemented (Year 1 to Year 3). These results are small and not statistically significant because the comparison schools chosen for this sensitivity analysis performed better than expected in the follow-up period, whereas the comparison schools in the main analysis did not. This finding

---

[2]Because the study period is shifted back, schools in the study sample now have fewer years of baseline data (three to four years rather than four to five years). Although trends are less reliably estimated when there are only three years of baseline data, this limitation was deemed acceptable for the purposes of this sensitivity analysis, so that all study schools could be retained in the analysis.

Appendix Table E.1

Estimated Effects on High School Outcomes, by Sensitivity Analysis

Outcome and Follow-Up Year	Main Sample	Design Validation[a]	Adjusting for School Characteristics[b]	Propensity Score Matching[c]
Graduation rate (%)				
Year 0		*-1.55*		
Year 1	4.44	3.02	2.37	3.30
Year 2	5.51	4.04	3.75	0.79
Year 3	7.50 *	6.67	6.03	4.54
Number of schools	19	18	19	18
Dropout rate (%)				
Year 0		*0.55*		
Year 1	-1.06	-0.77	-1.03	-0.54
Year 2	-1.19	-1.24	-1.21	-0.33
Year 3	-1.49	-1.58	-1.46	-1.09
Number of schools	32	31	32	32
Attendance rate (%)				
Year 0		*-0.63* *		
Year 1	0.04	-0.45	-0.22	0.01
Year 2	0.00	-0.41	-0.22	0.37
Year 3	0.12	-0.28	-0.12	0.69
Number of schools	32	31	32	32
ELA state test scores (z-scores)[d]				
Year 0		*-0.05*		
Year 1	0.08 **	0.04	0.09 **	0.06
Year 2	0.09 *	0.06	0.09 *	0.08
Year 3	0.10 *	0.05	0.08	0.12 **
Number of schools	21	21	21	21

(continued)

Appendix Table E.1 (continued)

Outcome and Follow-Up Year	Main Sample	Design Validation[a]	Adjusting for School Characteristics[b]	Propensity Score Matching[c]
Math state test scores (z-scores)[d]				
Year 0		*-0.05*		
Year 1	0.01	-0.01	0.00	-0.01
Year 2	0.05	0.04	0.04	0.06
Year 3	0.07	0.04	0.02	0.10
Number of schools	21	21	21	21

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the "Main Sample" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. A two-tailed test was applied to estimated effects. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]To estimate the effects in this column, the selection of comparison schools was conducted using the same process as was used for the main study sample, but in this case pretending that CIS schools' last baseline year was their first year of CIS implementation.

[b]Estimated effects in this column are adjusted for the percentage of students at a school in each racial/ethnic group, the percentage of students eligible for free or reduced-price lunches, school enrollment, the pupil/teacher ratio, and the percentage of male students, for each school year in the study.

[c]Estimated effects in this column are based on comparison schools that were chosen using propensity scores as an overall measure of the difference between schools, rather than the Euclidian distance.

[d]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

casts doubt on the idea that the comparison schools in the main analysis provide the right counterfactual trend in state test scores for the CIS schools. That doubt makes it difficult to conclude that the CIS model had effects on high school ELA state test scores.

Appendix Table E.2

Estimated Effects on Middle School Outcomes, by Sensitivity Analysis

Outcome and Follow-Up Year	Main Sample	Design Validation[a]	Adjusting for School Characteristics[b]	Propensity Score Matching[c]
Attendance rate (%)				
Year 0		*0.31*		
Year 1	0.21	0.45	0.21	-0.18
Year 2	-0.01	0.28	0.03	-0.37
Year 3	0.06	0.49	0.24	-0.42
Number of schools	39	38	39	33
ELA state test scores (z-scores)[d]				
Year 0		*-0.07*		
Year 1	-0.05	-0.12 **	-0.04	-0.06
Year 2	-0.08	-0.14 **	-0.07	-0.11
Year 3	-0.11 *	-0.17 ***	-0.10 *	-0.14 *
Number of schools	23	24	23	19
Math state test scores (z-scores)[d]				
Year 0		*-0.06*		
Year 1	-0.08	-0.11 *	-0.09 *	-0.08
Year 2	-0.07	-0.12	-0.09	-0.13
Year 3	-0.10	-0.13	-0.12 *	-0.15
Number of schools	23	24	23	19

(continued)

Hypothesis 3: There Are Group-Specific Changes in School Demographic Characteristics in the Follow-Up Period

Another possible threat to the credibility of this study's findings is that the CIS schools or the comparison schools experienced different demographic shifts after CIS was launched. Such demographic shifts — if they are specific to either the CIS or comparison schools — would make the comparison schools' deviations from trend in the follow-up period a less credible counterfactual for the CIS schools.

Appendix Table E.2 (continued)

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the "Main Sample" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. A two-tailed test was applied to estimated effects. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]To estimate the effects in this column, the selection of comparison schools was conducted using the same process as was used for the main study sample, but in this case pretending that CIS schools' last baseline year was their first year of CIS implementation.

[b]Estimated effects in this column are adjusted for the percentage of students at a school in each racial/ethnic group, the percentage of students eligible for free or reduced-price lunches, school enrollment, the pupil/teacher ratio, and the percentage of male students, for each school year in the study.

[c]Estimated effects in this column are based on comparison schools that were chosen using propensity scores as an overall measure of the difference between schools, rather than the Euclidian distance.

[d]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

To test this hypothesis, a CITS design was used to estimate the "effect" of the CIS model on the demographic characteristics of students in the study schools. This sensitivity analysis finds no evidence of statistically significant effects on students' characteristics, which suggests that the CIS schools and the comparison schools did not experience different demographic shifts in the follow-up period.

As an extension of this analysis, the CIS model's effect on school outcomes was also reestimated, in this case also controlling for schools' demographic and structural characteristics in each year of the study.[3] As shown in Appendix Tables E.1 to E.3 in the "Adjusted for School Characteristics" column, adjusting for school characteristics does not appreciably change the size of the estimated effects on high school graduation rates, dropout rates, elementary school attendance rates, or middle school ELA state test scores (even though in some cases the results are no longer statistically significant). This result suggests that the estimated effects from the main analysis are not confounded with a group-specific change in the population of students served by CIS or comparison schools in the follow-up period.

---

[3]The analysis controlled for the percentage of students in each racial or ethnic group, the percentage of students eligible for free or reduced-price lunches, school enrollment, the pupil/teacher ratio, and the percentage of male students.

Appendix Table E.3

Estimated Effects on Elementary School Outcomes, by Sensitivity Analysis

Outcome and Follow-Up Year	Main Sample	Design Validation[a]	Adjusting for School Characteristics[b]	Propensity Score Matching[c]
Attendance rate (%)				
Year 0		*-0.06*		
Year 1	0.16	0.28	0.14	0.12
Year 2	0.27	0.35 *	0.25	0.22
Year 3	0.41 **	0.58 **	0.40 **	0.29
Number of schools	60	61	60	58
ELA state test scores (z-scores)[d]				
Year 0		*0.00*		
Year 1	0.02	-0.02	0.02	-0.03
Year 2	0.05	-0.01	0.04	-0.02
Year 3	0.04	-0.03	0.03	-0.03
Number of schools	53	53	53	50
Math state test scores (z-scores)[d]				
Year 0		*0.03*		
Year 1	-0.01	0.00	-0.01	-0.07
Year 2	-0.08	-0.03	-0.09	-0.10
Year 3	-0.08	-0.06	-0.08	-0.08
Number of schools	53	53	53	50

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the "Main Sample" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. A two-tailed test was applied to estimated effects. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]To estimate the effects in this column, the selection of comparison schools was conducted using the same process as was used for the main study sample, but in this case pretending that CIS schools' last baseline year was their first year of CIS implementation.

[b]Estimated effects in this column are adjusted for the percentage of students at a school in each racial/ethnic group, the percentage of students eligible for free or reduced-price lunches, school enrollment, the pupil/teacher ratio, and the percentage of male students, for each school year in the study.

[c]Estimated effects in this column are based on comparison schools that were chosen using propensity scores as an overall measure of the difference between schools, rather than the Euclidian distance.

[d]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

Hypothesis 4: The Matching Method Did Not Result in a Credible Group of Comparison Schools

As explained in Appendix C, the main analysis in this study uses the Euclidian distance as a composite measure of the "dissimilarity" between the CIS and non-CIS schools for the purposes of matching. An alternative (and more popular) measure of distance is the propensity score.[4] The propensity score is a weighted composite of a school's characteristics and values on outcome measures, where the weight of each characteristic or measure is proportional to its ability to predict whether or not a school adopts the CIS model. Two schools with the same propensity score have the same predicted probability of adopting the CIS model, based on their baseline characteristics and values on outcome measures. The difference between two schools' propensity scores provides a measure of their "dissimilarity." The propensity score was not used to choose comparison schools in this study because it is trickier to use this approach when schools have begun implementing an intervention in different years (as they did in this study).

The literature on matching indicates that the choice of matching method should not matter, as long as the pool of comparison schools from which to choose is suitable and as long as preintervention data on the outcome measures of interest are available.[5] It is therefore a useful sensitivity analysis to determine whether the overall findings change when a different distance metric is used to select comparison schools. If the findings do change, then it would indicate that the comparison schools may not provide a credible counterfactual for the CIS schools.

For this sensitivity analysis, comparison schools were chosen using propensity score matching based on the same matching outcome measures and characteristics that were used to calculate the Euclidian distance. The propensity score was calculated by fitting the following logistic regression model to a cross-sectional data set that includes the CIS schools and the non-CIS schools in the comparison pool:

$$logit(CIS_j) = \alpha + \sum_k \beta_{kt} CHAR_{kt} + \sum_{2006}^{2009} \varphi_d STARTYR_j + \varepsilon_j$$

where j denotes schools.

The variables are defined as follows:

[4]Rosenbaum and Rubin (1983).

[5]Cook, Shadish, and Wong (2008); Somers, Zhu, Jacob, and Bloom (2013).

CIS_j = Dichotomous indicator for whether school *j* is a CIS school (equal to 1 if it is a CIS school, 0 if a school in the comparison pool)

$CHAR_{kt}$ = School characteristic or outcome measure *k* used for matching. These are the same characteristics and outcome measures that are used to calculate the Euclidian distance (see Appendix C). For characteristics where the value in the last baseline year is used for matching, the variable in the propensity score model is the value of that characteristic in the last baseline year. For characteristics where the entire baseline slope is the matching variable, there are five variables in the propensity score model, one for each baseline year.

$STARTYR_j$ = A set of dichotomous indicators for the first year of CIS implementation (2005-2006 to 2008-2009) for CIS schools and the comparison schools in their pool. These indicators are equal to 1 if a school started in year *d* and 0 otherwise.

ε_j = Random error term for school *j*.

The estimated coefficients from this logistic regression represent the relationship between schools' outcome measures/characteristics and their log odds of adopting the CIS model.[6] The coefficients from the regression can then be used to obtain the predicted probability that a school will adopt the CIS model given its baseline characteristics and outcome measure values: the *propensity score*.[7]

---

[6]Further controlling for start year (*STARTYR*) ensures that the estimated coefficients represent the average relationship between the matching characteristics and the probability of adopting the CIS model within each start year.

[7]In practice, the *logit* of the propensity score is used for matching, for three reasons. Because the logit transformation makes the propensity score linear, it is more relevant for assessing the results of linear modeling adjustments. Second, linear propensity scores tend to yield distributions with more similar variances and symmetry. Third, linear propensity scores are easier to relate to benchmarks in the literature on adjustments for covariates, which are based on linearity assumptions. See Rubin (2001).

To conduct propensity score matching, the data had to be reshaped and restructured. First, the panel data set for the study was "flattened," so that there was one row per school and characteristics and outcome measures that varied over time were expressed as multiple variables (DROPOUT2002, DROPOUT2003, etc.). Next, a data set was created for each CIS start year (2005-2006 to 2008-2009) that included the CIS schools
(continued)

Propensity scores were calculated for elementary, middle, and high schools and for schools in Texas and North Carolina. These scores were then used as the metric for choosing the two comparison schools most "similar" to the CIS schools.

The CIS model's effect was reestimated using a CITS design, based on these new comparison schools. On average, the comparison schools selected using propensity score matching have similar baseline characteristics and values on outcome measures as the sample of comparison schools chosen using the Euclidian distance. However, the actual comparison schools in the two samples are different. Only 31 percent of the comparison schools in the main analysis are also in the propensity score group.

As shown in Appendix Tables E.1 to E.3 in the "Propensity Score Matching" column, the CIS model's estimated effects on elementary school attendance rates and middle school ELA state test scores are similar in size to the estimated effects from the main analysis. However, the CIS model's estimated effects on dropout rates and graduation rates are smaller when comparison schools are chosen using propensity score matching. The estimated effect on graduation rates in this sensitivity analysis is 4.5 percentage points in Year 3 (compared with 7.5 percentage points in the main analysis), and the estimated effect on dropout rates is -1.1 percentage points in Year 3 (compared with -1.5 percentage points in the main analysis). These smaller effects reflect the fact that the comparison schools selected using propensity score matching performed better than expected during the follow-up period (that is, their deviations from trend are slightly larger and more positive than those of the comparison schools used in the main analysis). These results call into question the credibility of the comparison schools used in the main analysis to estimate effects on graduation and dropout rates. They also align with the results of the sensitivity analyses that were conducted to explore the graduation-rate findings specifically (see the next section).

---

that started in that year and all comparison schools in the pool. In each of these data sets, the outcome variables were renamed to reflect the school year relative to the CIS start year of the CIS schools in that data set. For example, DROPOUT5 would be the dropout rate five years before the start of CIS, DROPOUT4 would be the dropout rate four years before, etc. With the variables scaled in relative terms, it was then possible to append the four data sets (one for each start year) together. The final data set includes all CIS schools as well as the comparison schools in the pool, each of which appears four times in the data set (once for each possible CIS start year). Note that if the CIS schools had all started implementing the CIS model at the same time, then it would not have been necessary to include the comparison schools multiple times.

Hypothesis 5: The Comparison Schools Have Less Incentive to Improve their Outcomes

High School Graduation and Dropout Rates

As discussed in Chapter 3, there are subgroup differences in the CIS model's estimated effect on graduation and dropout rates. The CIS model's estimated effect on these outcomes is not statistically significant for the CIS high schools that had higher baseline graduation rates (and lower baseline dropout rates), whereas it is large and statistically significant for the CIS high schools that were performing worse (see Table 3.2).

There could be large estimated effects for the latter group because the CIS model is more effective for lower-performing schools, or there could be large effects because the comparison schools do not provide the right counterfactual and the CIS models' effect is overstated. Appendix Table E.4 presents the baseline trends in graduation and dropout rates for the higher- and lower-performing CIS high schools (see Appendix A for a discussion of how baseline trends are estimated). As shown in this table, the higher-performing CIS high schools and their comparison schools had similar baseline trends in their graduation and dropout rates, whereas the three lowest-performing CIS high schools had much lower graduation rates (and higher dropout rates) than their comparison schools, and their graduation rates had been decreasing more rapidly when they began implementing the model.[8]

It is an essential question whether the comparison schools for the three lowest-performing CIS schools provide the right counterfactual, and whether the large estimated effects for this subgroup overstate the true effect of the CIS model. To explore this hypothesis, a new set of comparison schools was chosen for the three lowest-performing CIS schools, this time by matching their baseline graduation rates *only* (as opposed to selecting matches using multiple outcome measures and school characteristics). To maximize the quality of the match, only one comparison school (and not two) was chosen for each CIS school.

After re-matching, there remains a large difference in baseline graduation rates between the three CIS schools and their new comparison schools, due to limitations in the pool of non-CIS schools. However, as shown in Appendix Table E.5 (compared with Appendix Table E.4), the new comparison schools have slightly more similar baseline graduation rates, and their baseline graduation rates were declining at the same rate on average as the CIS schools.

[8]The baseline difference in graduation rates between this subgroup of schools and their comparison schools is greater than 0.25 standard deviations, violating the What Works Clearinghouse standard for baseline equivalence. See What Works Clearinghouse (2014a).

Appendix Table E.4

Baseline Trends in High School Graduation and Dropout Rates, by Match Quality and Baseline Outcome Measures

Outcome Measure	CIS Schools	Comparison Schools	Estimated Difference	P-Value
Schools that are well matched and higher performing				
Graduation rate (%)				
In the last baseline year (intercept)	76.7	76.8	-0.1	0.913
Annual change (slope)	-1.9	-2.4	0.5	0.108
Number of schools	5	9		
Dropout rate (%)				
In the last baseline year (intercept)	3.2	2.9	0.3	0.146
Annual change (slope)	0.5	0.5	-0.1	0.229
Number of schools	11	17		
Schools that are not well matched and are lower performing				
Graduation rate (%)				
In the last baseline year (intercept)	50.1	79.8	-29.8 ***	<0.001
Annual change (slope)	-5.2	-2.1	-3.2 ***	<0.001
Number of schools	3	5		
Dropout rate (%)				
In the last baseline year (intercept)	8.5	4.1	4.4 ***	<0.001
Annual change (slope)	1.4	1.0	0.3	0.129
Number of schools	3	5		

SOURCE: MDRC calculations based on school-level data on student outcome measures from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated intercepts and slopes of the baseline trends for CIS and comparison schools, based on four or five years of preintervention data on the outcome measure of interest. The values in the "Estimated Difference" column are the differences between CIS and comparison schools with respect to the intercepts and slopes of their baseline trends. Rounding may cause slight discrepancies in calculating differences.

A two-tailed test was applied to estimated differences between CIS schools and comparison schools. The statistical significance of estimated differences is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

Appendix Table E.5

Baseline Trends in High School Graduation and Dropout Rates for Schools That Are Not Well Matched in the Main Analysis, After Re-Matching

Outcome Measure	CIS Schools	Comparison Schools	Estimated Difference		P-Value
Graduation rate (%)					
In the last baseline year (intercept)	50.1	76.6	-26.5	***	<0.001
Annual change (slope)	-5.2	-4.3	-0.9		0.218
Number of schools	3	2			
Dropout rate (%)					
In the last baseline year (intercept)	8.5	4.5	4.0	***	<0.001
Annual change (slope)	1.4	1.2	0.2		0.571
Number of schools	3	1			

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated intercepts and slopes of the baseline trends for CIS and comparison schools, based on four or five years of preintervention data on the outcome measure of interest. The values in the "Estimated Difference" column are the differences between CIS and comparison schools with respect to the intercepts and slopes of their baseline trends. Rounding may cause slight discrepancies in calculating differences.

A two-tailed test was applied to estimated differences between CIS schools and comparison schools. The statistical significance of estimated differences is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

As shown in Appendix Table E.6, the CIS model's estimated effect on graduation rates based on this new group of comparison schools is half of its original size (12.3 percentage points, compared with 27.5 percentage points in the main analysis — see Table 3.2). This change reflects the fact that the new comparison schools' average graduation rate also improved in the follow-up period (they deviated from their baseline trend by 12.7 percentage points).

A similar sensitivity test was conducted for dropout rates — matching schools using baseline trends in dropout rates only. Associated results are shown in Appendix Tables E.5 and E.6. Based on the new comparison schools, the estimated effect on dropout rates for the three lowest-performing CIS schools is -2 percentage points, compared with -7 percentage points in the original sample (see Table 3.2).

Appendix Table E.6

Estimated Effects on High School Graduation and Dropout Rates for Schools That Are Not Well Matched in the Main Analysis, After Re-Matching

	Deviation from Baseline Trend			P-Value for
	CIS	Comparison	Estimated	Estimated
Outcome and Follow-Up Year	Schools	Schools	Effect	Effect
Graduation rate (%)				
Year 1	7.82 †††	-1.92	9.74 *	0.069
Year 2	11.94 †††	-1.70	13.64 **	0.029
Year 3	25.03 †††	12.73 ††	12.30 *	0.075
Number of schools	3	2		
Dropout rate (%)				
Year 1	-2.45	1.04	-3.49	0.160
Year 2	-6.06 †††	-2.80	-3.26	0.255
Year 3	-8.90 †††	-6.77 ††	-2.13	0.506
Number of schools	3	1		

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated deviations from baseline trend for each group of schools, by follow-up year. The values in the "Estimated Effect" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed test was applied to estimated deviations and estimated differences between CIS schools and comparison schools. The statistical significance of estimated deviations is indicated as: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

These results suggest that the estimated effects for the three lowest-performing schools may overstate the CIS model's true effect for this group of schools. However, it is not possible to determine by how much the effect is overstated.

Elementary School Attendance Rates

The baseline difference in attendance rates at the elementary school level is larger than conventional standards (that is, it is greater than 0.25 standard deviations). An additional sensitivity test was therefore conducted in which a new set of comparison schools was chosen for all of the CIS elementary schools, this time by matching their baseline attendance rates only.[9] As shown in Appendix Table E.7, after re-matching there is no baseline difference in attendance rates between the CIS schools and the new comparison schools, nor is there any difference in the rate at which attendance rates were declining in the baseline period. As shown in Appendix Table E.8, the CIS model's estimated effects on attendance rates based on this new sample of comparison schools are positive and consistent with the main findings.

Appendix Table E.7

Baseline Trends in Elementary School Attendance Rates, After Re-Matching

Outcome Measure	CIS Schools	Comparison Schools	Estimated Difference	P-Value
Attendance rate (%)				
In the last baseline year (intercept)	96.1	96.1	0.0	0.205
Annual change (slope)	-0.1	-0.1	0.0	0.316
Number of schools	24	29		

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated intercepts and slopes of the baseline trends for CIS and comparison schools, based on four or five years of preintervention data on the outcome measure of interest. The values in the "Estimated Difference" column are the differences between CIS and comparison schools with respect to the intercepts and slopes of their baseline trends. Rounding may cause slight discrepancies in calculating differences.

A two-tailed test was applied to estimated differences between CIS schools and comparison schools. The statistical significance of estimated differences is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[9]This difference does not appear to be caused by any particular group of schools, so all of the elementary schools were re-matched.

Appendix Table E.8

Estimated Effects on Elementary School Attendance Rates, After Re-Matching

	Deviation from Baseline Trend			
Outcome and Follow-Up Year	CIS Schools	Comparison Schools	Estimated Effect	P-Value for Estimated Effect
Attendance rate (%)				
Year 1	0.30 †††	0.20 ††	0.11	0.465
Year 2	0.44 †††	0.29 †††	0.15	0.357
Year 3	0.61 †††	0.35 †††	0.26	0.178
Number of schools	24	29		

SOURCES: MDRC calculations based on school-level data on student outcomes from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Research Data Center at Duke University.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated deviations from baseline trend for each group of schools, by follow-up year. The values in the "Estimated Effect" column are the differences between CIS schools and comparison schools with respect to their deviations from baseline trend. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed test was applied to estimated deviations and estimated differences between CIS schools and comparison schools. The statistical significance of estimated deviations is indicated as: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

Appendix F

Findings for Schools in Texas

This appendix provides the results of a subgroup analysis based on the middle schools and elementary schools in Texas. Results for Texas high schools are not shown because all but one of the 14 Communities In Schools (CIS) high schools in the study are located in Texas. As explained in Chapter 1, the implementation of the CIS model is decentralized and therefore CIS National can have different relationships with different states' affiliates. For this reason, it is important to determine whether the CIS model has a larger effect in some states than others. The subsample of schools in Texas in this study is large enough to examine the CIS model's average effect in that state. The following tables are included in this appendix:

- Appendix Tables F.1 and F.2 present the baseline characteristics and outcome measures of CIS and comparison schools in Texas. These tables show that, like the schools in the full study sample, the Texas CIS schools serve more low-income students than the comparison schools, but their baseline attendance rates and scores on state tests are similar, though in some cases there are statistically significant differences between them.

- Appendix Tables F.3 and F.4 present the middle and elementary schools' deviations from their baseline trends and the CIS model's estimated effects in each follow-up year. The magnitude and pattern of estimated effects for schools in Texas is similar to those observed for the full study sample.

Appendix Table F.1

Characteristics and Outcome Measures of CIS Schools and Comparison Schools in the Last Baseline Year, Middle Schools in Texas

Baseline Characteristic or Outcome Measure	CIS Schools	Comparison Schools	Estimated Difference		P-Value
Structural and demographic characteristics					
Free or reduced-price lunch (%)	57.7	50.0	7.6	**	0.025
Racial/ethnic composition (%)					
Black	26.2	26.6	-0.4		0.924
Hispanic	42.8	38.1	4.7		0.368
White	26.8	31.2	-4.4		0.296
Other	4.2	4.2	0.1		0.951
Enrollment	798.2	961.0	-162.8		0.145
Pupil/teacher ratio	13.9	16.0	-2.1	***	0.004
Title I status[a] (%)	53.8	53.8	0.0		NA
Location (%)					
City/town	92.3	92.3	0.0		NA
Rural	7.7	7.7	0.0		NA
Student outcome measures					
Attendance rate (%)	95.5	95.9	-0.4	**	0.032
Proficient on state tests (%)					
ELA	81.2	81.2	0.0		0.996
Math	69.1	68.4	0.7		0.779
State test scores (z-scores)[b]					
ELA	-0.29	-0.22	-0.07		0.165
Math	-0.14	-0.15	0.01		0.903
Number of schools	13	20			

(continued)

Appendix Table F.1 (continued)

SOURCES: MDRC calculations based on data from the Common Core of Data and school-level data on student outcomes from the Texas Education Agency website.

NOTES: The values in the column labeled "CIS Schools" are the observed means in the last baseline year for schools that implemented CIS. The "Comparison Schools" values in the next column are the observed means in the last baseline year for matched comparison schools. The values in the "Estimated Difference" column are the differences between CIS and comparison schools in the last baseline year. Rounding may cause slight discrepancies in calculating sums and differences. The percentage of Title I and urban schools is exactly the same in the CIS and comparison schools, because CIS and comparison schools were matched exactly with respect to their Title I statuses and their locations (rural/urban).

A two-tailed t-test was applied to differences between CIS and comparison schools. Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

NA = not applicable.

[a]This percentage represents the Title I statuses of study schools in the baseline period. By the third follow-up year, the percentage of Title I-eligible schools had risen. Among middle schools, 62 percent of CIS schools and 65 percent of comparison schools were eligible for Title I funds by the third follow-up year.

[b]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

Appendix Table F.2

Characteristics and Outcome Measures of CIS Schools and Comparison Schools in the Last Baseline Year, Elementary Schools in Texas

Baseline Characteristic or Outcome Measure	CIS Schools	Comparison Schools	Estimated Difference		P-Value
Structural and demographic characteristics					
Free or reduced-price lunch (%)	79.8	75.2	4.7	**	0.018
Racial/ethnic composition (%)					
Black	17.7	18.7	-1.0		0.556
Hispanic	62.3	61.7	0.6		0.839
White	16.5	16.9	-0.4		0.879
Other	3.5	2.7	0.8		0.400
Enrollment	637.3	649.9	-12.6		0.782
Pupil/teacher ratio	14.7	14.0	0.8	*	0.054
Title I status (%)	100.0	100.0	0.0		NA
Location (%)					
City/town	100.0	100.0	0.0		NA
Rural	0.0	0.0	0.0		NA
Student outcome measures					
Attendance rate (%)	96.1	96.5	-0.3	**	0.026
Proficient on state tests (%)					
ELA	81.3	83.0	-1.7	**	0.020
Math	79.6	80.3	-0.7		0.616
State test scores (z-scores)[a]					
ELA	-0.28	-0.26	-0.02		0.535
Math	-0.13	-0.16	0.03		0.349
Number of schools	19	27			

(continued)

Appendix Table F.2 (continued)

SOURCES: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from the Texas Education Agency website.

NOTES: The values in the column labeled "CIS Schools" are the observed means in the last baseline year for schools that implemented CIS. The "Comparison Schools" values in the next column are the observed means in the last baseline year for matched comparison schools. The values in the "Estimated Difference" column are the differences between CIS and comparison schools in the last baseline year. Rounding may cause slight discrepancies in calculating sums and differences. The percentage of Title I and urban schools is exactly the same in the CIS and comparison schools, because CIS and comparison schools were matched exactly with respect to their Title I statuses and their locations (rural/urban).

A two-tailed t-test was applied to differences between CIS and comparison schools. Statistical significance levels are indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

NA = not applicable.

[a]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

Appendix Table F.3

Estimated Effects on Attendance Rates and State Test Scores, Middle Schools in Texas

	Deviation from Baseline Trend						
Outcome	CIS Schools	Comparison Schools		Estimated Effect	Lower 90% CI	Upper 90% CI	P-Value for Estimated Effect
Attendance rate (%)							
Year 1	0.20	-0.06		0.26	-0.13	0.65	0.274
Year 2	0.08	0.06		0.02	-0.45	0.49	0.935
Year 3	0.38	0.27		0.11	-0.44	0.66	0.752
Number of schools	13	20					
ELA state test scores (z-scores)[a]							
Year 1	0.01	0.05		-0.04	-0.12	0.03	0.307
Year 2	0.04	0.11	†††	-0.07	-0.16	0.01	0.149
Year 3	0.02	0.13	†††	-0.11 *	-0.21	-0.01	0.084
Number of schools	6	11					
Math state test scores (z-scores)[a]							
Year 1	0.00	0.05		-0.05	-0.16	0.05	0.412
Year 2	0.02	0.07		-0.05	-0.17	0.08	0.529
Year 3	-0.01	0.08		-0.09	-0.25	0.06	0.316
Number of schools	6	11					

SOURCE: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from the Texas Education Agency website.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated deviations from baseline trend for each group of schools. The values in the "Estimated Effect" column are the difference between CIS schools and comparison schools with respect to their deviations from baseline trend. The values in the "Lower 90% CI" and "Upper 90% CI" columns are the 90% confidence intervals for the estimated effects. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed test was applied to estimated deviations and estimated differences between CIS schools and comparison schools. The statistical significance of estimated deviations is indicated as: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

Appendix Table F.4

Estimated Effects on Attendance Rates and State Test Scores, Elementary Schools in Texas

	Deviation from Baseline Trend							
Outcome	CIS Schools		Comparison Schools		Estimated Effect	Lower 90% CI	Upper 90% CI	P-Value for Estimated Effect
Attendance rate (%)								
Year 1	0.35	†††	0.09		0.25 *	0.01	0.50	0.082
Year 2	0.42	†††	-0.01		0.43 ***	0.17	0.69	0.007
Year 3	0.61	†††	0.19	†	0.42 **	0.11	0.73	0.027
Number of schools	19		27					
Reading state test scores (z-scores)[a]								
Year 1	0.02		0.01		0.01	-0.08	0.10	0.844
Year 2	0.04		0.01		0.04	-0.07	0.14	0.572
Year 3	0.08		0.06		0.01	-0.12	0.14	0.866
Number of schools	16		23					
Math state tests scores (z-scores)[a]								
Year 1	0.01		-0.01		0.01	-0.11	0.14	0.859
Year 2	-0.04		0.06		-0.10	-0.26	0.05	0.253
Year 3	-0.06		0.05		-0.11	-0.29	0.07	0.304
Number of schools	16		23					

SOURCE: MDRC calculations based on data from the Common Core of Data and on school-level data on student outcomes from the Texas Education Agency website.

NOTES: The values in the "CIS Schools" and "Comparison Schools" columns are the estimated deviations from baseline trend for each group of schools. The values in the "Estimated Effect" column are the difference between CIS schools and comparison schools with respect to their deviations from baseline trend. The values in the "Lower 90% CI" and "Upper 90% CI" columns are the 90% confidence intervals for the estimated effects. Rounding may cause slight discrepancies in calculating sums and differences.

A two-tailed test was applied to estimated deviations and estimated differences between CIS schools and comparison schools. The statistical significance of estimated deviations is indicated as: ††† = 1 percent; †† = 5 percent; † = 10 percent. The statistical significance of estimated effects is indicated as: *** = 1 percent; ** = 5 percent; * = 10 percent.

[a]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

Appendix G

Statistical Power and Minimum Detectable Effects

A common way to convey a study's statistical power is through the minimum detectable effect (MDE). Formally, the MDE is the smallest true program impact that can be detected with a reasonable degree of power (in this case, 80 percent) for a given level of statistical significance (in this case, 10 percent for a two-tailed test). In a comparative interrupted time series (CITS) design, the MDE decreases (that is, smaller effects can be detected) when the number of schools in the study increases or when the number of years of baseline data increases. All else being equal, the MDE is larger for later follow-up years (that is, larger for estimated effects in Year 3 than Year 1), because the reliability of trend projections decreases the further the projection is extended through time.[1]

Appendix Table G.1 shows the MDEs for this study, as well as the number of schools included in the analysis for each outcome measure.[2] As a reference point, the table also shows the mean value of each outcome in the last baseline year. Note that the sample size — which is an important determinant of the MDE — varies from outcome measure to outcome measure depending on the amount of historical data available. For attendance rates, for example, data are available as early as the 2001-2002 school year, and therefore effects can be estimated for all Communities In Schools (CIS) schools in the study. However, for other outcomes, historical data are not available for as many school years, so the CIS schools that started implementing CIS earlier are excluded from the analysis (because they do not have four years of baseline data).[3]

Since CIS focuses on reducing the number of school dropouts, the two outcomes of greatest relevance are high school dropout rates and graduation rates. As shown in Table G.1, the study can detect effects ranging from 2 percentage points to 3 percentage points on dropout rates, depending on the follow-up year. The average dropout rate in the CIS study schools was about 5 percent before CIS was launched, so a reduction of 2 percentage points to 3 percentage points could be difficult for CIS to achieve, especially within three years. However, the study can detect effects on graduation rates ranging from 7 percentage points to 11 percentage points, which may be achievable because there was more room for improvement in this area: CIS schools' mean baseline graduation rate was 68 percent.

---

[1]Bloom (1999).

[2]The MDEs were obtained by multiplying the standard error of the estimated impact in each follow-up year by 2.5.

[3]For example, graduation-rate data are only available from 2002-2003 onward and only in Texas. Effects on this outcome cannot be estimated for the subset of Texas high schools that started implementing the CIS model in 2005-2006.

Appendix Table G.1

Minimum Detectable Effects by Outcome Measure and Follow-Up Year

	Number of Schools		Baseline	Minimum Detectable Effect		
Outcome Measure	CIS	Comparison	Mean[a]	Year 1	Year 2	Year 3
High schools						
Graduation rate (%)	8	11	68.1	7.44	9.47	10.92
Dropout rate (%)	14	18	5.1	-2.03	-2.35	-2.95
Attendance rate (%)	14	18	92.6	1.01	1.10	1.25
ELA state test scores (z-scores)[b]	9	12	-0.19	0.10	0.12	0.15
Math state test scores (z-scores)[b]	9	12	-0.18	0.13	0.13	0.16
Middle schools						
Attendance rate (%)	15	24	95.4	0.54	0.65	0.76
ELA state test scores (z-scores)[b]	8	15	-0.20	0.10	0.12	0.14
Math state test scores (z-scores)[b]	8	15	-0.10	0.13	0.15	0.19
Elementary schools						
Attendance rate (%)	24	36	95.9	0.35	0.42	0.47
ELA state test scores (z-scores)[b]	21	32	-0.33	0.11	0.13	0.17
Math state test scores (z-scores)[b]	21	32	-0.18	0.17	0.20	0.23

SOURCES: MDRC calculations based on school-level data on student outcome measures from state websites (the Texas Education Agency and the North Carolina State Department of Education) and from the North Carolina Education Research Data Center at Duke University.

NOTES: The minimum detectable effects in this table are calculated by multiplying the standard error of the estimated effects by 2.5. A statistical significance level of 10 percent is assumed.

ELA = English/language arts.

[a]The values in the "Baseline Mean" column are the predicted mean outcomes for CIS schools in the school year immediately before CIS was implemented (the intercepts of the baseline trends).

[b]State test scores were converted to z-scores based on the estimated student-level mean and standard deviation in test scores for the state.

References

Balfanz, Robert, Liza Herzog, and Douglas J. Mac Iver. 2007. "Preventing Student Disengagement and Keeping Students on the Graduation Path in Urban Middle-Grades Schools: Early Identification and Effective Interventions." *Educational Psychologist* 42, 4: 223-235.

Ballart, Xavier, and Clara Riba. 1995. "Impact of Legislation Requiring Moped and Motorbike Riders to Wear Helmets." *Evaluation and Program Planning* 18, 4: 311-320.

Bloom, Howard S. 1999. "Estimating Program Impacts on Student Achievement Using 'Short' Interrupted Time Series." MDRC Working Paper on Research Methodology. New York: MDRC.

Bloom, Howard S. 2003. "Using 'Short' Interrupted Time-Series Analysis to Measure the Impacts of Whole-School Reforms: With Applications to a Study of Accelerated Schools." *Evaluation Review* 27, 3: 3-49.

Bloom, Howard S., and James A. Riccio. 2005. "Using Place-Based Random Assignment and Comparative Interrupted Time-Series Analysis to Evaluate the Jobs-Plus Employment Program for Public Housing Residents." *Annals of the American Academy of Political and Social Science* 599, 1: 19-51.

Bloom, Howard S., Lashawn Richburg-Hayes, and Alison Rebeck Black. 2007. "Using Covariates to Improve Precision for Studies That Randomize Schools to Evaluate Educational Interventions." *Educational Evaluation and Policy Analysis* 29, 1: 30-59.

Borman, Geoffrey D., Gina M. Hewes, Laura T. Overman, and Shelly Brown. 2003. "Comprehensive School Reform and Achievement: A Meta-Analysis." *Review of Educational Research* 73, 2: 125-230.

Campbell, Donald T., and H. Laurence Ross. 1968. "The Connecticut Crackdown on Speeding: Time-Series Data in Quasi-Experimental Analysis." *Law and Society Review* 3, 1: 33-53.

Child Trends. 2015. "High School Dropout Rates: Indicators on Children and Youth." Website: www.childtrends.org/?indicators=high-school-dropout-rates.

Communities In Schools. 2011. *2011 Annual Report*. Arlington, VA: Communities In Schools.

Communities In Schools. 2013. *Changing the Picture: Communities In Schools Annual Report*. Arlington, VA: Communities In Schools.

Communities In Schools. 2014. *2014 Annual Report*. Arlington, VA: Communities In Schools.

Communities In Schools. 2015. *2015 Annual Report*. Arlington, VA: Communities In Schools.

Cook, Thomas D., Donald Thomas Campbell, and Arles Day. 1979. *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton Mifflin.

Cook, Thomas D., William R. Shadish, and Vivian C. Wong. 2008. "Three Conditions Under Which Experiments and Observational Studies Produce Comparable Causal Estimates: New Findings from Within-Study Comparisons." *Journal of Policy Analysis and Management* 27, 4: 724-750.

Corrin, William, and Thomas D. Cook. 1998. "Design Elements of Quasi-Experiments." In Arthur J. Reynolds and Harold J. Welber, ed., *Advances in Educational Productivity*, Vol. 7. Greenwich, CT: JAI Press.

Corrin, William, Leigh Parise, Oscar Cerna, Zeest Haider, and Marie-Andrée Somers. 2015. *Case Management for Students at Risk of Dropping Out: Implementation and Interim Impact Findings from the Communities In Schools Evaluation*. New York: MDRC.

Dee, Thomas S., and Brian Jacob. 2011. "The Impact of No Child Left Behind on Student Achievement." *Journal of Policy Analysis and Management* 30, 3: 418-446.

Figlio, David N. 2015. *Experimental Evidence of the Effects of the Communities in Schools of Chicago Partnership Program on Student Achievement*. Northwestern University Working Paper. Evanston, IL: Northwestern University.

Fretheim, Atle, Stephen B Soumerai, Fang Zhang, Andrew D. Oxman, and Dennis Ross-Degnan. 2013. "Interrupted Time-Series Analysis Yielded an Effect Estimate Concordant with the Cluster-Randomized Controlled Trial Result." *Journal of Clinical Epidemiology* 66, 8: 883-887.

Hallberg, Kelly, Vivian C. Wong, and Thomas D. Cook. 2016. "Evaluating Methods for Selecting School-Level Comparisons in Quasi-Experimental Designs: Results from a Within-Study Comparison." Unpublished paper.

Hill, Carolyn J., Howard S. Bloom, Alison Rebeck Black, and Mark W. Lipsey. 2008. "Empirical Benchmarks for Interpreting Effect Sizes in Research." *Child Development Perspectives* 2, 3: 172-177.

Ho, Daniel E., Kosuke Imai, Gary King, and Elizabeth A. Stuart. 2007. "Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference." *Political Analysis* 15, 3: 199-236.

ICF International. 2008. *Communities in Schools National Evaluation Volume 1: School-Level Report. Results from the Quasi-Experimental Study, Natural Variation Study, and Typology Study*. Fairfax, VA: ICF International.

Jacob, Robin T., Roger D. Goddard, and Eun Sook Kim. 2014. "Assessing the Use of Aggregate Data in the Evaluation of School-Based Interventions: Implications for Evaluation Research and State Policy Regarding Public-Use Data." *Educational Evaluation and Policy Analysis* 36, 1: 44-66.

Kemple, James J., Corinne M. Herlihy, and Thomas J. Smith. 2005. *Making Progress Toward Graduation: Evidence from the Talent Development High School Model*. New York: MDRC.

Levin, Henry, Clive Belfield, Peter Muennig, and Cecilia Rouse. 2007. *The Costs and Benefits of an Excellent Education for All of America's Children*. New York: Teachers College Press.

Moore, Kristin A., Selma Caal, Rachel Carney, Laura Lippman, Weilin Li, Katherine Muenks, David Murphey, Dan Princiotta, Alysha Ramirez, Angela Rojas, Renee Ryberg, Hannah Schmitz, Brandon Stratford, and Mary Terzian. 2014. *Making the Grade: Assessing Evidence for Integrated Student Supports*. Bethesda, MD: Child Trends.

Mulford, Harold A, Johannes Ledolter, and Jerry L. Fitzgerald. 1992. "Alcohol Availability and Consumption: Iowa Sales Data Revisited." *Journal of Studies on Alcohol and Drugs* 53, 5: 487-494.

Murnane, Richard J. 2013. "U.S. High School Graduation Rates: Patterns and Explanations." *Journal of Economic Literature* 51, 2: 370-422.

Office of Civil Rights. 2016. *2013-2014 Civil Rights Data Collection, a First Look: Key Data Highlights on Equity and Opportunity Gaps in our Nation's Public Schools.* Washington, DC: U.S. Department of Education, Office of Civil Rights.

Porowski, Allan, and Aikaterini Passa. 2011. "The Effect of Communities In Schools on High School Dropout and Graduation Rates: Results From a Multiyear, School-Level Quasi-Experimental Study." *Journal of Education for Students Placed at Risk* 16, 1: 24-37.

Rosenbaum, Paul R., and Donald R. Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects." *Biometrika* 70: 41-55.

Rubin, Donald R. 2001. "Using Propensity Scores to Help Design Observational Studies: Application to the Tobacco Litigation." *Health Services and Outcomes Research Methodology* 2, 3-4: 169-188.

Schneeweiss, Sebastian, Malcolm Maclure, Bruce Carleton, Robert J. Glynn, and Jerry Avorn. 2004. "Clinical and Economic Consequences of a Reimbursement Restriction of Nebulised Respiratory Therapy in Adults: Direct Comparison of Randomised and Observational Evaluations." *BMJ* 328, 7,439: 560.

Schott Foundation for Public Education. 2012. *The Urgency of Now: The Schott 50 State Report on Public Education and Black Males*. Cambridge, MA: Schott Foundation for Public Education.

Shadish, William R., Thomas D. Cook, and Donald T. Campbell. 2002. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton Mifflin.

Simonsen, Brandi, George Sugai, and Madeline Negron. 2006. "Schoolwide Positive Behavioral Supports: Primary Systems and Practices." *Teaching Exceptional Children* 40, 6: 32-40.

Somers, Marie-Andrée, Pei Zhu, Robin Tepper Jacob, and Howard S. Bloom. 2013. "The Validity and Precision of the Comparative Interrupted Time Series Design and the Difference-in-Difference Design in Educational Evaluation." MDRC Working Paper. New York: MDRC.

St. Clair, Travis, Thomas D. Cook, and Kelly Hallberg. 2014. "Examining the Internal Validity and Statistical Precision of the Comparative Interrupted Time Series Design by Comparison With a Randomized Experiment." *American Journal of Evaluation* 35, 3: 311-327.

Stetser, Marie C., and Robert Stillwell. 2014. *Public High School Four-Year On-Time Graduation Rates and Event Dropout Rates: School Years 2010-11 and 2011-12*. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Swanson, Christopher B. 2010. "Progress Postponed." *Education Week* 29, 34: 22-23.

What Works Clearinghouse. 2014a. *What Works Clearinghouse Procedures and Standards Handbook, Version 3.0.* Washington, DC: U.S. Department of Education, Institute of Education Sciences.

What Works Clearinghouse. 2014b. *WWC Evidence Review Protocol for Dropout Prevention Interventions*. Washington, DC: U.S. Department of Education, Institute of Education Sciences.

Wong, Manyee, Thomas D. Cook, and Peter M. Steiner. 2009. "No Child Left Behind: An Interim Evaluation of Its Effects on Learning Using Two Interrupted Time Series Each With Its Own Non-Equivalent Comparison Series." Northwestern University Institute for Policy Research Working Paper Series WP-09-11. Evanston, IL: Northwestern University.

Earlier MDRC Publications on Communities In Schools

Case Management for Students at Risk of Dropping Out
Implementation and Interim Impact Findings from the Communities In Schools Evaluation
2015. William Corrin, Leigh Parise, Oscar Cerna, Zeest Haider, and Marie-Andrée Somers.

About MDRC

MDRC is a nonprofit, nonpartisan social and education policy research organization dedicated to learning what works to improve the well-being of low-income people. Through its research and the active communication of its findings, MDRC seeks to enhance the effectiveness of social and education policies and programs.

Founded in 1974 and located in New York City and Oakland, California, MDRC is best known for mounting rigorous, large-scale, real-world tests of new and existing policies and programs. Its projects are a mix of demonstrations (field tests of promising new program approaches) and evaluations of ongoing government and community initiatives. MDRC's staff bring an unusual combination of research and organizational experience to their work, providing expertise on the latest in qualitative and quantitative methods and on program design, development, implementation, and management. MDRC seeks to learn not just whether a program is effective but also how and why the program's effects occur. In addition, it tries to place each project's findings in the broader context of related research — in order to build knowledge about what works across the social and education policy fields. MDRC's findings, lessons, and best practices are proactively shared with a broad audience in the policy and practitioner community as well as with the general public and the media.

Over the years, MDRC has brought its unique approach to an ever-growing range of policy areas and target populations. Once known primarily for evaluations of state welfare-to-work programs, today MDRC is also studying public school reforms, employment programs for ex-offenders and people with disabilities, and programs to help low-income students succeed in college. MDRC's projects are organized into five areas:

- Promoting Family Well-Being and Children's Development
- Improving Public Education
- Raising Academic Achievement and Persistence in College
- Supporting Low-Wage Workers and Communities
- Overcoming Barriers to Employment

Working in almost every state, all of the nation's largest cities, and Canada and the United Kingdom, MDRC conducts its projects in partnership with national, state, and local governments, public school systems, community organizations, and numerous private philanthropies.

mdrc USING INTEGRATED STUDENT SUPPORTS TO KEEP KIDS IN SCHOOL A QUASI-EXPERIMENTAL EVALUATION OF COMMUNITIES IN SCHOOLS APRIL 2017

mdrc
BUILDING KNOWLEDGE
TO IMPROVE SOCIAL POLICY

New York
16 East 34th Street
New York, NY 10016
Tel: 212 532 3200

California
475 14th Street
Oakland, CA 94612
Tel: 510 663 6372

www.mdrc.org