4-26-2022

Evaluation of Communities in Schools of Kalamazoo

Randall W. Eberts
W.E. Upjohn Institute for Employment Research, eberts@upjohn.org

Zachary Brown
W.E. Upjohn Institute for Employment Research, brown@upjohn.org

Upjohn Author(s) ORCID Identifier:
https://orcid.org/0000-0002-9711-5466

Citation
Eberts, Randall W. and Zachary Brown. 2022. "Evaluation of Communities in Schools of Kalamazoo." Upjohn Institute Technical Report No. 22-42. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research. https://doi.org/10.17848/tr22-042

W.E. UPJOHN INSTITUTE FOR EMPLOYMENT RESEARCH

Evaluation of Communities in Schools of Kalamazoo

Upjohn Institute Technical Report No. 22-042

April 2022

Randall W. Eberts and Zachary Brown
W.E. Upjohn Institute for Employment Research

W.E. Upjohn Institute for Employment Research
300 South Westnedge Avenue
Kalamazoo, Michigan 49007-4686 U.S.A.

communications@upjohn.org
269.343.5541

Upjohn.org

TABLE OF CONTENTS

ABSTRACT

This report evaluates the impact of intensive student support services provided by Communities in Schools of Kalamazoo (CISK) to elementary students of Kalamazoo Public Schools (KPS). These intensive services typically follow a case-management format and include activities related to academic assistance, basic needs, enrichment/motivation, and life/social skills. Our evaluation examines the following student outcomes: the value-added of NWEA tests for reading and math, attendance rates, and the number of days of unexcused absences. The first two outcomes, reading and math scores, measure student achievement, and the latter two outcomes measure student engagement and student (and parent) behavior. Although KPS considers elementary grades to include kindergarten, we consider only grades one through five in this evaluation, primarily because of the lack of testing in kindergarten. We use both difference-in-differences and a panel event-study methodology, but we prefer the panel event-study approach because it incorporates the dynamics of students receiving CISK services. Based on this approach, we find attendance rates to be the only student outcome significantly affected by CISK services. To understand how these impacts affect outcomes later in their educational careers, we relate elementary-school attendance rates to the number of high school dropouts and graduates for four cohorts of fifth graders. We find that elementary attendance rates show a strong and statistically significant negative association with high school dropouts and a statistically significant and positive association with high school graduates. We conclude, based on our findings, that CISK's intensive services can boost elementary attendance rates, which in turn can reduce the number of high school dropouts and increase the number of high school graduates. Achieving all three outcomes helps to reach the overarching goals of CISK.

I. INTRODUCTION

Local affiliates of Communities in Schools (CIS) work to connect students with integrated support services that help them stay on the path to graduation. The Kalamazoo, Michigan, affiliate of CIS, known as Communities in Schools of Kalamazoo (CISK), focuses on the needs of students attending Kalamazoo Public Schools (KPS). This affiliate provides three tiers of assistance to KPS schools. Tier I includes services that are broadly available to all students at a KPS school that participates with CISK. Tiers II and III provide more intensive services, typically following an individual case-management format, to those students that are determined by CISK to be most in need of attention. At CISK, supportive service categories include not only academic assistance but also basic needs and resources.

This report assesses the impact on student outcomes of Tier II or Tier III services provided by CISK. It does not assess the impact of Tier I services, which are offered to any student attending that school. Three categories of student outcomes will be included in the analysis: 1) education outcomes, as measured by standardized test scores in reading and math; 2) an engagement outcome, as measured by student attendance; and 3) a behavioral outcome, proxied by days of unexcused absences. The evaluation uses a newly constructed longitudinal data set, the Community Data System (CDS), which follows KPS students from kindergarten through graduation by cohort.[1] The data system, along with detailed information about each student participating in CISK, allows for the construction of comparison groups that match as closely as possible the demographics of those students in the treatment group. A propensity score

[1] The Community Data System also includes birth records and postsecondary affiliation. At one time it included employment data from UI wage records. Only KPS variables and CISK variables are used in this evaluation, though.

matching methodology is used to construct the comparison group, and an event study approach is used to estimate the dynamic impacts of those services on student outcomes.

Panel Event-Study Approach

More specifically, this event study approach, in addition to the more traditional difference-in-differences (DID) approach, is used to estimate the impact of CISK services on first through fifth graders in KPS. While the primary purpose of this evaluation is to pinpoint the effects of CISK services on elementary students, we also examine the effect of CISK services on selected outcomes of KPS middle schoolers, the results of which are included in the appendix. An event study approach accommodates the variation in years in which elementary students, for example, first encounter CISK services. As shown in Clarke and Schythe (2020), it is an extension of a difference-in-differences model. Once a student receives services, we hypothesize that the effects of those services can linger throughout the rest of a student's elementary grades, even if the student does not receive any services after that first encounter. Of course, it may be the case that students, once they receive CISK services for the first time, go on to receive them the next year and even in subsequent years. The event study approach allows for that possibility as well. Another attractive feature of the approach is that students who do not receive CISK services at any time during their elementary school years are considered to be members of the comparison group, along with those who received services in subsequent years. For example, a student who received services in 2016, regardless of his or her grade level at the time, will be in the comparison group in 2015, along with those who did not receive any services throughout elementary school.

The CIS Model

Bill Milliken founded Communities in Schools in 1977. According to the national CIS website, 1.61 million students in 2,900 schools received services from 133 organizations and licensees in the United States during the last school year (Communities in Schools 2022). Its primary purpose is to reduce high school dropout rates by integrating preventive services available to the entire school with intensive, targeted, and sustained services for those students who display significant risk of dropping out.[2] The CIS model uses site coordinators and other staff located directly in schools to assess students' needs and then target and coordinate the distribution of services and goods to meet those individual needs. These services and goods may come from other community organizations, as well as from CIS staff. The CIS model is different for elementary schools, middle schools, and high schools. In elementary schools, the model focuses on improving attendance rates by reaching out to and engaging parents. In middle schools, CIS emphasizes student behavior. In high schools, the model provides services intended to prevent students from dropping out and to help them progress toward graduation.

The CIS model provides three tiers of services to students identified as demonstrating moderate to high risk of dropping out of high school. The first tier is the least intensive. It includes preventive services that are available to all students, including those with the highest risk. Those services are considered "whole school" services, since they are available to everyone attending that school. The second and third tiers include activities that are more intensive and administered to a large extent through case management. These services are exclusionary and

[2] This information has been gleaned from the Communities in Schools national website under the headings "Mission" and "History." Warren (2005) also discusses the merits of the communities-in-schools model, although he calls it "the service approach," and describes how it is positioned to help reform urban school policy and practice. Child Trends (2014) assesses the evidence for integrated student supports, similar to the Tier II or Tier III services CIS makes available to students participating in CIS.

available to students who exhibit moderate- to high-risk behavior in engagement, academics, and behavior. Most of the CIS activities are directed toward moderate-risk students, not high-risk students. The exact criteria vary across affiliates, but for CISK the primary target group for Tier II or III services includes students with between 10 and 25 absences a year, one or more days of suspension, a core GPA of less than 2.5, and standardized math and reading test scores between the 25th and 65th percentiles. For most schools, both within KPS and in other school districts, Tier II or III services are targeted to about 10 percent of the student population.

National CIS Evaluations

During the past 20 years, the national office of CIS has contracted with two consulting firms to evaluate the CIS model. ICF International conducted the first evaluation, which included two levels of analyses (ICF International 2010). The first level entailed a school-level quasi-experimental study that used propensity score matching to select a comparison group from schools in three districts located across the country. Three categories of student outcomes were aggregated at the school level: 1) dropout and graduation, 2) attendance rate, and 3) academic performance on standardized math and reading assessments. For the treatment group, schools were classified as either high implementers or low implementers based on a rubric that differentiated between the two types of implementers. "High implementers" were considered to be schools that excelled in four domains: 1) planning, 2) needs assessment, 3) service delivery, and 4) monitoring and adjustment. The outcomes were measured at baseline and three years after implementation. Based on this approach, ICF International found that high implementers had considerably greater effects on reducing dropout rates and increasing on-time graduation than schools in the non-CIS comparison group and schools that were categorized as partial implementers. The study reported consistently positive but small improvements in attendance, particularly in high-implementing elementary schools. Similar results were found for math and reading assessments in high-implementing middle schools.

The second level of analysis used an RCT technique to separate individual students into treatment and comparison groups. The student outcomes differed from the first-level analysis. On-time high school graduation was replaced with retention, and a category of behavioral problems and discipline was added, which included the metrics of out-of-school suspensions and disciplinary referrals. The results were similar to those found in the first level analyses—at least one of the metrics, if not more, within a student-outcome category was shown to have a positive effect.

In its ongoing commitment to continuous improvement, CIS asked MDRC, a nationally recognized evaluation and consulting organization, to continue with an evaluation of the activities of some of its affiliates along the same lines as ICF International had done.[3] MDRC used a methodology very similar to that proposed and used by ICF. MDRC conducted its evaluation in two phases. The first phase examined schoolwide or "whole school" effects using a quasi-experimental methodology (Somers and Haider 2017). The treatment group included 53 schools in North Carolina and Texas, comprising 14 high schools, 15 middle schools, and 24 elementary schools. These schools began implementing the CIS model in 2005. A comparison group of 78 schools (18 high schools, 24 middle schools, and 36 elementary schools) was selected from schools in the same counties as the treatment group schools. No schools in the comparison group used the CIS model; however, these schools generally exhibited the same characteristics and demographics as the CIS schools. Using an interrupted time-series model, the whole-school study found that for elementary schools, attendance rates improved in schools

[3] Another factor in selecting MDRC soon after ICF International had completed its evaluation was the participation of CIS in the federal Social Innovation Fund (SIF) grant program.

implementing the CIS model more than they did in a comparison group of similar schools. However, the study found no effect on attendance in middle and high schools, and no effect on test scores in elementary schools. In middle schools, English/language arts test scores did not improve in schools implementing the CIS model, whereas they did improve in the comparison group of middle schools.[4] 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. High school students in the treatment group experienced lower dropout rates and greater graduation rates than students in the comparison group schools, but it was unclear whether CIS actually was responsible for the difference.

The second phase of the MDRC evaluation used a randomized control trial (RCT) to construct a treatment group and a comparison group (Parise et al. 2017). Researchers estimated the impact of Tier II services of the CIS model by comparing student outcomes between the two groups. This approach is the closest to ours, since we used an event study approach to estimate the effect of Tier II or III services on individual students. In addition to the difference in methodology between our study and that of MDRC, another difference occurred between our study and MDRC's second phase, in that we focused on elementary students, whereas MDRC examined middle school and high school students. The MDRC study looked at students in 14 middle schools and 10 high schools but no students in elementary schools. The evaluation was conducted in the 2012–2013 and 2014–2015 school years; the final report was published in 2017.

The students in MDRC's second-phase evaluation were randomly assigned to one of two groups. The treatment group provided case-management services to its students, while the

[4] David Figlio (2015) found that students in CIS schools in K–8 grades (which include middle school) in the Chicago Public Schools (CPS) performed better than those in comparison CPS schools, and that the difference was statistically significant.

comparison group continued with business as usual. This included access to whatever forms of support were normally available, such as CIS Tier I services. Since Tier I services are available to anyone who attends one of the schools in the sample, the evaluation essentially estimates the relative effectiveness of Tier II (case-managed) services versus Tier I services.[5] The phase-two RCT experiment was conducted across two consecutive years. Depending on student responses, treatment-group students in the first year of the experiment received services like those the comparison group did, which were primarily Tier I services. In Year Two, the services became more individualized and intensive. Students with a moderate risk of dropping out were more likely to receive CIS services than were high-risk students. Unfortunately, since the experiment lasted only for two years, it was not possible to track students through twelfth grade to see whether the interventions affected high school graduation. The study, instead, examined nonacademic mediating outcomes and more traditional school outcomes.

The MDRC study found no difference between case-managed and non-case-managed students in their participation in school and non-school-sponsored extracurricular activities, or in their educational goals and expectations. The study did find that case management had positive and statistically significant effects on students' engagement with school, their educational attitudes, and their belief that education had value in their lives. However, it found similar results for students who received services not related to the case-management approach. As for more traditional student outcomes, such as attendance and course performance, the MDRC study concluded that case-management services did not influence students' outcomes. Thus, the

[5] At the time of the MDRC's evaluation, CIS services were categorized as either Tier I or Tier II services; Tier III did not exist, and what became Tier III services were included in Tier II. It was because of this evaluation that MDRC recommended adding a third tier to the other two. CIS adopted that recommendation.

MDRC study did not show that CIS case management improved students' attendance, course performance, or behavior.

Differences between the MDRC Study and This Evaluation

Three major differences emerge between the MDRC study and ours. First, our evaluation considers only the integrated service component of the CIS model. It does not examine the effects on student outcomes of Tier I services, those services that are available to every student at the school that offers them. Second, the MDRC evaluation used random assignment to construct the various treatment and comparison groups; we used quasi-experimental approaches. Random assignment is considered the gold standard of evaluation methods. The advantage of randomization is that with sufficient sample size, it minimizes estimation bias. It makes the treatment and comparison groups comparable with respect to both observed and unobserved factors.

However, random assignment does have disadvantages. One drawback is the expense of collecting the appropriate data from each treatment and control group. Since it is not known, a priori, which student will belong to which group, one must use surveys to collect data on students in the two groups. Administrative data, for example, does not include information about members of a comparison group since, by definition, administrative data includes information about only those receiving services. Furthermore, power calculations could demand larger samples, which in turn require more resources from the investigators. Also, since information is collected only after random selection has taken place, one cannot depend upon possessing data before that time. Therefore, MDRC could examine only two years of student outcomes. Because of this, say Parise et al. (2017, p. ES-6), "it was not possible to track students through high school graduation." Since we use a quasi-experimental design, we employ administrative data from KPS and CISK to track students from first grade through the end of high school. Yet another drawback of random assignment is that it denies individuals in the control group access to the treatment. This can raise ethical issues of whether an individual should be denied treatment that may be beneficial, and of what the sequence of treatment should be if the evaluation consists of more than one round of tests.[6]

The third difference between the two studies is that MDRC examined students in middle school and high school, whereas our primary focus is students in elementary school.[7] According to the CIS national office, services in elementary school are geared toward improving attendance rates and engagement in school, whereas the primary purpose of services in middle school is to address behavioral issues. And in high school, the purpose of the CIS model is to reduce dropout rates and increase graduation rates. While our primary focus for elementary students is on attendance, we also test to determine whether CISK services have a significant impact on academics and behavioral outcomes.

The results from our evaluation are consistent with what was found by the latest MDRC evaluation of affiliates in the states of North Carolina and Texas, with one notable exception: We found that CISK services increased attendance rates of elementary students between the school years 2014–2015 and 2018–2019. The coefficients are statistically significant. MDRC found no statistically significant effect of CIS services on traditional and nontraditional student outcomes. We were then able to show that elementary attendance rates were statistically

[6] Another issue is the length of time an individual must wait for treatment, if indeed the treatment could benefit that individual. Many times, individuals in the comparison group may adopt other treatments than the one under evaluation and confound the results.

[7] The primary focus of our evaluation is on elementary students, but we do include middle schoolers to compare our results with other studies. The results are shown in the appendix,

significantly related to a reduction in the number of high school dropouts and an increase in the number of high school graduates from the same cohort of students.

Outline of this Report

The remainder of this report is divided in six major sections. Section II describes the construction of the two samples used throughout the study. Section III examines in detail the activities of CISK staff as recorded in the databases CISK prepares for its own use and for the national office. We examine the types of services as categorized by tier, by provider, by school building, and by description of the service. We tabulate the services by year, by grade level, and by hours engaged in the service. This analysis provides the insight that is used in subsequent sections.

Section IV describes construction of the comparison and treatment groups used in the study. That section then details the panel event-study methodology and the difference-in-differences (DID) approach and discusses the relationship between the two. Differences in the average value of the covariates and student outcomes are examined to gain a better sense of the differences between the treatment and comparison groups. Section V posts the results of the estimates of the impact of CISK services on student outcomes, using both the DID approach and the panel event-study methodology. Simple statistics related to the panel event study are displayed, and results are presented and discussed.

Section VI shows empirically the relationship between attendance rates of elementary students and the number of high school dropouts and graduates. Section VII discussions the implication of the estimates, particularly with respect to the MDRC evaluation.

II. SAMPLES OF STUDENTS

We use two samples of KPS students in this study. Both samples include elementary students, which in KPS are kindergartners through fifth graders, but for this study we include only first through fifth graders. We do not include kindergartners for two reasons: First, kindergartners do not take state standardized tests that can be compared with tests taken in higher grades, and second, CISK serves only 3.6 percent of kindergartners. Without kindergartners included, CISK serves 9.0 percent, or 2,610 students, of the nearly 29,000 first through fifth graders attending KPS elementary schools from 2015 through 2019.

The first sample (referred to as Sample A) comprises students from the same cohort starting in first grade and ending with fifth grade. To belong to this cohort, students must attend KPS in each of the school years from 2014–2015 (first grade) through 2018–2019 (fifth grade). The first sample includes only first graders the first year, second graders the second year, third graders the third year, and so forth up through fifth grade. The second sample (Sample B) includes all elementary students, except kindergartners, in each of the five school years. Because Sample B includes all elementary students each school year, it does not require that students must attend KPS each year of elementary school and thus is a much larger sample than Sample A.

As with any school district, KPS has attrition among students. As shown in Table 1, the first-grade cohort in KPS starts with 1,212 students in 2014–2015, and by the next year the district had lost 154 students to various types of attrition, leaving 1,058 second graders who had started in KPS the year before. By fifth grade, only 805 of the initial 1,212 members of the first-grade class had attended KPS each year and progressed adequately so that they were not held back to repeat a grade.

Table 1 Counts of Students in Various Grades by Year, Sample A

2015		2016		2017		2018		2019	
Grade	Count	Grades	Count	Grades	Count	Grades	Count	Grades	Count
1	1,212	1–2	1058	1–3	948	1–4	869	1–5	805
2	1,071	2–3	1036	2–4	933	2–5	853		
3	1,125	3–4	1010	3–5	911				
4	1,095	4–5	977						
5	1,059								

SOURCE: Authors' analysis of the Community Data System.

When the file is reshaped into a longitudinal file, the number of students times the number of interventions becomes the number of student-years. That is, each of the 805 students in Sample A is seen for five years, once in each grade. Of course, because of the construction of Sample A, only a student in the grade captured for the specific school year is in the sample; the other grades have zero students. Therefore, a cross tabulation of year by grade yields Table 2.

Table 2 Number of Students in Sample A by Year and Grade

	Grade					
School year	1st	2nd	3rd	4th	5th	Total
2014–2015	805	0	0	0	0	805
2015–2016	0	805	0	0	0	805
2016–2017	0	0	805	0	0	805
2017–2018	0	0	0	805	0	805
2018–2019	0	0	0	0	805	805
Total	805	805	805	805	805	4,025

SOURCE: Authors' analysis of Community Data System.

Sample B, as shown in Table 3, includes all first through fifth graders who attended KPS each year, but unlike Sample A, Sample B does not require those students to be enrolled in KPS for their entire five years in elementary school. Students can come into the district at any time during their five years in elementary school and leave at any time. Beyond the cohort of students that attend KPS from first through fifth grade, Sample B also includes all five grades for each of the five years. Therefore, Sample A is a subset of Sample B.

Table 3 Number of Students in Sample B by School Year and Grade

	Grade					
School year	1st	2nd	3rd	4th	5th	Total
2014–2015	1,140	1,180	1,053	1,127	1,059	5,559
2015–2016	1,104	1,105	1,153	1,020	1,133	5,515
2016–2017	1,134	1,096	1,066	1,148	1,014	5,458
2017–2018	1,113	1,095	1,080	1,021	1,115	5,424
2018–2019	1,135	1,087	1,065	1,077	1,009	5,373
Total	5,626	5,563	5,417	5,393	5,330	27,329

SOURCE: Authors' analysis of Community Data System.

To be more specific, Sample A includes 805 students in each grade per year, totaling 4,025. Sample B, which expands to all five elementary grades each year, includes a total of 27,329 student-years. Sample B, because it not only includes all five elementary grades but also includes all elementary students in each grade, regardless of whether they attended KPS during their entire elementary experience, is nearly seven times the size of Sample A. Sample B enables more precision in the estimates not just because of a greater number of students, but also because it includes the grades the students are in each year.

III. CISK ACTIVITIES

Delivery of CISK Services

CISK records detailed information about the services received by students whom it determines eligible for each of the three tiers of assistance.[8] CISK exclusively serves students in the Kalamazoo Public Schools, which between the school years of 2014–2015 and 2018–2019 enrolled an average of just over 14,000 K–12 students a year. CISK serves roughly 5.7 percent of those students. The largest percentage of KPS students it serves are in the kindergarten through

[8] Throughout this report, we will use a shorthand for the activities that CISK engages in with students. Typically, we would refer to the services and supports that CISK offers, but for brevity's sake we will combine the terms "services" and "supports" into the single word "services."

fifth grade elementary schools (8.1 percent), followed by students in the sixth through eighth grade middle schools (4.4 percent), and then those in the ninth through twelfth grade high schools (2.7 percent).

Table 4 shows the number of students receiving CISK services each year and the total number of students in each grade (i.e. the "grade count") in Sample A. In this sample, CISK served a total of 595 students, or 10.5 percent of enrollment, including 86 first graders in school year 2014–2015, 138 second graders in 2015–2016, 191 third graders in 2016–2017, 76 fourth graders in 2017–2018, and 104 fifth graders in 2018–2019. Similarly, Table 5 displays the number of CISK students and total grade count in Sample B. Even though the number of students in Sample B is much larger than in Sample A, the percentage of CISK participants is lower by 1 percentage point.

Table 4 Counts of Students and CISK Participants by Elementary Grades in Sample A

Grade	2015	2016	2017	2018	2019	Total students in grade	% CISK students
First	86					1,212	7.1
Second		138				1,171	11.8
Third			191			1,125	17.0
Fourth				76		1,095	6.9
Fifth					104	1,059	9.8
Total						5,662	10.5

SOURCE: Authors' analysis of the CISK database.

Table 5 Counts of Students and CISK Participants by Elementary Grades in Sample B

Grade	2015	2016	2017	2018	2019	Total CISK students	Total students	% CISK students
First	86	102	122	46	87	443	5,626	7.9
Second	98	138	166	72	98	572	5,563	10.3
Third	98	140	191	58	126	613	5,417	11.3
Fourth	43	125	203	76	100	547	5,393	10.1
Fifth	24	54	180	72	104	434	5,330	8.1
Total	349	559	862	324	515	2,609	27,329	9.5

SOURCE: Authors' analysis of the CISK database.

Service Providers

Four categories of service providers are included in the CISK database: 1) CISK staff, 2) community partners, 3) school staff, and 4) volunteers. Table 6 displays, by grade, the number of students served by the four providers in Sample A. Of the four categories of service providers, CISK staff serves the most students, and most of that effort is concentrated in the elementary grades, particularly the middle elementary grades. After CISK staff, community partners and volunteers typically rank second and third in the portion of students they impact.

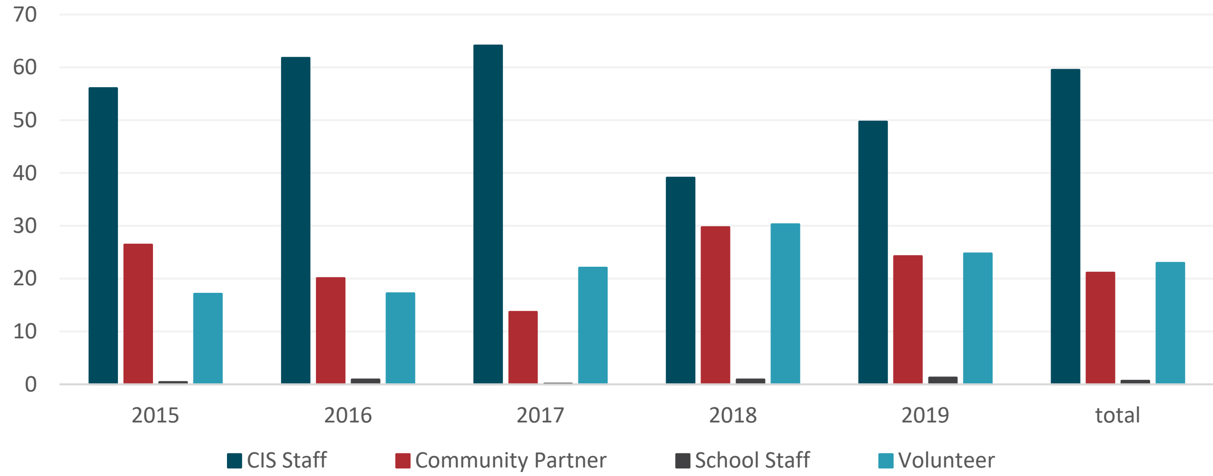
Table 6 Count of Services by Year and Grade in Sample A

	2015, 1st grade		2016, 2nd grade		2017, 3rd grade	
Service providers	*N*	%	*N*	%	*N*	%
CIS staff	509	52.26	1,272	57.07	2,065	61.81
Community partner	290	29.77	497	22.30	486	14.55
School staff			11	0.49	1	0.03
Volunteer	175	17.97	449	20.14	789	23.62
Total	974	100.00	2,229	100.00	3,341	100.00
	2018, 4th grade		**2019, 5th grade**		**Total**	
	N	%	*N*	%	*N*	%
CISK staff	90	16.57	980	55.37	3,936	55.54
Community partner	74	13.63	328	18.53	1,347	19.01
School staff	12	2.21	16	0.90	24	0.34
Volunteer	367	67.59	446	25.20	1,780	25.12
Total	543	100.00	1,770	100.00	7,087	100.00

SOURCE: Authors' analysis of CISK database. Totals may not sum to 100.00 because of rounding.

Of the four categories of service providers, displayed in Figure 1, the majority of support services for KPS students comes from CISK staff. Support services can last for as long as 360 hours a year or for as little as 15 minutes. It can also be listed multiple times for a specific student. Students received as many as 140 separate service elements a year (school year 2018–2019) and as few as one (each school year). The median number of activities ranges from 16 to 25 a year. CISK staff provided nearly 60 percent of the 58,000 activities for all five years combined and, except for school year 2017–2018, at least 50 percent of the activities each year. One can also see from the figure that school staff provide relatively few services, with a combined total of 0.67 percent, or 394 services. Community partners and volunteers provided nearly equal percentages of the remaining 26,000 activities.

Figure 1 Percentage of CISK Services Offered by Provider Type, School Years 2015–2019



SOURCE: Authors' analysis of CISK database.

Table 7 breaks out the service providers by year for Sample B, similar to what was done using Sample A in Table 6. The sample proportions of responsibilities are shown across types of providers. CISK staff handle the most responsibility, in terms of number of services performed, while KPS school staff have the fewest number of duties.

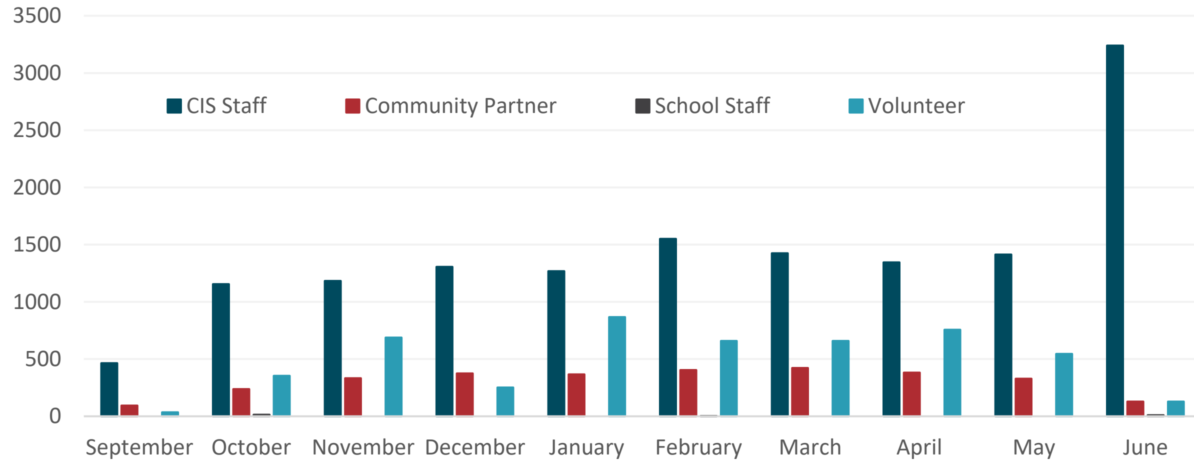
Table 7 Count of CISK Services by Year in Sample B

	2015	2016	2017	2018	2019	Total	% Total
CISK staff	3,179	7,880	14,363	2,380	6,947	34,749	57.1
Community partner	1,498	2,563	3,069	1,809	3,388	12,327	20.2
School staff	23	113	26	54	178	394	0.7
Volunteer	971	2,196	4,944	1,942	3,458	13,411	22.0

SOURCE: Authors' analysis of CISK database.

Figure 2 displays the number of services by month and provider type during the school year. Even though the busiest time for CISK staff is June, the end of the school year, CISK staff perform more services each month than any other type of provider. The June spike results from CISK staff closing their files for students at the end of the school year.

Figure 2 Number of Services by Provider Type and Month, for School Year 2016–2017



SOURCE: Authors' analysis of CISK database.

Classification of Student Support Categories

Tables 8 and 9 show the provision of CISK services by provider type. For both samples, CISK staff provide much of the Tier II or Tier III services, and community partners offer most of the Tier I services. It is evident that the classification of certain services has changed. This follows from the earlier discussion of a change in classification because of the MDRC study. In more recent years, after MDRC's recommendation, services have been distributed more evenly across tier classifications.

Service Support Categories and Number of Hours

CISK categorizes services into at least 10 student support groups. Examining the categories across the years, we find that before 2015–2016, most categories were classified as Tier III activities, whereas in more recent years, the same categories were classified more evenly as Tier I and Tier II as well as Tier III activities. For example, as shown in Table 10, academic assistance during and before school year 2014–2015 was classified only as a Tier III service,

Table 8 Count of Services by Provider Type, Tier, Year, and Grade Cohort, Sample A

	2015			2016			2017		
Provider type	Tier I	Tier II	Tier III	Tier I	Tier II	Tier III	Tier I	Tier II	Tier III
CIS staff	0	0	509	94	379	799	91	649	1,325
Community partner	0	0	290	205	134	158	132	155	199
School staff	0	0	0	10	0	1	0	1	0
Volunteer	0	0	175	13	114	322	4	497	288
Tier share of total	0.0	0.0	1.00	0.144	0.281	0.574	0.068	0.390	0.542

	2018			2019			Total		
Provider type	Tier I	Tier II	Tier III	Tier I	Tier II	Tier III	Tier I	Tier II	Tier III
CIS staff	2	60	28	141	704	135	328	1,792	2,796
Community partner	12	1	61	116	26	186	465	316	894
School staff	12	0	0	16	0	0	38	1	1
Volunteer	153	52	162	42	188	216	212	851	1,163
Tier share of total	0.330	0.208	0.462	0.178	0.519	0.303	0.118	0.334	0.548

SOURCE: Authors' analysis of CISK database.

Table 9 Count of Services by Provider Type, Tier, Year, and Grade Cohort, Sample B

	2015			2016			2017		
Provider Type	Tier I	Tier II	Tier III	Tier I	Tier II	Tier III	Tier I	Tier II	Tier III
CIS staff	3	0	3,176	501	2,917	4,462	397	4,600	9,366
Community partner	0	2	1,496	878	950	735	587	1,329	1,153
School staff	0	0	23	47	39	27	0	24	2
Volunteer	0	0	971	202	661	1,333	48	2,993	1,903
Tier share of total	0.001	0.000	0.999	0.128	0.358	0.514	0.046	0.399	0.555

	2018			2019			Total		
Provider Type	Tier I	Tier II	Tier III	Tier I	Tier II	Tier III	Tier I	Tier II	Tier III
CIS staff	227	563	1,590	926	3,558	2,463	2,054	11,638	21,057
Community partner	794	56	959	1,218	259	1,911	3,477	2,596	6,254
School staff	51	1	2	59	89	30	157	153	84
Volunteer	525	475	842	729	1,019	1,710	1,504	5,148	6,759
Tier share of total	0.262	0.180	0.558	0.210	0.353	0.438	0.118	0.321	0.561

SOURCE: Authors' analysis of CISK database.

Table 10 Comparison of Services by Categories across Several Years

	2012–2013	2014–2015				2015–2016			
	Tier	Tier				Tier			
Service support categories	III	I	II	III	Total	I	II	III	Total
Academic assistance	741	0	0	1,347	1,347	116	1,187	1,510	2,813
Basic needs/resources	1,013	3	2	1,736	1,741	836	1,536	484	2,856
Behavior intervention	13	0	0	267	267	142	94	53	289
Case management		0	0	451	451	0	248	3,521	3,769
College/career preparation	18	0	0	137	137	21	160	58	239
Community service	3	0	0	12	12	16	3	1	20
Enrichment/motivation	869	0	0	539	539	90	561	215	866
Family engagement	28	0	0	140	140	43	4	27	74
Life/social skills	562	0	0	628	628	145	611	306	1,062
Physical fitness/health	102	0	0	174	174	166	57	8	231
Professional mental health	49	0	0	224	224	7	106	374	487
Site coordination		0	0	11	11	46	0	0	46

SOURCE: Authors' analysis of CISK database.

as were all the other activities. However, in the next year, school year 2015–2016, 42 percent of academic assistance services were classified as Tier II services and 54 percent of those services were classified as Tier III services, with the remaining 4 percent classified as Tier I services.

CISK also records the number of hours that each of the four types of service providers spends on various activities with each student. Table 11 displays the total number of hours recorded for each activity for the grade and school year included in the study. In total hours, the service with the most time spent on it is academic assistance, but the category with the largest

Table 11 Number of Hours and Count of Students (*N*) by Grade, School Year, and Support Category in Sample B

	2015			2016			2017		
Service support categories	Total hours	Mean	*N*	Total hours	Mean	*N*	Total hours	Mean	*N*
Academic assistance	3,407	14.81	230	10,439	19.81	527	13,330	16.81	793
Basic needs/resources	1,985	5.70	348	4,056	7.85	517	4,542	10.59	429
Behavior intervention	109	3.12	35	108	2.44	44	107	3.81	28
Case management	77	1.30	59	258	0.39	660	861	0.68	1,268
College/career preparation	8	2.67	3	238	9.91	24	54	4.50	12
Enrichment/motivation	5,392	61.27	88	7,345	62.24	118	9,243	36.11	256
Family engagement	36	1.63	22	38	2.50	15	10	1.90	5
Life/social skills	949	9.68	98	1,354	7.61	178	4,259	10.00	426
Physical fitness	17	0.64	27	469	12.34	38	547	11.63	47
Professional mental health	115	1.79	64	261	2.67	98	298	5.13	58
Total	12,095	12.42	974	24,676	11.07	2,229	33,274	9.96	3,341

	2018			2019			Total		
	Total hours	Mean	*N*	Total hours	Mean	*N*	Total hours	Mean	*N*
Academic assistance	1,968	4.54	434	4,876	4.17	1,170	34,021	10.79	3,154
Basic needs/resources							10,583	8.18	1,294
Behavior intervention	33	4.71	7	68	0.82	83	425	2.16	197
Case management	2	0.28	8	77	1.17	66	1,275	0.62	2,061
College/career preparation	240	20.00	12	9	1.50	6	549	9.63	57
Enrichment/motivation	237	14.79	16	693	3.61	192	22,910	34.19	670
Family engagement	8	4.00	2	153	7.29	21	244	3.75	65
Life/social skills	79	2.94	27	2,325	17.35	134	8,967	10.39	863
Physical fitness	326	16.29	20	614	10.06	61	1,973	10.22	193
Professional mental health	40	2.34	17	92	1.46	63	806	2.69	300
Total	2,933	5.40	543	8,929	4.96	1,802	81,906	9.21	8,889

SOURCE: Authors' analysis of CISK database.

average time spent on it is enrichment/motivation—more than three times the average hours spent per activity of academic assistance, the category with the next highest mean. In terms of total hours, the category "Basic needs/resources" ranks third, but the time commitment to that service does not reveal the true value of that service, since it may take less time to get resources to those in need of them than to tutor an academic lesson.

Schools Attended by Students Receiving Tier II or Tier III Services

Tables 12 and 13 show the number of students receiving Tier II or Tier III services by elementary school in KPS. Table 12 displays the cohort of students passing through KPS without leaving KPS, whereas Table 13 exhibits all elementary school students in KPS. The total in Table 13 includes all first through fifth graders in each of the five years displayed in the table. During the five-year period, 8.7 percent of the 24,841 elementary students received Tier II or Tier III services. For the cohort of elementary students in the first sample, 12.4 percent of 3,453 students received these intensive services.

Table 12 Number of Students Receiving Tier II or Tier III Services by School and Year in Sample A

	School year									
	2015		2016		2017		2018		2019	
School	Treated	Total	Treated	Total	Treated	Total	Treated	Total	Treated	Total
Arcadia	6	47	4	51	8	50	11	49	7	49
Prairie Ridge	4	65	4	61	4	57	0	54	3	53
Edison	10	34	15	36	17	39	0	42	2	44
Lincoln	4	35	7	31	13	29	0	33	0	36
Milwood	7	52	6	45	11	48	0	63	4	64
Northeastern	2	40	5	35	7	30	0	25	5	29
Northglade	8	28	9	28	12	28	0	29	0	30
Woods Lake	10	59	26	68	28	70	2	75	10	84
Parkwood-Upjohn	1	66	5	65	5	65	4	70	3	70
Spring Valley	6	35	4	31	5	35	6	34	6	41
Washington	3	48	4	46	16	46	5	49	5	34
King-Westwood	3	62	1	65	0	65	0	105	0	106
Woodward	7	62	11	51	8	49	7	53	9	43
El Sol	5	49	4	49	9	45	9	43	7	44
Total	76	682	106	664	143	656	44	724	61	727

NOTE: Three elementary schools in KPS—Greenwood, Indian Prairie, and Winchell—decided not to participate in the CISK program, so their numbers were not included in the table; nor were summer school totals and an unknown category. Thus, the total number of students in the sample each year does not add to 805, which is the number of students each year in the cohort in sample A.
SOURCE: Authors' computation of KPS and CISK data.

Table 13 Number of Students Receiving Tier II or Tier III Services by School and Year in Sample B

School	School year									
	2015		2016		2017		2018		2019	
	Treated	Total	Treated	Total	Treated	Total	Treated	Total	Treated	Total
Arcadia	25	346	25	399	52	366	48	358	36	377
Prairie Ridge	30	512	44	497	54	429	0	429	19	437
Edison	35	319	55	281	81	279	1	297	14	309
Lincoln	22	261	32	268	65	249	0	257	0	289
Milwood	26	374	31	320	47	335	0	381	35	370
Northeastern	11	305	20	250	34	211	0	209	20	242
Northglade	24	148	30	142	53	172	0	205	11	215
Woods Lake	43	512	79	554	100	544	15	556	49	574
Parkwood-Upjohn	9	472	17	497	24	490	17	488	17	502
Spring Valley	17	312	22	302	52	292	44	301	27	305
Washington	28	439	33	372	93	370	17	332	45	238
King-Westwood	18	499	14	506	1	523	0	531	1	510
Woodward	37	422	49	323	60	311	25	294	45	222
El Sol	24	274	32	278	54	282	36	280	43	291
Total	349	5,205	483	4,989	770	4,853	203	4,918	362	4,881

NOTE: Three elementary schools in KPS—Greenwood, Indian Prairie, and Winchell—decided not to participate in the CISK program and were not included in the table.
SOURCE: Authors' computation of KPS and CISK data.

Table 14 displays the number of elementary students receiving Tier II or Tier III services by neighborhood for Sample B. Since some students have missing values for place of residence and the neighborhoods are only within the city of Kalamazoo, the neighborhood table has fewer students than the school-based tables. For the total sample, 10,866 students (compared with 24,841 students in the school-based table) are included in the table. According to KPS records, the number of students who live in Kalamazoo County but do not live in the city is 6,902. Two other categories are included in the KPS records—living in Michigan and living outside Michigan—which together account for 2,870 elementary students.

Table 14 Number of Students Receiving Tier II or Tier III Services by Neighborhood for the five years within the City of Kalamazoo in Sample B

Neighborhood	Treated	Total
Westnedge Hill	8	348
Milwood South	26	632
Winchell	38	449
Milwood North	28	429
Oakwood	8	186
South Westnedge East	33	636
Colony Farm	0	49
Burke Acres	54	522
South Westnedge West	36	341
Southside	36	449
Edison Central	163	1,555
Stuart	75	742
Vine	34	534
Westwood West and Arcadia	24	511
Edison North	57	312
Northside West	129	1,260
Northside East	50	347
Eastside	71	793
Knollwood	72	573
WMU	1	31
CBD	16	167
Total	959	10,866

NOTE: Three elementary schools in KPS—Greenwood, Indian Prairie, and Winchell—decided not to participate in the CISK program and were not included in the table.
SOURCE: Authors' computation of KPS and CISK data.

IV. EVALUATION METHODOLOGIES

A panel event study is used to evaluate the CISK program for elementary students and middle schoolers in KPS during the aforementioned school years. A panel event study is an extension of the standard two-way fixed effect (sometimes called difference-in-differences) model, in which a single "post event" indicator is included for all periods after the occurrence of the services received by the treated students (Clarke and Schythe 2020).[9] The estimating equation is written as follows:

[9] Difference-in-differences (DID) and more recently panel event-study methodologies have been used frequently in the economics literature to estimate the causal effect of interventions. A few applications of DID include Stevenson and Wolfers (2006), Goodman (2019), Donald and Lang (2007), and Dee et al. (2021). We include only two of the many papers that examine DID within multiple time periods: Callaway and Sant'Anna

$$y_{st} = \alpha + \beta\ postevent_{st} + \mu_s + \lambda_t + X_{st}\acute{\Gamma} + \epsilon_{st} \quad , \tag{1}$$

where $postevent_{st} = 1\ [t \geq Event_s]$ and X_{st} represent covariates added to the equation. Not apparent from the simple model in Equation (1) is that the full set of event leads allow for the inspection of parallel trends in the pretreatment period, and the policy lags allow for the inspection of the temporal nature of treatment effects—noting, for example, any dynamics in the appearance of effects. This methodology exploits the variation in the timing of KPS students' receiving services from CISK to evaluate the impact of CISK services on those students.

Following loosely the paper by Goodman-Bacon (2018), we explore difference-in-differences estimates in instances when the time varies in which groups of students receive treatment. The results are reported in the document prior to the event study description and results. Exploiting the variation in the timing of treatment is the hallmark of this evaluation, as it is for so many DID applications, and we find little statistically significant difference between treatment and comparison group outcomes for different time periods.

Construction of Treatment and Comparison Groups

Treatment groups for both samples include students who participated in CISK Tier II or Tier III activities for at least one of the five years (again, not considering kindergartners). We assume that once a student receives CISK services, the potential effects of the services linger for the rest of his or her time in elementary school. This is a standard assumption for most event study analyses. Therefore, before CISK intervention, the students are in the comparison group, but during the year of the intervention and thereafter, the students are in the treatment group. Tier I services are omitted from the analysis because every student in a school that offers Tier I

(2021) and Goodman-Bacon (2018). The panel event-study approach offers a rich array of leads and lags, as recently described by Freyaldenhoven, Hansen, and Shapiro (2019) and Freyaldenhoven et al. (2021).

services has access to them, and there is no attempt to exclude any student attending that school from receiving those services.

Sample A

The construction of the treatment group varies according to the construction of the student sample. The treatment group in the first sample (Sample A), which includes only students who remain in KPS all five years of elementary school, is constructed by considering students who received CISK Tier II or Tier III services.

Table 15 Total Number of Students in Sample A Receiving CISK Tier II or Tier III Services, by Grade

	A	B	C	D	E	F
	Comparison	First-time treatment	Total Tier II or Tier III interventions	Lingering effects	Total students	% treated
Grade						
1st	729	76	76	0	805	9.4
2nd	685	44	107	13	805	13.3
3rd	646	39	148	11	805	18.4
4th	623	23	44	138	805	5.5
5th	606	17	61	138	805	7.6
Total	3,289	199	436	300	4,025	10.8

NOTE: Subtracting column B from column C yields the number of times students received more than one intervention from CISK. In this sample, students received more than one intervention 237 times. Add the 199 first-time interventions to the 237 multiple interventions, and the total number of Tier II or Tier III services is 436. Adding the lingering effects (300) to the total interventions (436) yields 736 student-years.
SOURCE: Authors' analysis of Community Data System and the CISK database.

Since the potential effect of a CISK Tier II or Tier III service is expected to linger during the rest of elementary school (and even into the upper grades, although not recorded here), it is helpful to categorize students from Table 15 into four groups. The first group (column A) includes students who have never received CISK Tier II or Tier III services and never will throughout elementary school. In the literature on event study analysis, these students are considered the pure benchmark or comparison group, and most statistical programs, including the one we use for this evaluation, designate those students with a missing value. In the first sample, the number of comparison students across the five grades is 3,289. The second group (column B) includes students who switch from receiving no services to receiving services for the first time. Column C, labeled "Total Tier II or Tier III interventions," includes all students who received those services. The fourth group of students (column D) includes those who have what we call a lingering effect from receiving an intensive service. The lingering effect is present for 300 of the 4,025 student-years in the sample.[10] For example, suppose a student is designated a CISK student for the first time and is placed in the treated group (column B). However, the potential effects of that service or services could last through the remainder of elementary school. Add any more interventions the student receives to the first-time treatment (column B), and the total number of interventions is displayed in column C.

The lingering effect, as we call it, includes the years after which a student received Tier II or Tier III services. For example, if a student first joins CISK in first grade, the individual is in the treated group from first grade until the end of fifth grade, with the lingering effects counted in the four years after the intervention. To complicate the example, if a student received a Tier II or Tier III service in the years after the first treatment, say the third year, then the lingering effects take place in the years an intervention is *not* present—the second, fourth, and fifth years. The student is still considered in the treated group for all five elementary years, but for different reasons than simply the four years after the first and only intervention.

Students who never received services from CISK act as pure controls—the counterfactual on which the estimation of impacts is based. Differences between those pure-control students and

[10] A student-year consists of a year of intervention for the student. The number of student-years is the number of years a student receives an intervention (receives Tier II or Tier III services) multiplied by the number of students. If a student has two interventions in the five years of elementary school, then that student has two student-years of interventions. If 20 students had this same sequence of interventions, then 40 student-years would be tallied for that sequence. Because of the construction of the first, smaller sample, each of the 805 students is present in each grade during the five years covered by the sample. Therefore, there are 4,025 student-years (805 students times five years) in the sample.

the students that received CISK services are anchored at 0 in the omitted base period—i.e., the first lag in Equation (1). Hence, lags and leads capture the difference between treated and comparison students, compared to the prevailing difference in the omitted base period. Unbiased estimation of "post event" treatment effects thus relies fundamentally on the so-called parallel trends assumption: in the absence of treatment, it is assumed that treated and comparison students would have maintained similar trends in the baseline period.

Many statistical programs that analyze event studies require the generation of a variable that follows a member of the treatment group from a predetermined beginning point in time to some predetermined time after that person receives the treatment. We use the program in Stata, "eventdd," which requires the generation of a variable we call "time to event." As shown in Table 16, this variable takes on a negative number before the event occurs and a positive number thereafter. It assumes a value of "0" during the first year the student receives Tier II or Tier III services from CISK. Before that time, it assumes a value down to −4, depending upon the year the student first received services. After the intervention, it takes on a value of between 1 and 4 for the number of remaining years in elementary school. For example, if a student first received a service in second grade, then the student is in the treatment group for that school year and for the three school years remaining in elementary school. On the other hand, the table also shows that if a student received a service in grade five, that year would be the only year in which she is considered in the treatment group, since it is at the end of the sample period and there is no time left for lingering effects to occur in. For the previous four years, she is in the comparison group.

Using the hypothetical situation depicted in Table 16, "Student ID 1" is in the comparison group in school year 2014–2015, but for the next school year that student switches into the treatment group and remains in the treatment group through fifth grade. "Student ID 2" is in the comparison group until fifth grade (the 2018–2019 school year), at which time the student switches into the treatment group for that last year.

Table 16 Hypothetical Sequence of Participation in CISK Activities for Sample A

Student ID	School year	CISK service	Postevent	Time to event	Treatment (including lingering effects)	Comparison
1	2014–2015		0	−1	0	1
1	2015–2016	2015–2016	1	0	1	0
1	2016–2017		1	1	1	0
1	2017–2018		1	2	1	0
1	2018–2019		1	3	1	0
2	2014–2015		0	−4	0	1
2	2015–2016		0	−3	0	1
2	2016–2017		0	−2	0	1
2	2017–2018		0	−1	0	1
2	2018–2019	2018–2019	1	0	1	0
…	…	…	…	…	…	…

SOURCE: Authors' calculations.

Sample B

Constructing the treatment and comparison groups for Sample B is the same as for Sample A except the number of students is much larger and students are not required to stay in KPS for their entire elementary school experience. Sample B includes the 4,025 students in Sample A who entered KPS in first grade and remained in KPS through fifth grade. In addition to those students in the first sample, the larger sample includes students who entered KPS from elsewhere sometime during elementary school or left to go elsewhere. The number of students either entering or exiting KPS during the five-year period adds 1,316 more students, for a total of 5,341. The second and much larger difference in Sample B compared to A is that it includes students from all elementary grades (grades 1 to 5) each year instead of only one grade per year as in the first sample. This criterion adds 21,988 to the total number of students. Considering both criteria, the larger sample is 23,304 students greater than the 4,025 students in the smaller sample, for a total of 27,329 student-years in Sample B.[11]

As shown in Table 17, 1,268 students received CISK Tier II or Tier III services for the first time, or 4.6 percent of the 27,341 student-years. There were 2,208 student-years of total intensive treatment (receiving either Tier II or Tier III services), or 8.1 percent. If we include the 999 students who experienced what we call a lingering effect of the Tier II or Tier III services, we get 3,207 student-years, or 11.7 percent of the total 27,341 student-years. The pure comparison group is made up of those students who never received a Tier II or Tier III service during their elementary school experience. These students account for 88.3 percent of all student-years.

Table 17 Students in Sample B Receiving Tier II or Tier III Services, by Category and Year

	A	B	C	D	E	F
School year	Comparison	First-time treatment	Total Tier II or Tier III interventions	Lingering effects	Total students	% treated
2014–2015	5,211	349	349	0	5,560	6.3
2015–2016	4,953	239	496	68	5,517	9.0
2016–2017	4,614	348	793	55	5,462	14.5
2017–2018	4,697	108	209	520	5,426	3.9
2018–2019	4,659	224	361	356	5,376	6.7
Total	24,134	1,268	2,208	999	27,341	8.1

NOTE: Subtracting column B from column C yields the number of times students received more than one intervention from CISK. In this sample, students received more than one intervention 940 times. Add the 1,268 first-time interventions to the 940 multiple interventions, and the total number of Tier II or Tier III services is 2,208. Adding the lingering effects (999) to the total interventions (2,208) yields 3,207 student-years.
SOURCE: Authors' analysis of Community Data System and CISK database.

Table 18 illustrates hypothetical examples for three students, labeled in the first column as Student ID 1, 2, and 3. The first student (ID 1) shown in Table 18 received CISK Tier II or Tier III services for the first time in school year 2014–2015. This is denoted in the "First service"

[11] Note that there is a difference of 12 students between the two tables and thus between the two data files that account for the numbers in the two tables.

column as a 1 and in the “Time to event” column as a 0. The second student (ID 2) received CISK Tier II or Tier III services in school year 2018–2019, for which a 1 appears in the “First service” column and a 0 in the “Time to event” column. The third student (ID 3) received two groups of services, the first in 2014–2015 and the second in 2016–2017. In this case, that student is in the treatment group all five years, but for reasons different from the first student. The third student received her first CISK Tier II or Tier III treatment in 2014–2015, as did the first student. However, the third student received another treatment in 2016–2017, after a hiatus of one year. But because of the rule that the student remains in the treatment group to account for lingering effects after the first enrollment, the student was in the treatment group all five years of elementary school. Even though a student may have multiple interventions, as illustrated by the third student in Table 18, the “Time to event” variable starts only at the first occurrence of an intervention. After that, the lingering effect assumption is used, even though a subsequent intervention may occur.

Table 18 Hypothetical Sequence of Participation in CISK Activities for Sample B

Student ID	School year	First CISK service	First service	Time to event	Treatment (including lingering effects)	Comparison
1	2014–2015	2014–2015	1	0	1	1
1	2015–2016		0	1	1	0
1	2016–2017		0	2	1	0
1	2017–2018		0	3	1	0
1	2018–2019		0	4	1	0
2	2014–2015		0	−4	0	1
2	2015–2016		0	−3	0	1
2	2016–2017		0	−2	0	1
2	2017–2018		0	−1	0	1
2	2018–2019	2018–2019	1	0	1	0
3	2014–2015	2014–2015	1	0	1	0
3	2015–2016		0	1	1	0
3	2016–2017	2016–2017	0	2	1	0
3	2017–2018		0	3	1	0
3	2018–2019		0	4	1	0

SOURCE: Authors’ compilation of hypothetical example.

V. EVALUATION RESULTS

Four student outcomes are examined in the evaluation—two are academic outcomes and two are behavioral outcomes. The two academic outcomes are based on the value-added of the NWEA tests, as measured by the difference in NWEA reading and math scores between the fall and spring administration of the tests.[12] The NWEA tests are considered to be among the most comprehensive for the subjects considered, and the difference between the two tests—with one taken at the beginning of the school year and the other toward the end—offers an accurate way to assess a student's academic gain.[13] The other two student outcomes consist of engagement as measured by attendance rate and behavior as proxied by unexcused absences. The attendance rate, measured as the percentage of days in the classroom out of the total number of school days present, considers the commitment to attending school. The number of days of unexcused absences is a measure of an elementary student's adverse behavior (and of parents' behavior) by recording the number of days a student misses class but does not report that he or she is absent. The four outcome measures are examined for each of the two student samples.

T-tests of Covariates and Student Outcome Variables

We conduct tests to examine differences in the values of selected covariates and the four student outcomes between the treatment and comparison groups. More specifically, a *t*-test is estimated for each covariate and student outcome. Of the variables describing student characteristics in Sample A, only the shares of black students and students on free or reduced-price lunch were statistically significantly different between the two groups. In the treatment group, 56.6 percent of the students are black, whereas in the comparison group, 37.8 percent of

[12] Chetty et al. (2014a,b) describe the value-added of standardized test scores similar to what are used here.
[13] Thum and Kuhfield (2020) describe the NWEA assessments for math and reading.

students are black. Similarly, 97.7 percent of the students in the treatment group are eligible for free or reduced-price lunch, but only 75.1 percent of the students in the comparison group are eligible. For the outcome variables in Table 17, only the number of unexcused absences is statistically significantly different between the two groups. For the treatment group, there are an average of 7.77 days of unexcused absences; for the comparison group, 6.01 days. Table 19 displays the differences between values for the two groups, along with their *t*-statistics.

Table 19 *T*-tests of Differences in Various Covariates and Outcome Measures between Treatment and Comparison Groups in Sample A, for All Years

Black	FRPL	Hispanic	Male
0.188	0.226	0.026	−0.012
(7.63)	(10.85)	(1.36)	(0.49)
NWEAread_VA	**NWEAmath_VA**	**Attendance rate**	**Unexcused absences**
1.032	−0.081	−0.004	1.76
(1.81)	(−0.16)	(−1.25)	(4.57)

NOTE: *T*-tests were conducted using a regression approach in which the various covariates and outcomes were regressed against "treated2_3" individually. *T*-statistics are in parentheses.
SOURCE: Authors' analysis of KPS and CISK data.

As shown in Table 20, the results of the covariates or independent variables of the event study in Sample B follow closely those in Sample A, except for Hispanic students. Hispanic students join blacks and students on free or reduced-price lunch as statistically significantly different between the two groups. For the treatment group, 56.1 percent of students are black, whereas for the comparison group 40.0 percent of students are black. Some 98.2 percent of students in the treatment group are eligible for free or reduced-price lunch, and 79.7 percent of students in the comparison group are eligible. Unlike in Sample A, the difference between Hispanics in the two groups is statistically significant: 17.4 percent of students in the treatment group are Hispanic, whereas 13.2 percent in the comparison group are. For student outcome measures, only the difference between the two groups in the value-added of the NWEA math test is *not* statistically significantly different: The value-added of the NWEA reading test is 12.99 for the treatment group and 12.47 for the comparison group. For the attendance rate, the average value is 93.8 percent for the treatment group and 92.8 percent for the comparison group. For the number of days of unexcused absences, the treatment group averages 7.52 days and the comparison group 6.54.

Table 20 *T*-tests of Differences in Various Covariates and Outcome Measures between Treatment and Comparison Groups in Sample B, for All Years

Black	FRPL	Hispanic	Male
0.192	0.185	0.042	0.006
(17.9)	(21.1)	(5.50)	(0.53)

NWEAread_VA	NWEAmath_VA	Attendance rate	Unexcused absences
0.522	0.422	0.010	0.985
(2.02)	(1.67)	(4.84)	(5.38)

NOTE: *T*-tests were conducted using a regression approach in which the various covariates and outcomes were regressed against the binary variable signifying the individual participated in CISK.
SOURCE: Authors' analysis of KPS and CISK data.

Propensity Score Matching

Using propensity score matching, we attempt to replicate as closely as possible the criteria used by CISK to select students to receive intensive services. According to the executive director of CISK, the following criteria are used to select elementary students for intensive services:

- Absences 10–25 days (attendance)
- Suspended one or more days (behavior)
- NWEA spring math test between 25th–65th percentile (academics)
- NWEA spring reading test between 25th–65th percentile (academics)
- Core GPA of less than 2.5.[14]

The first four criteria are used in the propensity matching procedure, since we do not have ready access to students' GPAs. The student characteristics that are most relevant are poverty status

[14] In an email correspondence dated February 9, 2021, the executive director of CISK noted that "due to a variety of factors we don't limit enrollment for case management or the afterschool program to these criteria, which results in us typically working with students that are *below* one or more [of] the ranges identified below—especially with respect to NWEA scores for elementary students." A similar explanation of the selection process was used in the MDRC evaluation of CIS affiliate programs in North Carolina and Texas.

(free or reduced-price lunch program), race, ethnicity, gender, and the school the student attended.

The propensity matching procedure is estimated using a cross-sectional database for both samples. Since a student in the comparison group never receives CISK services in elementary school, the treated group includes students who receive intensive services in at least one grade during the five years in a KPS elementary school. Using the Stata program “psmatch2,” we match comparison group members with those from the treatment group for each year from 2015 through 2019. Included in the matching analysis are the various selection factors, listed above, used by CISK to select students into the treated group each year.

The results of the propensity matching procedure are reported in Table 21. The procedure implements full Mahalanobis matching to adjust for pretreatment observable differences between a group of treated and a group of untreated students. The intervention takes on the value of 1 if the individual student is in the treated group (receives intensive services or Tier II or Tier III activities) and 0 if in the comparison group (receives no services in any of the years). Since the dependent variable is binary, a logit estimation is used with standard errors on the treatment effects, computed under the assumption of independent observations, fixed weights, homoskedasticity, and independence with the propensity score.

The results show a strong (statistically significant) positive relationship between selection into the treated group and black students (versus whites and Asians), but not as much for Hispanics (versus whites and Asians), since only a few of the coefficients are statistically significant. We also find that students who fall within the 25th and 65th percentiles of the spring NCEW math and reading tests are more likely to be selected for CISK intensive services. However, the statistical significance of the estimates varies by grade and year. For some grades

Table 21 Propensity Score Estimates for All Grades, Selected Years, Larger Sample

Black#FRPL#male#Hispanic	2015		2016		2017		2018		2019	
0 0 0 1	0.65	(0.57)	1.46	(2.32)	1.68	(3.14)			1.76	(2.72)
0 0 1 0	1.19	(2.07)	0.43	(0.92)	0.69	(1.74)	−0.05	(−0.03)	−0.96	(−1.39)
0 0 1 1	0.76	(0.68)	1.75	(2.97)	−0.16	(−0.15)			1.07	(1.31)
0 1 0 0	1.79	(3.31)	1.64	(4.18)	2.21	(6.48)	3.01	(2.94)	1.51	(3.60)
0 1 0 1	2.56	(4.72)	2.28	(5.69)	2.67	(7.59)	3.36	(3.25)	2.11	(4.92)
0 1 1 0	2.03	(3.82)	2.08	(5.44)	2.27	(6.68)	3.39	(3.33)	1.67	(4.01)
0 1 1 1	2.13	(3.83)	2.31	(5.80)	2.65	(7.49)	3.85	(3.75)	2.41	(5.72)
1 0 0 0	2.35	(3.66)	2.21	(4.82)	1.85	(4.09)	2.75	(2.37)	1.68	(3.15)
1 0 1 0	1.27	(1.45)	1.54	(2.91)	1.83	(4.05)	2.38	(1.93)	1.38	(2.42)
1 1 0 0	2.74	(5.30)	2.52	(6.78)	2.82	(8.53)	3.02	(2.96)	1.80	(4.44)
1 1 1 0	2.83	(5.48)	2.37	(6.36)	2.79	(8.41)	3.62	(3.57)	2.17	(5.43)
Attendance rate	12.08	(4.00)	8.76	(4.55)	10.43	(6.83)	7.29	(2.90)	4.96	(2.94)
Days of unexcused absences	0.06	(3.00)	0.04	(3.07)	0.07	(6.16)	0.04	(2.18)	0.04	(2.79)
NWEA math 25th–65th pctile										
1st grade	0.29	(1.03)	0.08	(0.32)	−0.46	(−2.10)	−0.68	(−1.65)	0.34	(1.30)
2nd grade	0.15	(0.56)	0.53	(2.36)	−0.12	(−0.62)	0.33	(1.01)	−0.07	(−0.27)
3rd grade	0.77	(3.23)	0.61	(3.03)	0.13	(0.65)	0.75	(2.55)	0.39	(1.74)
4th grade	−0.03	(−0.09)	0.30	(1.34)	0.18	(1.06)	0.62	(2.24)	−0.48	(−1.75)
5th grade	−0.75	(−1.80)	−0.26	(−0.92)	0.09	(0.47)	−0.20	(−0.62)	0.35	(1.37)
NWEA reading 25th–65th pctile										
1st grade	0.57	(2.15)	0.19	(0.74)	0.05	(0.26)	0.38	(1.09)	−0.22	(−0.80)
2nd grade	0.91	(3.72)	0.21	(0.90)	0.21	(1.11)	−0.11	(−0.33)	0.43	(1.77)
3rd grade	0.59	(2.40)	0.23	(1.09)	0.15	(0.77)	−0.38	(−1.08)	0.58	(2.64)
4th grade	0.08	(0.27)	0.68	(3.11)	0.29	(1.71)	0.30	(1.03)	0.85	(3.71)
5th grade	0.03	(0.09)	−0.34	(−1.19)	0.40	(2.15)	0.57	(1.88)	0.01	(0.03)
Constant	−16.96	(−5.62)	−12.99	(−6.76)	−14.35	(−9.38)	−13.53	(−5.09)	−9.36	(−5.5)
Pscore range	(0, 0.33)		(0, 0.33)		(0, 0.38)		(0, 0.18)		(0, 0.27)	

NOTE: Since the treatment variable is binary, a logit analysis is used to estimate the coefficients. Z-scores are in parentheses. Because a student can be of only one race/ethnicity, some of the cells in the table have no observations and thus no coefficient. The first column indicates the various combinations of a student being black, male, Hispanic, or eligible for the free or reduced-price lunch program. A value of 1 indicates that a student is classified as one of the four categories, and a 0 indicates otherwise.
SOURCE: Authors' analysis of KPS and CISK data.

in some years, the estimates are statistically significant, but for other grades and years they are not. Attendance rates and the number of days of unexcused absences are statistically significant for all years, which means that these two criteria are consistently important for eligibility. Other variables that are statistically significant for all five years include students who are black, students who are black and eligible for free or reduced-price lunch, and students who are poor (as proxied by free or reduced-price lunch) and male. This could explain why 25 percent of the treated students are black and on free or reduced-price lunch and 27 percent are poor black males. Yet many of these students are not eligible for CISK services; roughly 16 percent each of both groups (poor blacks and poor black males) are in the comparison group.

Difference-in-Differences (DID) Estimates

Following the standard differences-in-differences strategy, we estimate the "average treatment effect on the treated" (ATET).[15] Two differences are used to compute a consistent estimate. The first difference computes the mean outcome of the treatment group before and after the treatment occurs. This difference subtracts out common time-invariant factors that occur before and after the time of treatment. However, this difference may not be enough because of the presence of additional factors that could confound the results. We may find that time-varying factors also come into play in affecting the mean outcome of the treatment group. For example, we may find that the change in the outcome of the treated group is due to time-varying confounders which extend across the time the treatment occurs. Therefore, the difference in the outcome of the treated group is not enough to estimate a consistent ATET. By constructing a comparison group with the assumption that the same confounding effects are present in this group as in the treatment group, except for the effect of the treatment, we can then subtract out these confounding effects between the treatment and comparison groups by taking a second difference—the difference between the difference before and after treatment for the treatment group as opposed to the comparison group.

One final condition must be met before the estimate of ATET is consistent. This condition is called the parallel-trends assumption. It states that, without the treatment, the trends of the treated and the outcome groups should be the same. One way to consider this assumption is to look at the graphs of the two groups before the treatment occurs. If the two trends are not

[15] We follow Angrist and Pischke (2009) for the estimation of the ATET using DID.

the same (that is, if they are statistically significantly different from each other), then the assumption stands.

We consider a variation of the difference-in-differences (DID) methodology when the treatment is binary and varies over time. In that way, there is a pretreatment period and a posttreatment period. In the pretreatment period, we test whether the average values of the treated group and the comparison group for the four outcomes—value-added of NWEA math and reading tests, attendance rate, and the number of unexcused absences—are statistically significantly different from zero. If they are, then the assumption described in the previous paragraph is satisfied. As a reminder, Sample A is by far the smaller of the two samples, with students remaining in KPS and progressing as a cohort from one grade to the next. The first two years constitute the preintervention period, and the last two represent the postintervention period. Unlike the event study, described below, we examine only the impact of the treatment for the year or years the student participates in CISK Tier II or Tier III services. We do not consider any lingering effects thereafter.

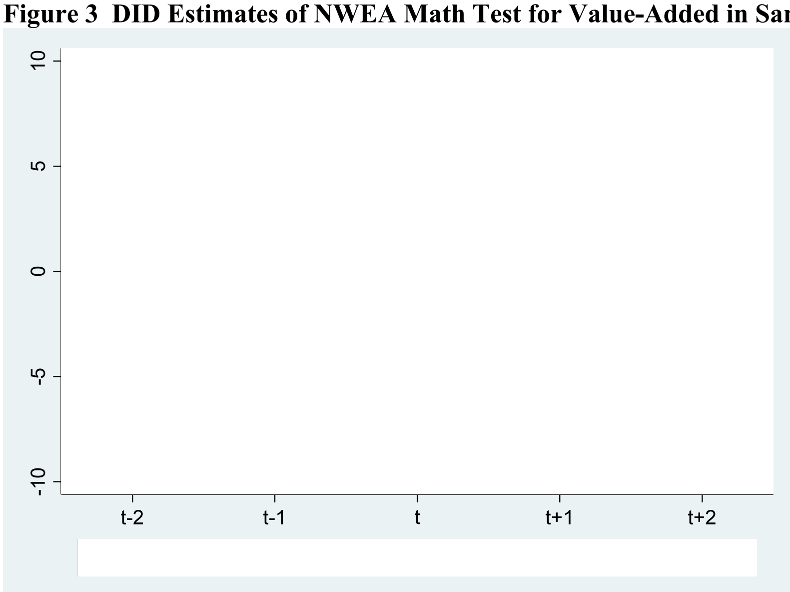
Table 22 shows the estimates of the 2017 intervention on the four outcomes considered for students in Sample A. In this case, only the value-added of the NWEA math test score (the difference between the NWEA test taken in the fall and the one in the spring of a given school year) is statistically significantly different from zero between the treated and comparison groups. Figure 3 shows the negative effect of the CISK participation during the intervention year compared to the differences in pre- and postintervention effects. The other three outcomes are not statistically significantly different from zero at the 95 percent confidence level. The parallel trend assumption passes for all four outcomes under one type of test and for three of the four in the other test.

Table 22 DID Estimates of Student Outcomes in Sample A

		2017 intervention	Parallel trend assumption using leads	Parallel trend assumption using time-trend
NWEAread_VA	DID	−1.58	F(2, 690)=1.72	F(1, 755)=3.95
	t-value	(−1.04)	"passed"	"not passed"
NWEAmath_VA	DID	−3.02	F(2, 708)=0.31	F(1, 756)=0.95
	t-value	(−2.09)	"passed"	"passed"
Attend_rate	DID	−0.010	F(2, 795)=0.49	F(1, 804)=0.86
	t-value	(−1.05)	"passed"	"passed"
Unexcused_days	DID	1.14	F(2, 799)=0.52	F(1, 804)=0.12
	t-value	(1.08)	"passed"	"passed"

NOTE: DID estimates based on the Stata program "tvdiff." The four outcomes are listed in the first column and the intervention is a binary variable that denotes enrollment of a student in CISK. The model used is OLS, two annual leads and lags are used around the outcome of 2017 for all grades, the robust method is used to estimate standard errors, and a *t*-test is used to estimate the time-trend approach for the statistical significance of the parallel trends. The term "passed" is used to show that the two tests of the parallel trend assumption passed---that is, that they are not statistically significantly different from zero.
SOURCE: Authors' analysis of KPS and CISK data.

Figure 3 DID Estimates of NWEA Math Test for Value-Added in Sample A, Based on Estimates in Table 22



NOTE: The vertical columns represent various confidence intervals for the estimates of the difference of the average value-added NWEA math test score between the treated group and the comparison group. We can see at time *t* that the estimate is negative and statistically significant. The vertical bars are shaded to indicate confidence intervals. The parallel trends assumption is satisfied by the estimates, since neither the estimate at point $t-2$ nor at point $t-1$ are statistically significant at a reasonable confidence level.
SOURCE: Authors' analysis of KPS and CISK data.

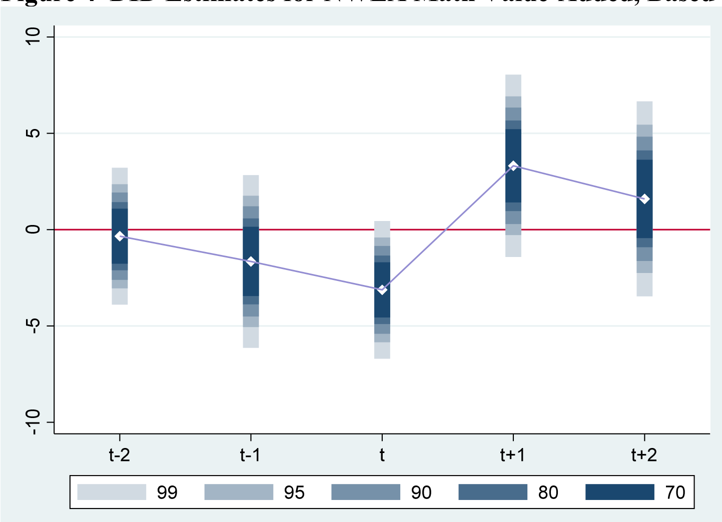
Table 23 shows DID estimates for the four student outcomes in Sample B. As with the estimates of the first sample, only the NWEA math value-added score is statistically significantly different from zero. For this sample, all the parallel trend assumptions pass, according to the *F*-tests associated with each outcome. Figure 4 displays graphically the estimates of the DID method associated with the value-added of the NWEA math test scores between the treated and comparison groups, which is the only outcome of the four that is statistically significant. If we were to rely only on the DID methodology for this evaluation, our conclusion would be that CISK Tier II or Tier III services reduce math test scores. However, the event study provides a richer analysis of the effects of CISK services.

Table 23 DID Estimates of Student Outcomes in Sample B

		2017 intervention	Parallel trend assumption using leads	Parallel trend assumption using time-trend
NWEAread_VA	DID	−1.43	F(2, 723)=1.33	F(1, 8859)=5.25
	t-value	(−0.96)	"passed'	"not passed"
NWEAmath_VA	DID	−3.13	F(2, 742)=0.50	F(1, 8867)=0.34
	t-value	(−2.26)	"passed"	"passed"
Attend_rate	DID	−0.007	F(2, 783)=0.60	F(1, 11401)=2.04
	t-value	(−0.73)	"passed"	"passed"
Unexcused_days	DID	0.86	F(2, 783)=1.15	F(1,
	t-value	(0.80)	"passed"	11401)=0.52"passed"

NOTE: The table displays the estimate of the DID model for the midpoint intervention with two leads before the intervention and two lags afterwards. The table also displays two different tests of the parallel trend assumption. The intervention variable, "treated2_3," is a binary variable in which 1 equals the treated group and 0 equals the comparison group. The DID model uses ordinary least squares for all estimations, and the value in parentheses is the *t*-statistic of the estimated coefficient.
SOURCE: Authors' analysis of KPS and CISK data.

Figure 4 DID Estimates for NWEA Math Value-Added, Based on Estimates in Table 23



NOTE: The vertical lines represent various confidence intervals for the estimates of the difference of the average value-added NWEA math test score between the treated group and the comparison group. We can see at time *t* that the estimate is negative and statistically significant. The vertical bars are shaded to indicate the various confidence intervals. The parallel trends assumption is satisfied by the estimates, since neither the estimate at point $t - 2$ or at point $t - 1$ are statistically significant at a reasonable confidence level. The variable "Treated2_3" refers to the treatment group that includes students who receive Tier 2 or Tier 3 services during a specific school year; a value of 1 signifies that a student is a member of the treatment group, and a value of 0 indicates that that student is a member of the comparison group.
SOURCE: Authors' analysis of KPS and CISK data.

Panel Event-Study Estimates

The panel event-study methodology is an econometric technique for panel data that compares outcomes before and after an event. In this evaluation, the event is the participation of KPS students in CISK Tier II or Tier III activities, and for both samples of students the effects of the event may linger, even though a subsequent event may not have occurred in the years following the initial event. A panel event study is analogous to difference-in-differences, but with a strong dynamic element.

Following the presentation by Clarke and Schythe (2020), we estimate the following event study model:

(2) $$y_{ijt} = \alpha + \delta_j + \gamma_t + \sum \beta_t \ \{t=\tau\} + X'\acute{\Gamma} + \varphi + \epsilon \ ,$$

where y_{ijt} is the outcome of student *i* in building *j* in time *t*, δ_i signifies the fixed effects across school buildings, γ_t is a fixed-year effect, X is a vector of time-variant control variables, and φ and ϵ are the typical error terms. The coefficient of interest is β_t, which estimates the effect of the event for each of the τ years included in the lags and leads of the estimation. The estimates capture the differences between treated and comparison students, compared to the prevailing difference in the omitted base period.

The event study corresponds to a DID-style model, and it estimates and plots a series of lag and lead coefficients related to an event. These lag and lead coefficients are all relative to the passage of an event of interest, which can occur at different moments in different units of the panel. The single binary variable compares the pre- and post-event period of the treatment group against the baseline of the comparison group. Lumping together the periods before and after the event into a single binary variable (1,0), instead of a dynamic structure, can be seen in the following equation:

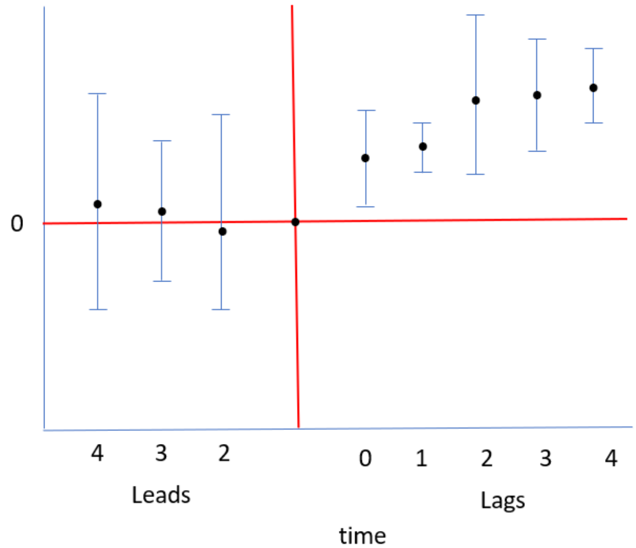
(3) $$y_{ijt} = \alpha + \delta + \gamma + \beta \{1[t \geq event_s]\} + X'\acute{\Gamma} + \varphi + \epsilon \,,$$

where X is a set of time-variant control variables, δ signifies the fixed effects across school buildings, γ is a fixed-year effect, and φ and ϵ are the typical error terms. Basically, the only difference between the two equations is the dynamic structure of the panel event-study approach versus the binary structure of the DID methodology related to the explanatory variables around the estimated β.

The estimates of the lags and leads of three model specifications are displayed in each of the tables for each student outcome in the two student samples. Model A is the simplest of the three specifications, with only the outcome and intervention variables included. Model B adds time and building fixed effects and adjusts the standard error for building clusters. Model C links the estimation more closely to the propensity score methods by reducing the sample to observations with a "pscore" of greater than its median value (0.06 for Sample B) for the sample. We have excluded additional covariates from the models, since they are already included in the propensity score matching estimations.

What follows is a set of estimates for each of the four student outcomes—value-added of NWEA reading and math tests, attendance rates, and the number of days of unexcused absences. The ideal figure, as shown in Figure 5, is one that shows no statistically significant relationship before the intervention (the lines portraying the confidence intervals intersect with the zero horizontal axis), and after the intervention the confidence interval lines are either completely above or completely below the horizontal axis and do not cross the axis, signifying that the estimates are

Figure 5 Ideal Configuration of Estimates for the Panel Event-Study Methodology



SOURCE: Drawn by the authors for illustrative purposes. Prior to lead 1, the coefficients are statistically insignificant; during and after lag 0, the estimates are positive and statistically significant.

statistically significantly different from zero. The statistical insignificance of the pre-event data is related to the parallel-trends assumption, and the statistical significance of the post-event data indicates the persistence of the effect over the included time. In the case of the ideal situation, as depicted in Figure 5, the average outcome of the treatment group is shown to be greater than that of the comparison group for all four lags.

Sample A Estimates

As before, four student outcomes are considered in the panel event-study estimation—value-added for the NWEA tests in reading and math, attendance rate, and the number of days of unexcused absences. Table 24 shows the estimates and *t*-statistics for the lags and leads in the panel event study for the value-added of NWEA reading tests. Model A includes only two variables, the outcome variable and the intervention variable. Model B adds a year and building effect. Model C restricts the sample size by including only that part of the sample that is greater than the median value of pscore, the estimated propensity score, which is generated from the propensity scoring program. Restricting the sample in this way uses only those values of pscore that are above the median, which means that the comparison group will include students who are closer to the observed characteristics of the treatment group.

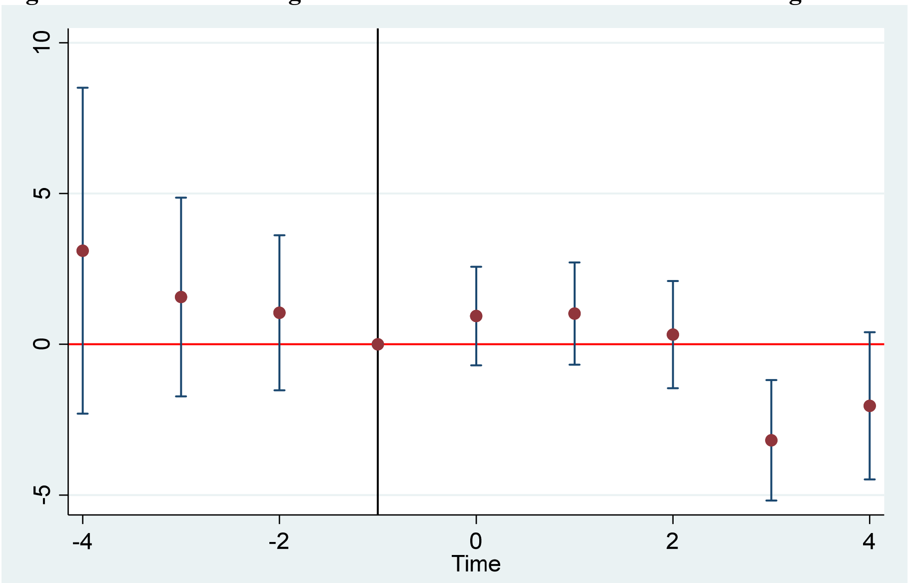
Table 24 Estimates from the Panel Event Study for NWEA Reading Value-Added in Sample A

	NWEAread_VA		
	A	B	C
Lead4	3.10	0.58	0.40
	(1.13)	(0.26)	(0.27)
Lead3	1.57	−0.87	−3.17
	(0.93)	(−0.47)	(−1.00)
Lead2	1.05	−1.38	−1.30
	(0.80)	(−1.32)	(−0.71)
Lag0	0.94	−0.26	−1.72
	(1.12)	(−0.47)	(−2.52)
Lag1	1.02	0.09	−0.30
	(1.18)	(0.08)	(−0.24)
Lag2	0.32	0.86	0.77
	(0.35)	(1.17)	(0.85)
Lag3	−3.19	−0.64	−0.82
	(−3.13)	(−0.71)	(−0.65)
Lag4	−2.04	2.72	2.29
	(−1.64)	(2.73)	(1.86)
Constant	11.83	13.8	14.84
	(60.87)	(18.62)	(15.12)
R-squared	0.005	0.114	0.102
Obs	3,541	3,541	2,019
Time / building FE		X	X
Pscore>0.32			X
Cluster(building) VCE		X	X

NOTE: Models A–C use the procedure "eventdd" found in Stata to estimate the coefficients. Model A estimates only leads and lags. Models B and C use year and building binary variables with the standard errors clustered around the buildings, and Model C adds the "if" statement based on the median value of pscore to reduce the sample by roughly half.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

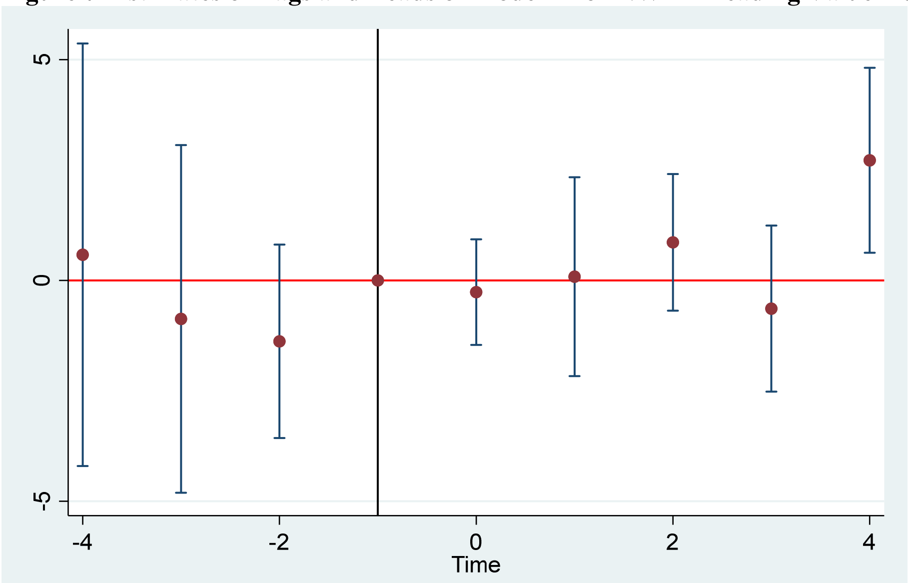
Of the three specifications of the models (Figures 6–8) estimating the average treatment effect of the student outcome of the value-added of the NWEA reading test, with the cluster computation of the standard errors, only Model B yields positive and statistically significant estimates, and this for only one of the four lags. While most specifications have a similar pattern of effects, only the fourth lead in Model B shows statistical significance in this way. At face value, it appears that KPS students who receive CISK services from Tier II or Tier III achieve higher test results in reading between fall and spring of the fourth year after they receive the services. The impact of the treatment appears to jump around before the fourth lead in Models B and C, and none of the impacts is statistically significant.

Figure 6 Estimates of Lags and Leads of Model A for NWEA Reading Value-Added in Sample A



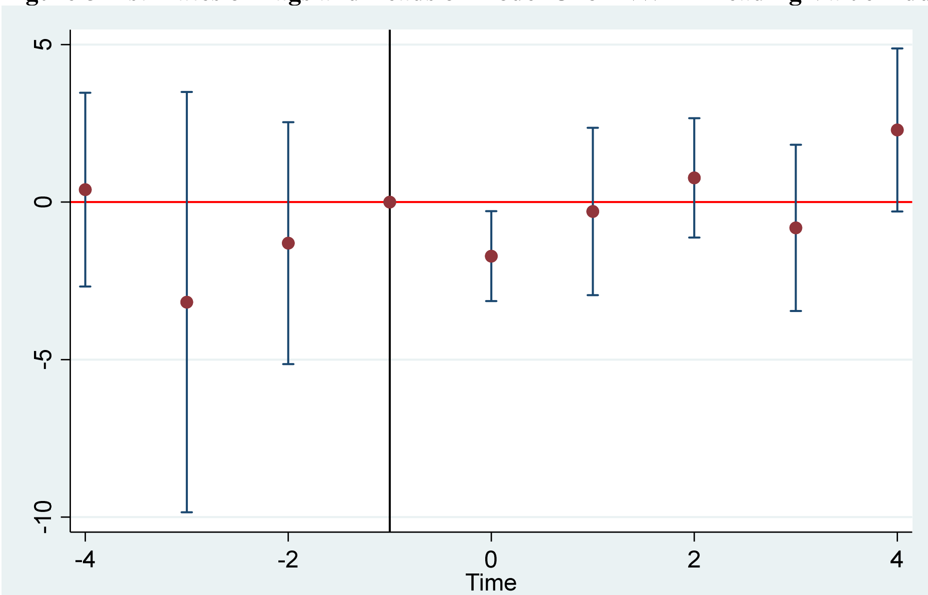
NOTE: Estimates are displayed in Table 24. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 7 Estimates of Lags and Leads of Model B for NWEA Reading Value-Added in Sample A



NOTE: Estimates are displayed in Table 24. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 8 Estimates of Lags and Leads of Model C for NWEA Reading Value-Added in Sample A



NOTE: Estimates are displayed in Table 24. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Considering the estimates of the three models for the value of the NWEA reading test, we conclude that there is no effect of CISK Tier II or Tier III services on this academic outcome. The effects are for the most part not statistically significant, and the results are scattered from positive to negative. For those coefficients that are statistically significant, the estimate of the third lag in Model A is negative, the estimate of the fourth lag in Model B is positive, and the estimate of the zero lag in Model C is negative.

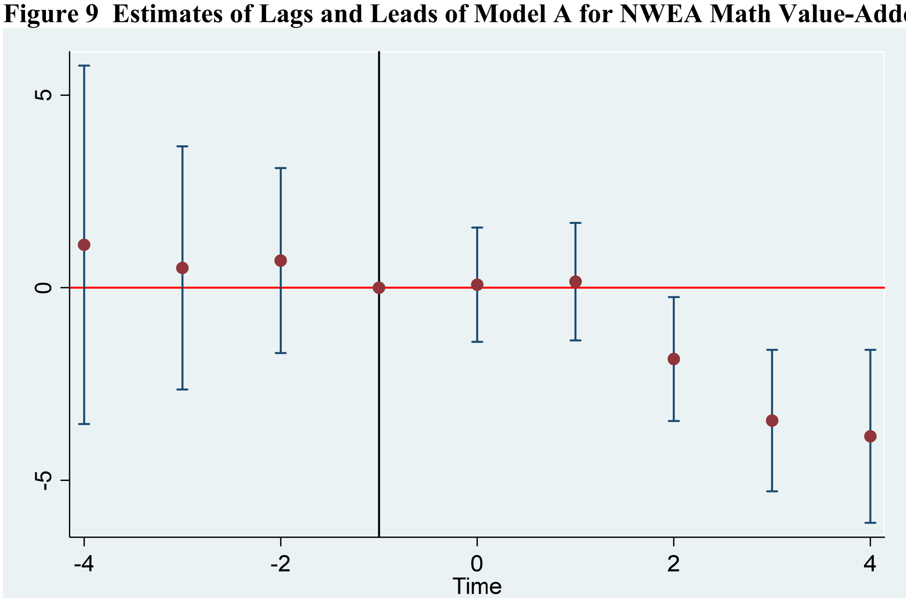
For the value-added of the NWEA math test (shown in Table 25 and Figures 9–11), Model A shows a clear decline from lag 2 and afterward. However, the other two models show statistically insignificant estimates for all lags and no negative tendency as shown in Model A. Therefore, it is unclear from Sample A what affect CISK services might have on the second academic outcome.

Table 25 Estimates from the Panel Event Study for NWEA Math Value-Added in Sample A

	NWEAmath_VA		
	A	B	C
Lead4	1.11	−1.00	2.79
	(0.47)	(−0.58)	(0.72)
Lead3	0.51	−0.99	−1.80
	(0.32)	(−0.60)	(−0.64)
Lead2	0.71	−0.80	−0.180
	(0.58)	(−1.04)	(−0.16)
Lag0	0.08	−0.39	−0.037
	(0.10)	(−0.54)	(−0.04)
Lag1	0.16	0.44	0.965
	(0.20)	(0.50)	(0.97)
Lag2	−1.85	−0.60	0.348
	(−2.26)	(−0.52)	(0.24)
Lag3	−3.45	−0.77	−0.374
	(−3.68)	(−0.86)	(−0.37)
Lag4	−3.86	0.81	0.419
	(−3.37)	(0.59)	(0.37)
Constant	14.57	16.40	14.98
	(81.49)	(18.89)	(16.16)
R-squared	0.008	0.098	0.114
Obs	3,552	3,552	2,034
Time / building FE		X	X
Pscore>0.32			X
Cluster(building) VCE		X	X

NOTE: Models A–C use procedure "eventdd" found in Stata. Models B and C use year and building binary variables with the standard errors clustered around the buildings, and Model C adds the "if" statement based on the median value of pscore to reduce the sample by roughly half.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 9 Estimates of Lags and Leads of Model A for NWEA Math Value-Added in Sample A

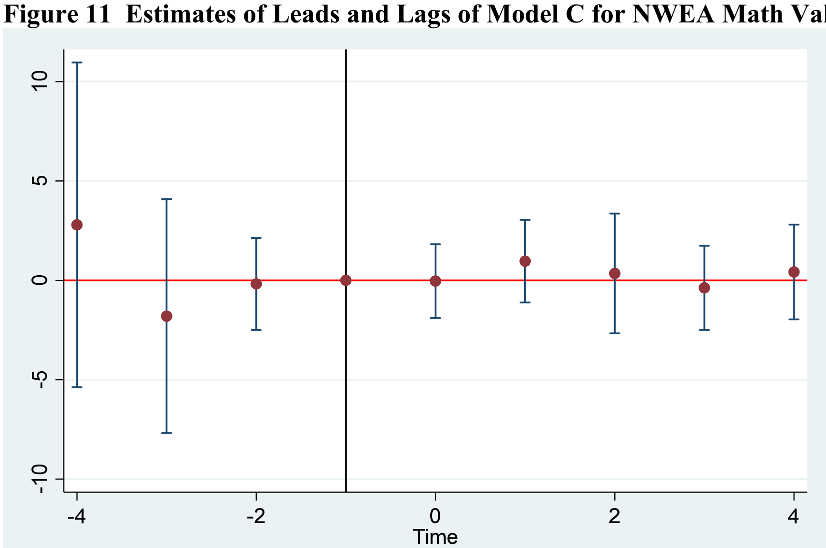


NOTE: Estimates are displayed in Table 25. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 10 Estimates of Lags and Leads of Model B for NWEA Math Value-Added in Sample A

NOTE: Estimates are displayed in Table 25. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 11 Estimates of Leads and Lags of Model C for NWEA Math Value-Added in Sample A



NOTE: Estimates are displayed in Table 25. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

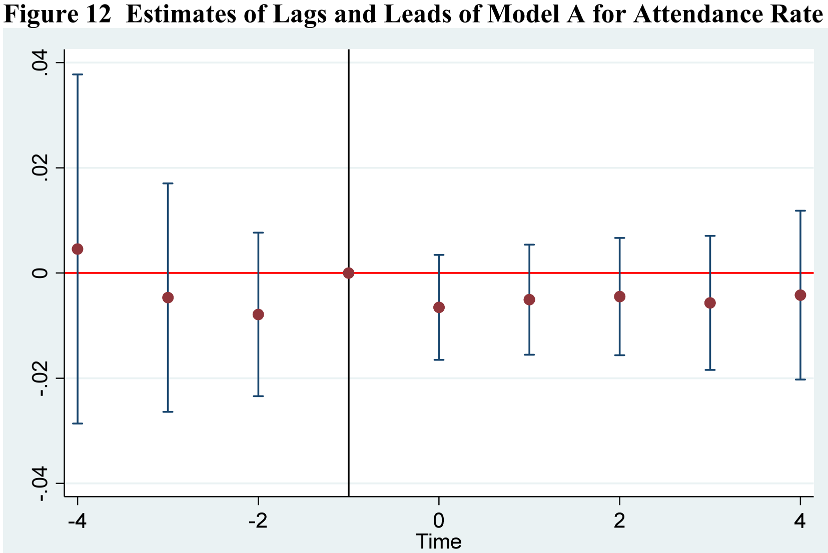
Table 26 displays the estimates of the three models (Figures 12–14) for attendance rate. A slightly clearer direction is seen for attendance rate, especially in Models B and C. In Models B and C, we see positive estimates that increase with greater lags. However, only Model C shows two lags with statistically significant coefficients, lag 2 and lag 4, although the same pattern is evident in both models.

Table 26 Estimates from the Panel Event Study for Attendance Rate in Sample A

	Attendance rate		
	A	B	C
Lead4	0.005	0.014	0.031
	(0.27)	(2.17)	(3.14)
Lead3	−0.005	0.0002	0.005
	(−0.42)	(0.03)	(0.48)
Lead2	−0.008	−0.0004	0.002
	(−0.99)	(−0.05)	(0.18)
Lag0	−0.007	0.0009	0.007
	(−1.28)	(0.20)	(1.21)
Lag1	−0.005	0.003	0.010
	(−0.95)	(0.73)	(1.68)
Lag2	−0.004	0.007	0.014
	(−0.79)	(1.38)	(2.93)
Lag3	−0.006	0.005	0.013
	(−0.87)	(0.52)	(0.95)
Lag4	−0.004	0.008	0.018
	(−0.51)	(1.04)	(2.36)
Constant	0.94	0.94	0.914
	(755.50)	(292.72)	(201.29)
R-squared	0.001	0.054	0.065
Obs	3,988	3,988	2,299
Time / building FE		X	X
Pscore>0.32			X
Cluster(building) VCE		X	X

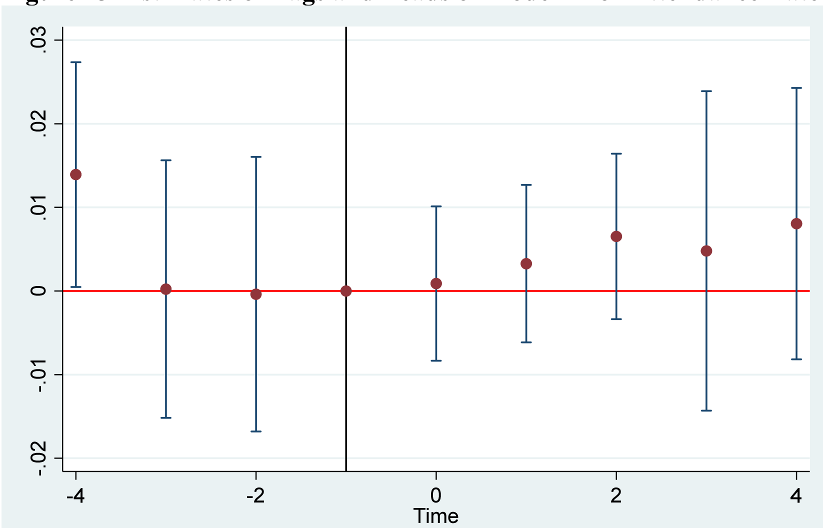
NOTE: Models A–C use procedure "eventdd," found in Stata. Models B and C use year and building binary variables with the standard errors clustered around the buildings, and Model C adds the "if" statement based on the median value of pscore to reduce the sample by roughly half.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 12 Estimates of Lags and Leads of Model A for Attendance Rate in Sample A



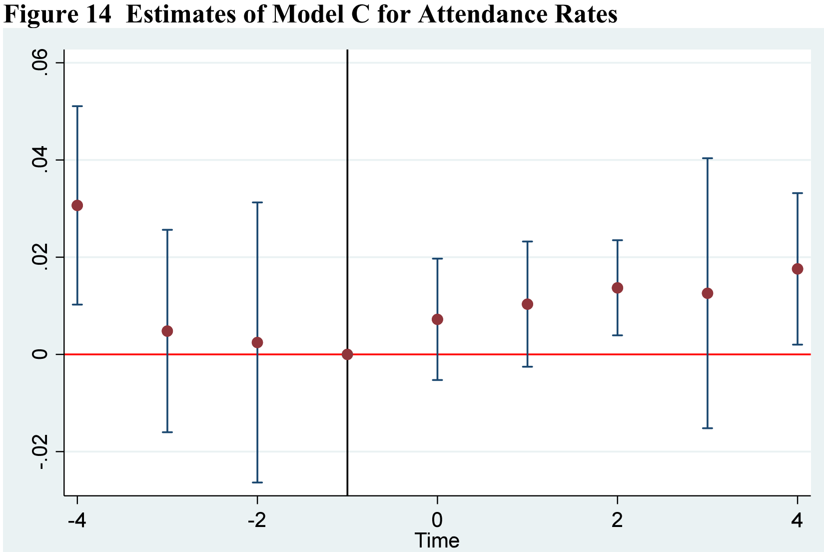
NOTE: Estimates are displayed in Table 26. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 13 Estimates of Lags and Leads of Model B for Attendance Rate in Sample A



NOTE: Estimates are displayed in Table 26. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 14 Estimates of Model C for Attendance Rates



NOTE: Estimates are displayed in Table 26. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

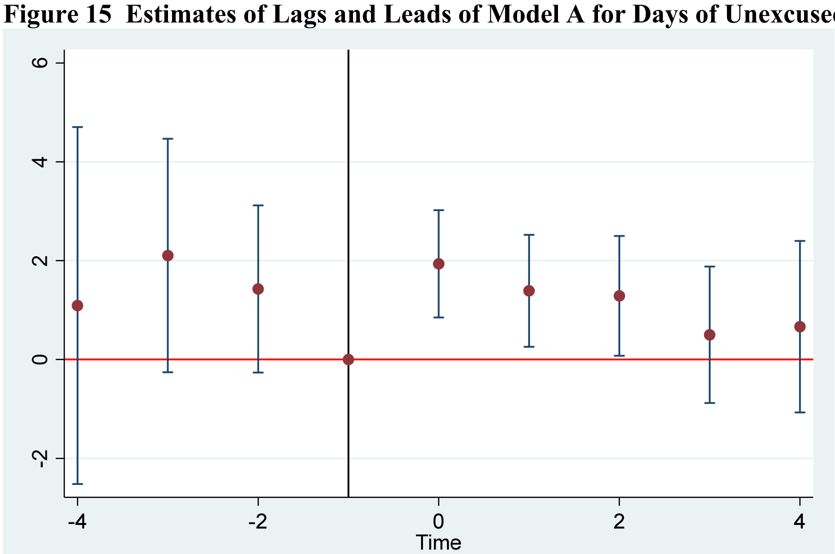
In Table 27, unexcused absences, measured in days, are included as a measure of student (and parent) adverse behavior. It is fair to say that the greater the number of days a year that a student has unexcused absences, the less engagement that student has with school and with his or her education. We find that CISK services have no effect on this student outcome. Few of the coefficients are statistically significant, and most estimates hover around the baseline (Figures 15–17).

Table 27 Estimates from the Panel Event Study for Days of Unexcused Absences in Sample A

	Unexcused days		
	A	B	C
Lead4	1.09	−0.98	−4.04
	(0.59)	(−1.04)	(−2.98)
Lead3	2.10	0.35	−0.450
	(1.75)	(0.33)	(−0.34)
Lead2	1.43	−0.57	−1.48
	(1.65)	(−0.71)	(−1.06)
Lag0	1.94	0.32	−0.717
	(3.50)	(0.42)	(−0.73)
Lag1	1.39	−0.18	−1.50
	(2.41)	(−0.27)	(−2.08)
Lag2	1.29	−0.89	−1.07
	(2.08)	(−0.12)	(−1.27)
Lag3	0.50	−0.43	−1.29
	(0.71)	(−0.62)	(−1.18)
Lag4	0.66	0.029	−0.534
	(0.75)	(0.03)	(−0.50)
Constant	5.91	9.16	13.06
	(43.80)	(27.84)	(31.32)
R-squared	0.006	0.132	0.162
Obs	4,022	4,022	2,324
Time / building FE		X	X
Pscore>.32			X
Cluster(building) VCE		X	X

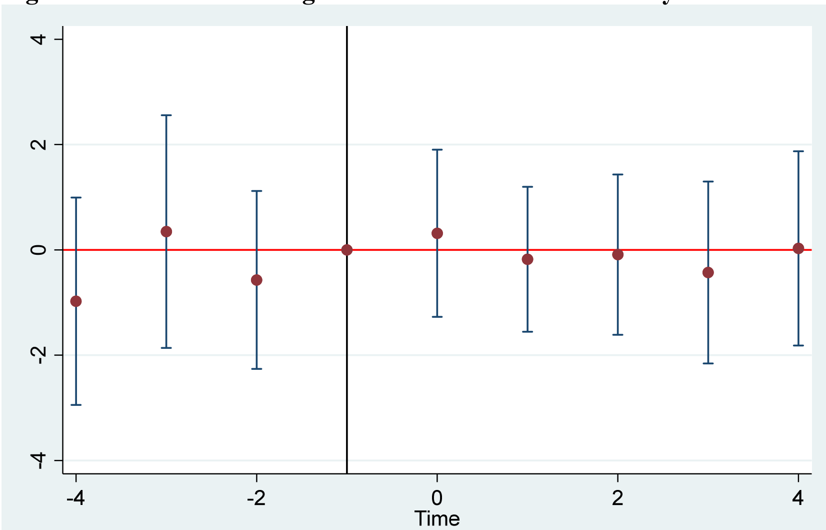
NOTE: Models A–C use procedure "eventdd" found in Stata. Models B and C use year and building binary variables with the standard errors clustered around the buildings, and Model C adds the "if" statement based on the median value of pscore to reduce the sample by roughly half.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 15 Estimates of Lags and Leads of Model A for Days of Unexcused Absences in Sample A



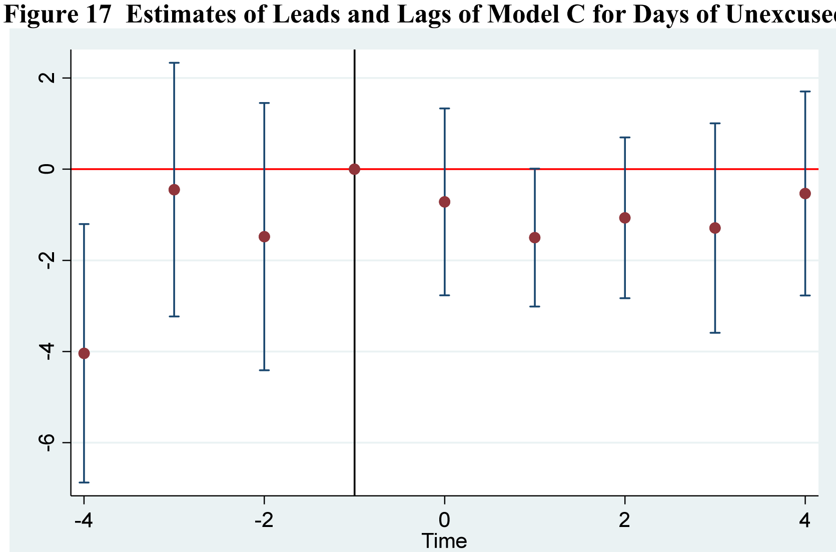
NOTE: Estimates are displayed in Table 27. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 16 Estimates of Lags and Leads of Model B for Days of Unexcused Absences in Sample A



NOTE: Estimates are displayed in Table 27. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 17 Estimates of Leads and Lags of Model C for Days of Unexcused Absences in Sample A



,
NOTE: Estimates are displayed in Table 27. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Sample B Estimates

The next set of estimates is based on Sample B, in which all students in grades one through five are included each year, instead of on Sample A, which includes only those students who remain in KPS throughout grades one through five and thus are in the same cohort from first through fifth grade. In Sample B, the number of students is significantly larger, which appears to help with the precision of the estimates. However, it may also be the mix of students that helps with precision, since each of the five grades is represented in this sample each year.

As we can see from Tables 28–31, many more of the estimates are statistically significant than was found for Sample A. However, most of the estimates still exhibit the same general pattern as was found in the first sample, albeit more pronounced for some. Estimates of leads and lags for the two academic outcomes in Sample B (Figures 18–23) show a more pronounced negative and downward direction of the estimates of the lag variables. In addition, many more of the coefficients are statistically significant. These estimates could lead us to conclude that CISK activities have a downward effect on student achievement for reading and that the treatment group is more negative than the comparison group. However, the parallel-trend assumption is not satisfied in this case, which nullifies the conclusion. The fact that many of the lead variables are statistically significant invalidates this assumption and throws into question the causal validity of the estimates. One could conclude that the trend established before the event continues afterward. Therefore, it is fair to deduce that CISK services have no causal effect on students' academic outcomes, as measured by the value-added of the two NWEA tests.

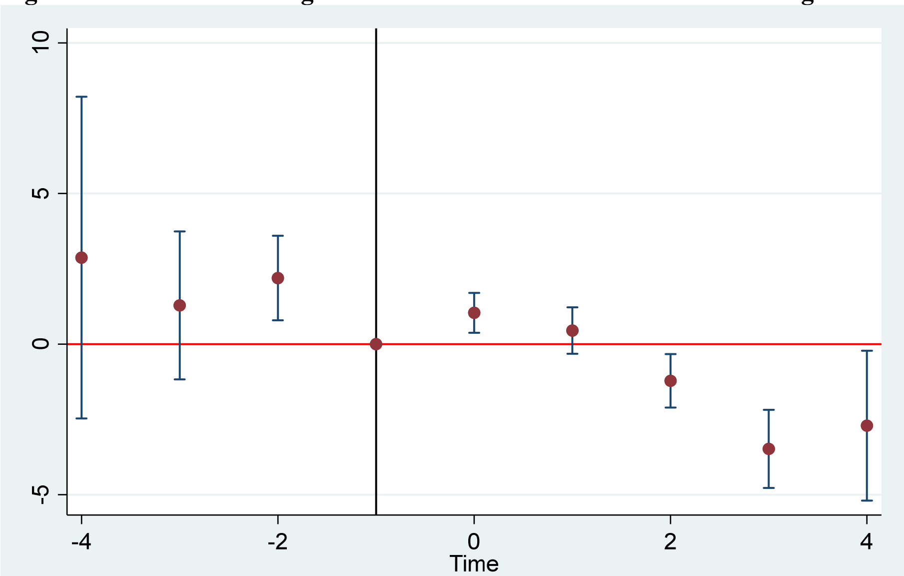
However, for Sample B, CISK services clearly have a causal effect on attendance rates, as shown in Table 30 and Figures 24–26, since the parallel-trend assumption is supported and all the lag coefficients are positive, with many being statistically significant. Therefore, there is no ambiguity in the positive effect of CISK Tier II or Tier III services on a student's attendance rate. Furthermore, all models support the parallel-trends assumption, in that all coefficients of the lead variables are statistically insignificant. The estimates suggest that CISK services increase attendance rates by approximately 1 percentage point during the year of intervention and up to 1.5 percentage points during the second year after intervention.

Table 28 Estimates from the Panel Event Study for NWEA Reading Test Value-Added in Sample B

	NWEAread_VA		
	A	B	C
Lead4	2.87	6.14	7.39
	(1.05)	(2.68)	(4.02)
Lead3	1.29	2.75	2.16
	(1.03)	(2.01)	(1.46)
Lead2	2.19	2.83	3.24
	(3.06)	(2.85)	(3.10)
Lag0	1.03	0.97	0.80
	(3.08)	(2.66)	(1.94)
Lag1	0.45	−0.53	−0.43
	(1.15)	(−1.29)	(−1.01)
Lag2	−1.22	−2.80	−2.72
	(−2.69)	(−5.51)	(−5.57)
Lag3	−3.48	−4.90	−4.03
	(−5.26)	(−6.65)	(−5.58)
Lag4	−2.71	−4.13	−4.38
	(−2.13)	(−3.68)	(−3.16)
Constant	12.50	10.29	9.87
	(156.94)	(30.87)	(25.60)
R-squared	0.0029	0.036	0.034
Obs	21,880	21,880	11,896
Time / building FE		X	X
Pscore>0.06			X
Cluster(building) VCE		X	X

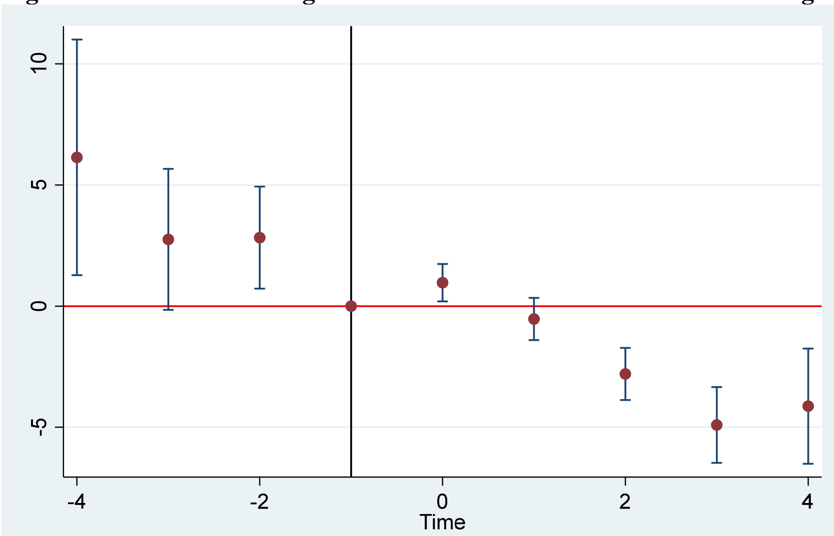
NOTE: Models A–C use procedure “eventdd” found in Stata. Models B and C use year and building binary variables with the standard errors clustered around the buildings, and Model C adds the “if” statement based on the median value of pscore to reduce the sample by roughly half.
SOURCE: Authors’ analysis of data from the Community Data System combined with the CISK database.

Figure 18 Estimates of Lags and Leads of Model A for NWEA Reading Test Value-Added in Sample B



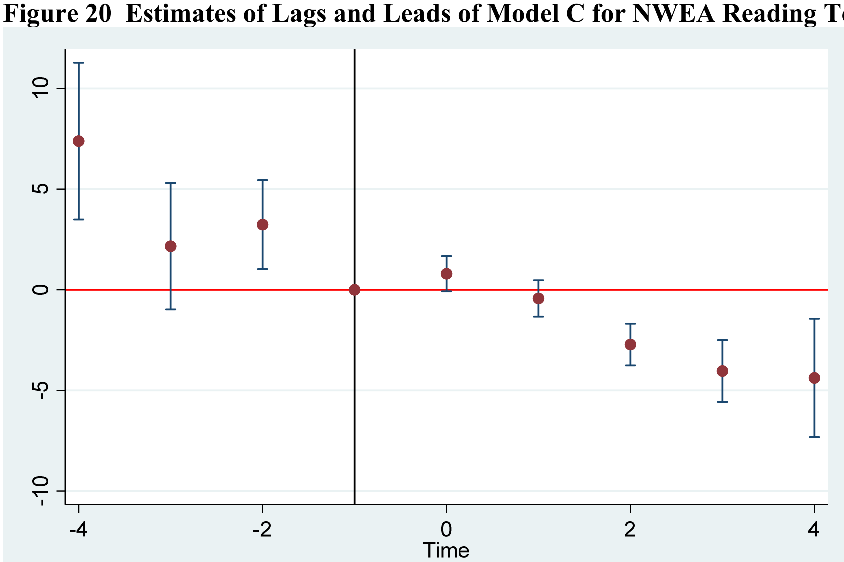
NOTE: Estimates are displayed in Table 28. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors’ analysis of data from the Community Data System combined with the CISK database.

Figure 19 Estimates of Lags and Leads of Model B for NWEA Reading Test Value-Added in Sample B



NOTE: Estimates are displayed in Table 28. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 20 Estimates of Lags and Leads of Model C for NWEA Reading Test Value-Added in Sample B



NOTE: Estimates are displayed in Table 28. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

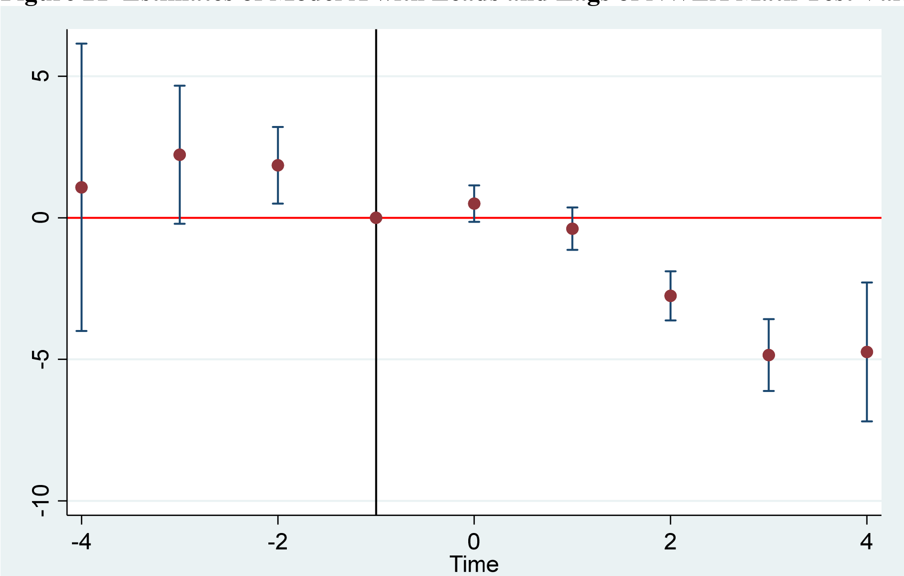
The number of days of unexcused absences is not causally affected by CISK services. Model A (Figure 27) estimates that all lag coefficients are positive and three are statistically significant, whereas Models B and C (Figures 28 and 29) show that all estimates are negative but that only one estimate (lag 2 in Model C) is statistically significant. The inconsistency in the results across the three models and the general statistical insignificance of the estimates lead us to conclude that CISK activities do not causally affect unexcused absences.

Table 29 Estimates from the Panel Event Study for NWEA Math Test Value-Added in Sample B

	NWEAmath_VA		
	A	B	C
Lead4	1.08	4.37	1.91
	(0.42)	(1.94)	(1.12)
Lead3	2.23	4.31	4.12
	(1.79)	(3.24)	(2.86)
Lead2	1.85	3.03	3.98
	(2.68)	(3.83)	(3.93)
Lag0	0.50	0.96	1.19
	(1.52)	(2.05)	(2.39)
Lag1	−0.39	−0.48	−0.50
	(−1.01)	(−0.79)	(−1.07)
Lag2	−2.76	−3.82	−3.65
	(−6.22)	(−6.10)	(−6.29)
Lag3	−4.85	−5.73	−5.87
	(−7.49)	(−8.63)	(−10.85)
Lag4	−4.74	−6.23	−6.39
	(−3.79)	(−3.78)	(−5.27)
Constant	15.40	13.16	12.36
	(197.91)	(27.72)	(19.51)
R-squared	0.006	0.042	0.044
Obs.	22,019	22,019	11,974
Time / building FE		X	X
Pscore>0.06			X
Cluster(building) VCE		X	X

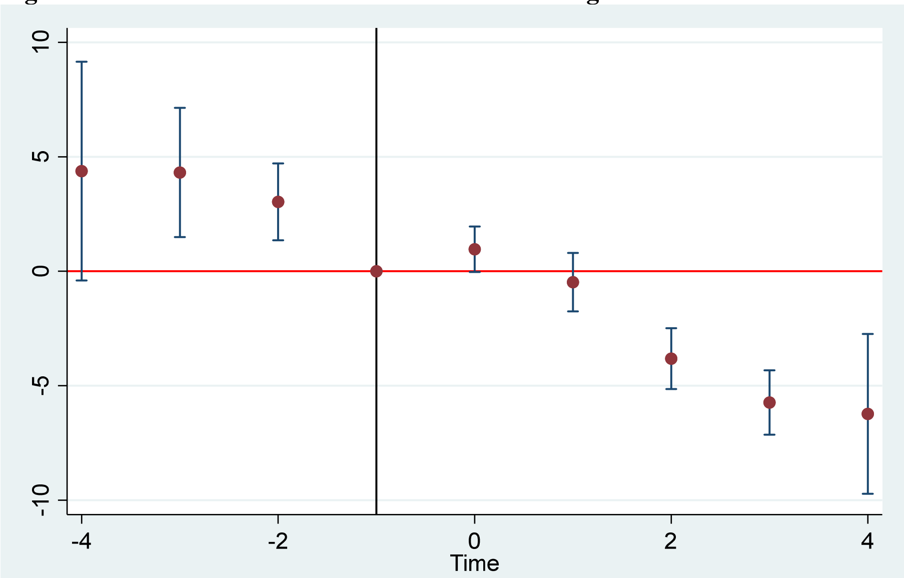
NOTE: Models A–C use procedure "eventdd" found in Stata. Models B and C use year and building binary variables with the standard errors clustered around the buildings, and Model C adds the "if" statement based on the median value of pscore to reduce the sample by roughly half.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 21 Estimates of Model A with Leads and Lags of NWEA Math Test Value-Added in Sample B



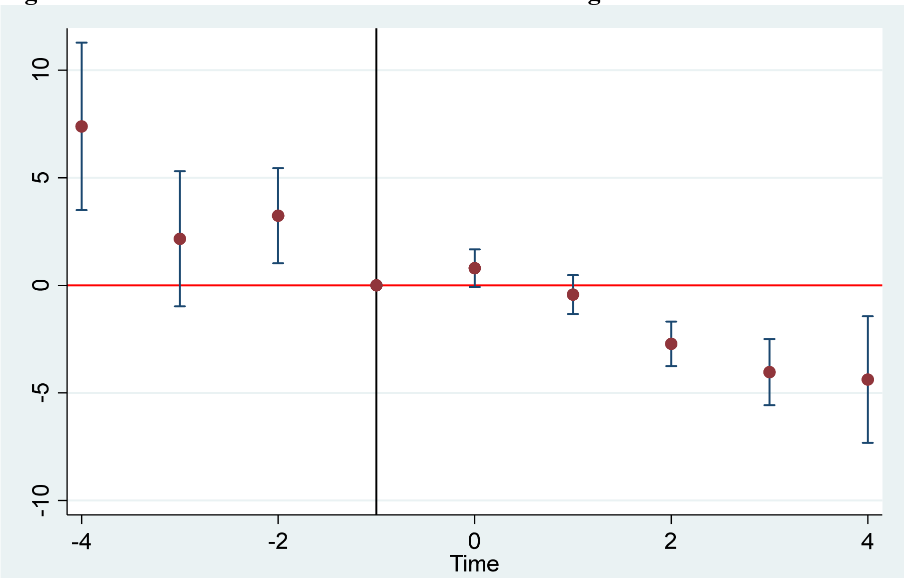
NOTE: Estimates are displayed in Table 29. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 22 Estimates of Model B with Leads and Lags of NWEA Math Test Value-Added in Sample B



NOTE: Estimates are displayed in Table 29. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 23 Estimates of Model C with Leads and Lags of NWEA Math Test Value-Added in Sample B



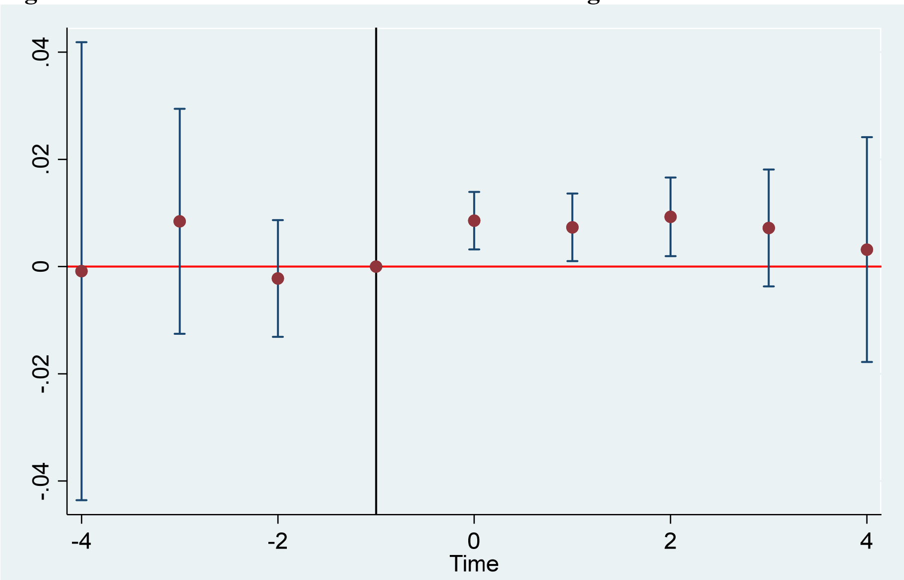
NOTE: Estimates are displayed in Table 29. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Table 30 Estimates from the Panel Event Study for Attendance Rate in Sample B

	Attendance rate		
	A	B	C
Lead4	−0.001	0.00	0.007
	(−0.04)	(0.00)	(0.79)
Lead3	0.008	0.009	0.003
	(0.79)	(1.78)	(0.42)
Lead2	−0.002	0.005	−0.0006
	(−0.40)	(1.15)	(−0.17)
Lag0	0.009	0.017	0.007
	(3.13)	(4.22)	(2.35)
Lag1	0.007	0.017	0.007
	(2.27)	(3.43)	(2.20)
Lag2	0.009	0.021	0.015
	(2.48)	(3.88)	(4.71)
Lag3	0.007	0.017	0.011
	(1.29)	(2.89)	(2.07)
Lag4	0.003	0.015	0.012
	(0.30)	(1.89)	(1.55)
Constant	0.93	0.93	0.94
	(1504.34)	(286.85)	(524.11)
R-squared	0.001	0.037	0.052
Obs	27,327	27,327	14,487
Time / building FE		X	X
Pscore>0.06			X
Cluster(building) VCE		X	X

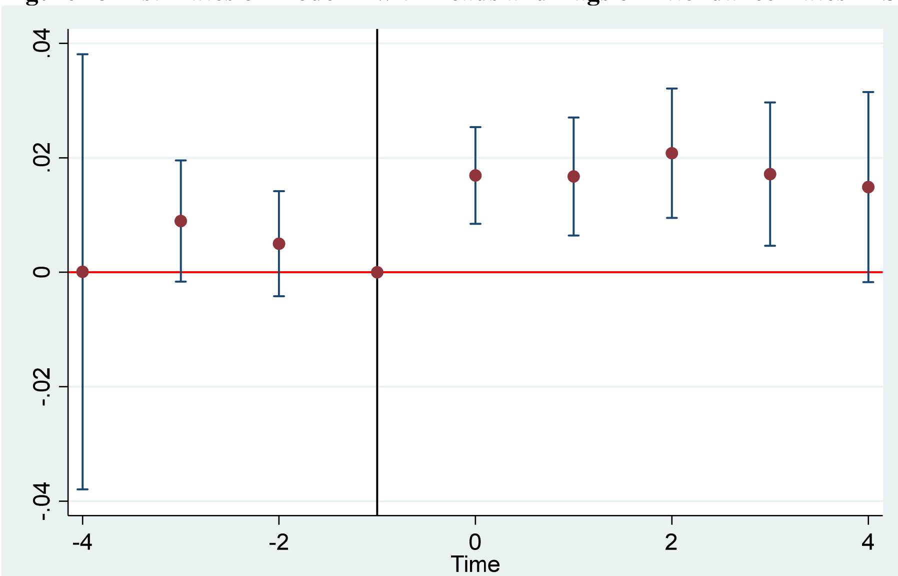
NOTE: Models A–C use procedure "eventdd" found in Stata. Models B and C use year and building binary variables with the standard errors clustered around the buildings, and Model C adds the "if" statement based on the median value of pscore to reduce the sample by roughly half.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 24 Estimates of Model A with Leads and Lags of Attendance Rates in Sample B



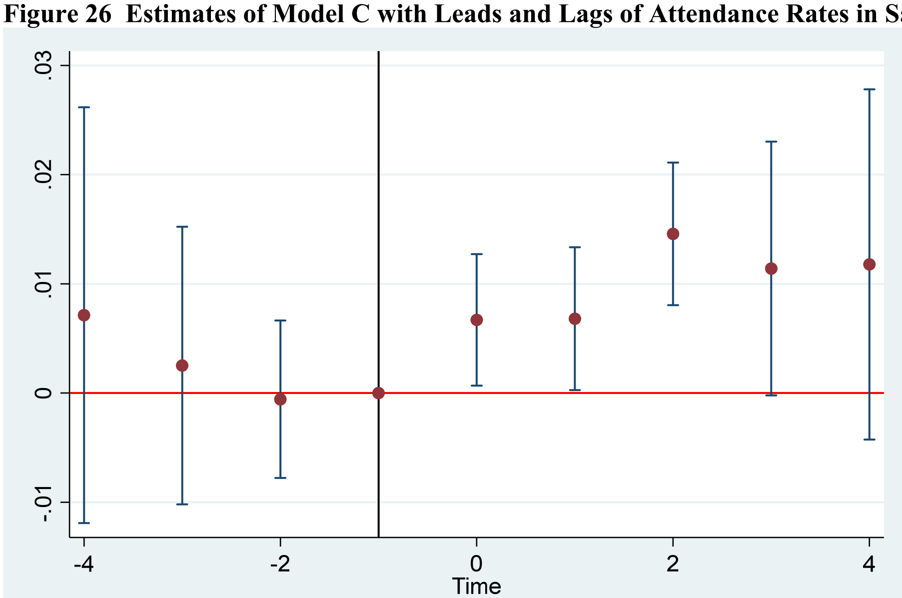
NOTE: Estimates are displayed in Table 30. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 25 Estimates of Model B with Leads and Lags of Attendance Rates in Sample B



NOTE: Estimates are displayed in Table 30. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 26 Estimates of Model C with Leads and Lags of Attendance Rates in Sample B



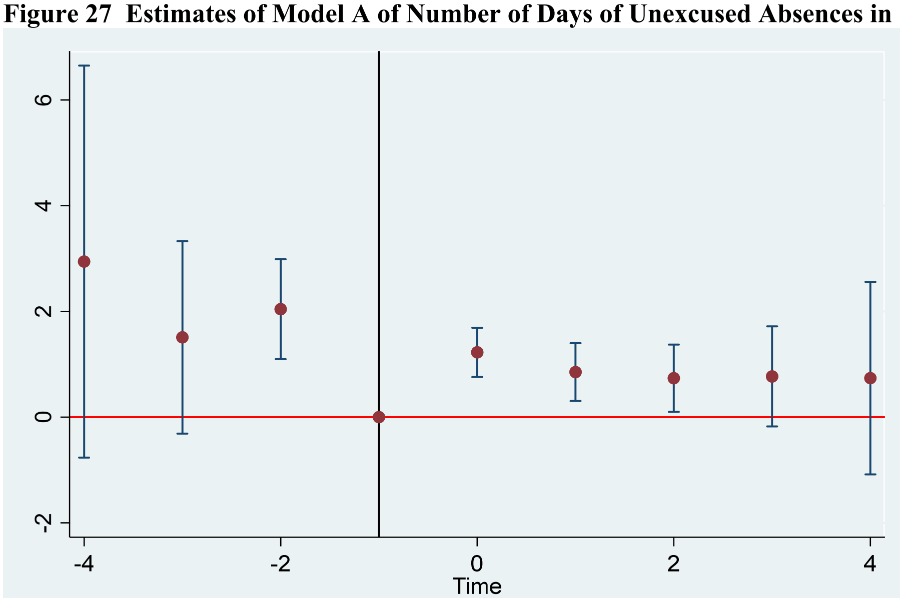
NOTE: Estimates are displayed in Table 30. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Table 31 Estimates from the Panel Event Study for Days of Unexcused Absences in Sample B

	Unexcused days		
	A	B	C
Lead4	2.94	2.12	0.14
	(1.56)	(1.00)	(0.09)
Lead3	1.51	0.78	0.59
	(1.62)	(1.05)	(0.61)
Lead2	2.04	0.63	0.22
	(4.24)	(0.98)	(0.41)
Lag0	1.23	−0.11	−0.45
	(5.16)	(−0.29)	(−1.26)
Lag1	0.85	−0.51	−0.48
	(3.05)	(−1.18)	(−1.04)
Lag2	0.74	−0.60	−1.19
	(2.26)	(−1.40)	(−3.10)
Lag3	0.77	−0.41	−1.03
	(1.59)	(−0.65)	(−1.30)
Lag4	0.74	−0.37	−1.78
	(0.79)	(−0.45)	(−1.84)
Constant	6.48	8.51	8.58
	(121.02)	(46.30)	(35.91)
R-squared	.002	0.095	0.060
Obs	27,327	27,327	14,487
Time / EWBuilding FE		X	X
Pscore>0.06			X
Cluster(EWBuilding) VCE		X	X

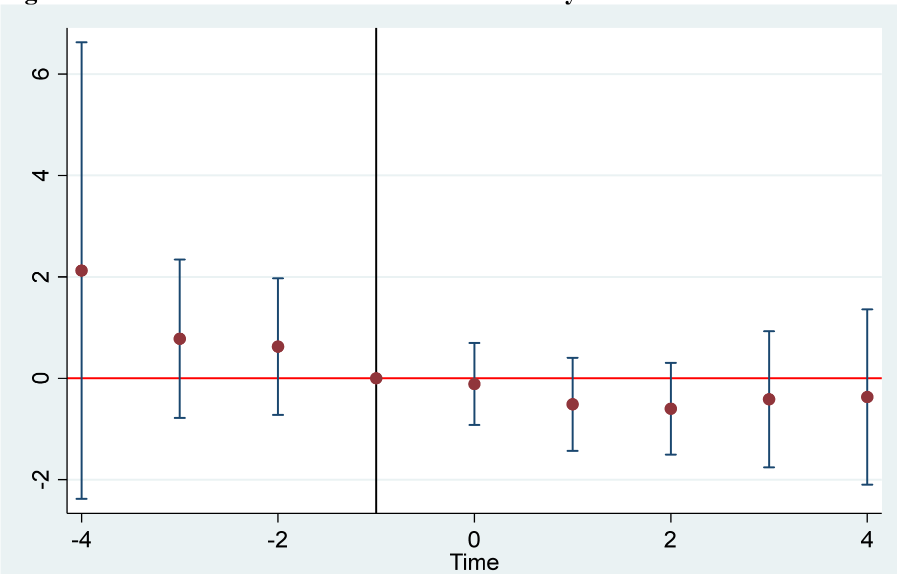
NOTE: Models A–C use procedure “eventdd” found in Stata. Models B and C use year and building binary variables with the standard errors clustered around the buildings, and Model C adds the “if” statement based on the median value of pscore to reduce the sample by roughly half.
SOURCE: Authors’ analysis of data from the Community Data System combined with the CISK database.

Figure 27 Estimates of Model A of Number of Days of Unexcused Absences in Sample B



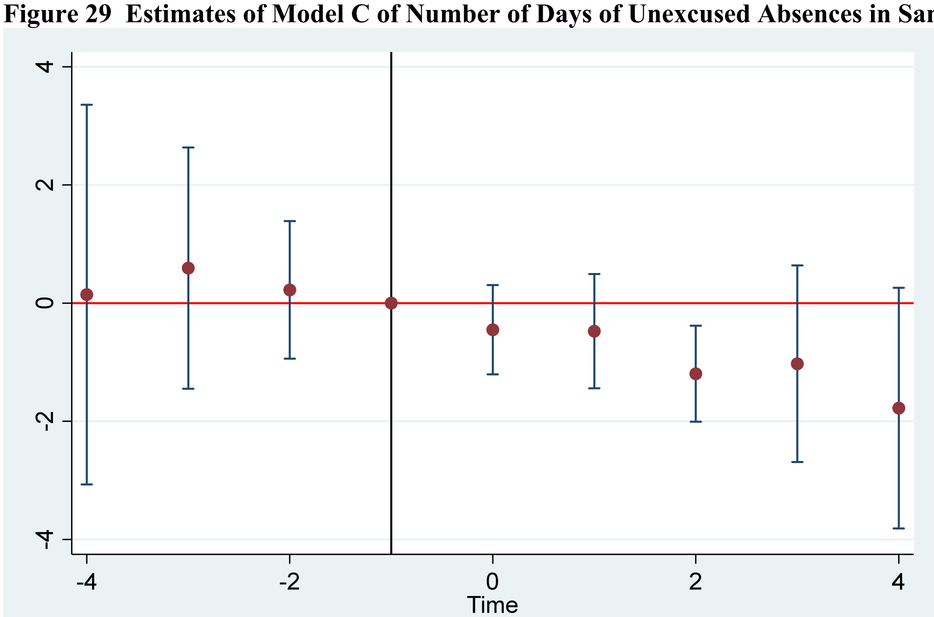
NOTE: Estimates are displayed in Table 31. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors’ analysis of data from the Community Data System combined with the CISK database.

Figure 28 Estimates of Model B of Number of Days of Unexcused Absences in Sample B



NOTE: Estimates are displayed in Table 31. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Figure 29 Estimates of Model C of Number of Days of Unexcused Absences in Sample B



NOTE: Estimates are displayed in Table 31. Dots designate point estimates with 95% confidence interval shown.
SOURCE: Authors' analysis of data from the Community Data System combined with the CISK database.

Attendance Rates and Student Demographics

Since attendance rates make up the only student outcome variable that is statistically significantly affected by the treatment, it is worth digging deeper into the relationship between attendance rates and CISK Tier II or Tier III services. This section examines attendance rates for various combinations of student demographic characteristics. Combinations of four student demographic variables are considered here: 1) eligibility for free and reduced-price lunch programs, 2) black, 3) female, and 4) Hispanic. Here we use a simple regression analysis by group, with elementary student attendance rates as the dependent variable and the interaction of four student characteristics as the independent variables. After regressing the dependent variable on interactions, we then use the "margins" function to switch all the coefficients relative to the omitted variable to average amounts. For each group, the average attendance rates are displayed in Table 32 by the various combinations of the four independent variables.

Table 32 Attendance Rates by Each Combination of Variables for Comparison and Treatment Groups in Sample B

	A	B	C	D
	Comparison	Treatment	Difference	*t*-stat
FRPL#black#female#Hispanic	Mean	Mean		
0000	0.9541	0.9412	−0.013	−1.12
0001	0.9478	0.9671	0.019	0.67
0010	0.9569	0.9613	0.004	0.40
0011	0.9479	0.9748	0.027	1.34
0100	0.9429	0.9443	0.001	0.07
0110	0.9437	0.9542	0.011	0.60
1000	0.9130	0.9362	0.023	3.46
1001	0.9318	0.9476	0.016	2.12
1010	0.9215	0.9349	0.013	1.85
1011	0.9357	0.9394	0.004	0.50
1100	0.9064	0.9321	0.026	5.58
1110	0.9156	0.9396	0.024	5.35

NOTE: The estimates are derived from the *ttest* procedure in Stata, with "if" statements restricting the sample for the combinations of the demographic variables listed in the table. Column A includes the mean of the comparison group for the various combination of demographic variables, Column B includes the mean of the treatment group, Column C is the difference between the two means, and Column includes the *t*-statistic for the differences.
SOURCE: Authors' analysis of KPS and CISK data.

The attendance rates for all combinations of demographic characteristics are basically higher for the treatment group than for the comparison group. The average attendance rate is 1.3 percentage points higher for the comparison group than for the treatment group, but all the demographic combinations are higher for the treatment group than for the comparison group. For example, the attendance rate for blacks only is very similar for the comparison group and the treatment group. The attendance rate for blacks on free or reduced-price lunch is the lowest of all combinations (0.9321 for the treatment group and 0.9064 for the comparison group). The highest attendance rate in the treatment group is for Hispanic female students (0.9748), and the highest rate for the comparison group of Hispanic female students is 0.9479.

VI. STATISTICAL RELATIONSHIP BETWEEN ELEMENTARY ATTENDANCE RATES AND HIGH SCHOOL DROPOUTS AND GRADUATES

The stated purpose of Community in Schools and the local affiliate Community in Schools of Kalamazoo is to reduce "high school dropout rates by integrating preventive services available to the entire school with intensive, targeted, and sustained services for those students who display significant risk of dropping out." So far, because of the large percentage of elementary students who receive CISK services, we have concentrated our evaluation of CISK services on elementary students. The strong and statistically significant effect of CISK services on the attendance rates of elementary students leads us to consider the statistical relationship between attendance rates of elementary students and the number of high school dropouts and graduates.

We examine whether higher attendance rates in elementary school reduce dropout rates in high school within the same student cohorts. The premise behind this concept is that greater student engagement through greater attendance in elementary school is expected to lead to fewer dropouts in high school. Although the number of students is reduced when we follow the same students from elementary grades through high school—primarily because of the flux of students coming in and out of KPS during that time—we still have enough sample size to estimate the relationship.

Table 33 displays the number of dropouts and graduates from KPS high schools (grades 9 through 12), regardless of whether students attended KPS at any other time in their schooling. The sample includes all students who were in high school at any time during the four school years starting with 2014–2015. Both the number of dropouts and the number of graduates are included; the percentages are also included and based on the relevant denominator for each year. We see that the number of dropouts was stable throughout these school years, hovering in the 120s except for Year 2020, when the number dropped to 91. That year, marking the first year of the pandemic, was the same year in which the number of high school graduates broke the 700-student mark. The lowest number of graduates, as well as the lowest percentage relative to the number of twelfth graders, appeared at the beginning of the six-year period, in 2016. The numbers of dropouts and graduates were supplied by the school district.

Table 33 Number of High School Dropouts in Grades 9 through 12 during Years 2016–2021, by School Year

	2015–2016	2016–2017	2017–2018	2018–2019	2019–2020	2020–2021
No. of dropouts	110	113	113	105	88	121
No. of HS grads	613	623	668	684	708	607
No. in grades 9–12	4,052	4,178	4,153	4,089	4,071	3,881
No. of 12th graders	906	852	909	896	905	821
Dropouts/9–12 (%)	2.71	2.70	2.72	2.57	2.16	3.12
HS grads/12th (%)	67.66	73.12	73.49	76.34	78.23	73.93

SOURCE: Authors' analysis of KPS data.

Table 34 shows the number of students who were KPS fifth graders during the school years 2010–2011 through 2013–2014 and then went on to attend KPS high school grades from school year 2015–2016 through 2020–2021. The table follows four cohorts of students from fifth grade through high school and shows the number of dropouts out of the students who attended fifth grade between the years ending in 2011 and 2014 in KPS and also attended high school (grades 9 through 12) in KPS after the normal advancement within the five-year period. That is, a student advancing at the normal rate who attended fifth grade in the school year ending in 2011 would be in 9th grade in the school year ending in 2015. Because of the longitudinal nature of the data and students entering and leaving KPS each year, a data set that follows students over time will have fewer students than a data set that slices the population of students each year.[16]

Table 34 shows the elementary schools that the fifth graders attended and the number of students from each school that dropped out of high school out of those students who still attended KPS. For example, for those students in fifth grade during the school year 2010–2011, 787 continued in high school, of which 72 dropped out of school by the end of twelfth grade. The

Table 34 Number and Percentage of High School Dropouts from KPS Students Who Attended 5th Grade Between 2011 and 2014 and Attended KPS in Grades 9 through 12, by 5th Grade Elementary School

Cohort	2011			2012			2013			2014			Total		
School	No.	Total	%	No.	Total	%	No.	Total	%	No.	Total	%	No.	Total	%
Total	72	787	9.2	46	796	5.8	35	768	4.6	37	778	4.8	190	3,129	6.1
Arcadia	2	40	5.0	0	34	0.0	2	37	5.4	1	34	2.9	5	145	3.4
Prairie Ridge	5	63	7.9	4	58	6.9	3	52	5.8	5	73	6.9	17	246	6.9
Edison	5	40	12.5	2	33	6.1	4	53	7.6	6	38	15.8	17	164	10.4
Lincoln	6	50	12.0	7	56	12.5	3	39	7.7	3	43	7.0	19	188	10.1
Milwood	12	75	16.0	7	75	9.3	4	67	6.0	4	66	6.1	27	283	9.5
Northeastern	6	37	16.2	0	36	0.0	0	32	0.0	1	31	3.2	7	136	5.1
Northglade	0	25	0.0	1	32	3.1	0	24	0.0	0	25	0.0	1	106	0.1
Woods Lake	9	56	16.1	7	62	11.3	4	55	7.3	1	59	1.7	21	232	9.0
Parkwood-Upjohn	4	53	7.6	2	68	2.9	3	70	4.3	2	73	2.7	11	264	4.2
Spring Valley	5	53	9.4	3	46	6.5	0	39	0.0	1	49	2.0	9	187	4.8
Washington	5	47	10.6	3	40	7.5	2	44	4.6	6	41	14.6	16	172	9.3
King Westwood	6	102	5.9	4	109	3.7	2	101	2.0	1	103	1.0	13	415	3.1
Winchell	0	77	0.0	1	74	1.4	1	62	1.6	2	66	3.0	4	279	1.4
Woodward	5	39	12.8	5	54	9.3	4	52	7.7	4	44	9.1	18	189	9.0
El Sol	2	30	6.7	0	19	0.0	3	41	7.3	0	33	0.0	5	123	4.1

SOURCE: Authors' analysis of KPS data.

elementary school with the highest number of students who dropped out was Milwood Elementary, but Northeastern edged out Milwood with the highest percentage of dropouts, with 16.2 percent. The table shows that fifth grade cohorts after the 2010–2011 school year had both

[16] Comparing the numbers of students in Tables 33 and 34 is consistent with the previous statement. The four cohorts of fifth graders account for 4,006 students who have gone through four years of high school or have dropped out in ninth grade. If one adds the filter that these students must have attended ninth grade, the number of students from these four cohorts falls to 3,129.

lower numbers of dropouts and lower percentages. One reason for the decline in dropouts over time is the greater number of years a student in the 2011 cohort has in which to drop out compared to a student in the 2016 cohort. The way the sample is constructed, fifth graders in the 2011 cohort have seven years to drop out of school, fifth graders in the 2012 cohort have six years, fifth graders in the 2013 cohort have five years, and those in the 2014 cohort have four years.[17]

Table 35 Number and Percentage of High School Graduates from KPS Students Who Attended 5th Grade Between 2011 and 2014 and Went On to Attend KPS in Grades 9 through 12, by 5th Grade Elementary School

Cohort	2011			2012			2013			2014			Total		
School	No.	Total	%	No.	Total	%	No.	Total	%	No.	Total	%	No.	Total	%
Total	518	787	66	562	796	71	506	768	66	453	778	58	2039	3129	65
Arcadia	23	40	57	25	34	74	25	37	68	21	34	62	94	145	65
Prairie Ridge	39	63	62	50	58	86	37	52	71	46	73	63	172	246	70
Edison	19	40	48	15	33	45	30	53	57	13	38	34	77	164	47
Lincoln	25	50	50	29	56	52	18	39	46	18	43	42	90	188	48
Milwood	46	75	61	47	75	63	37	67	55	30	66	45	160	283	57
Northeastern	20	37	54	24	36	67	18	32	56	16	31	52	78	136	57
Northglade	20	25	80	29	32	91	15	24	63	15	25	60	79	106	75
Woods Lake	31	56	55	48	62	77	38	55	69	31	59	53	148	232	64
Parkwood-Upjohn	39	53	74	52	68	76	50	70	71	49	73	67	190	264	72
Spring Valley	32	53	60	32	46	70	25	39	64	24	49	49	113	187	60
Washington	22	47	47	15	40	38	18	44	41	15	41	37	70	172	41
King Westwood	88	102	86	97	109	89	89	101	88	84	103	82	358	415	86
Winchell	68	77	88	52	74	70	53	62	85	57	66	86	230	279	82
Woodward	23	39	59	33	54	61	28	52	54	16	44	36	100	189	53
El Sol	23	30	77	14	19	74	25	41	61	18	33	55	80	123	65

SOURCE: Authors' analysis of KPS data.

For those students who were in fifth grade in school year 2010–2011, made expected progress during middle school, and stayed in KPS through high school, Table 35 shows that 518 graduated sometime between school years 2017–2018 and 2020–2021. Similarly, 562 students who were in fifth grade in 2011–2012 graduated between the years of 2018–2019 and 2020–2021. As with the students who dropped out, students in the earlier cohorts have more time to

[17] It should be noted that the entry/withdrawal record supplied by KPS may not capture all dropouts in the sample. However, KPS assured us that most dropouts are included in these records.

graduate from high school and thus should have higher graduation percentages than those in the more recent cohorts. However, somewhat surprisingly, the percentage of high school graduates is higher for the 2012 fifth grade cohort than for the 2011 cohort. Even so, the percentages fall steadily after that. King Westwood Elementary has the highest number of graduates for the 2010–2011 school year and the largest number of students in the sample. Except for one year, it has the highest percentage of graduates, always greater than 80 percent.

Before moving to the regression analysis of the relationship between attendance rates in fifth grade and the number of dropouts and graduates in high school, we examine the difference in elementary attendance rates between those who drop out and those who do not, as well as between those who graduate from high school and those who do not. Table 36 shows elementary attendance rates to be 2.5 percentage points lower for those students who eventually drop out compared with those who do not. High school graduation rates demonstrate the same relative difference. Both differences are statistically significant at the 95 percent confidence level. As before, these percentage-point differences reflect all four fifth-grade cohorts combined that attended KPS during high school.

Table 36 Means of Elementary Attendance Rates, High School Graduation Rates, and Dropout Rates for 5th-Grade Cohorts between 2011 and 2014 Who Attended KPS in Grades 9–12 (%)

	Attendance rate	HS graduation rate	Dropout rate
Dropout	92.6 (diff=0.025)	5.8	
Not a dropout	95.1 (*t*=7.16)	69.0	
HS grad	96.0 (diff=-0.030)		0.5
Not a HS grad	93.0 (*t*=-17.5)		16.4

NOTE: Sample created with ninthAttendee==1. Even though a student may have dropped out in 9th grade, that student can still find his or her way back to school at some later date and even graduate. Some students who dropped out of high school returned to graduate from high school in later years. The table also includes those students who did not graduate from high school and who did not dropout in ninth grade but left the district for other reasons.
SOURCE: Authors' analysis of KPS data.

For the regression results, we find a strong statistical relationship between attendance rates of fifth graders and the number of high school dropouts, as displayed in Table 37. A simple regression yields an estimated coefficient of −0.643, as shown under Model A, and the coefficient is statistically significant. Model B, which adds selected student demographic variables, reduces the estimates slightly but increases the adjusted *R*-squared substantially. Models C and D have the same structure as Models A and B except that they use logit instead of OLS. The results are similar. Models E through H are the same basic models as Models A through D for dropouts, but the models have replaced dropouts as the dependent variable with an indicator for high school graduation. The relationship between models with high school graduates mimics the models for dropouts.

Table 37 Statistical Relationship between Dropouts, Graduates, and Elementary Attendance Rates

Dependent Variable:	Number of high school dropouts				Number of high school graduates			
	Model A	Model B	Model C	Model D	Model E	Model F	Model G	Model H
Elementary attendance rate	−0.643	−0.569	−7.50	−6.45	3.02	2.54	15.08	13.08
	(−7.16)	(−6.22)	(−6.50)	(−5.37)	(17.54)	(14.89)	(15.27)	(12.86)
Constant	0.671	0.544	4.31	3.69	−2.22	−1.68	−13.67	−11.39
	(7.87)	(4.93)	(4.00)	(3.21)	(−13.54)	(−8.15)	(−14.58)	(−9.66)
Adj. R2	0.016	0.024	0.026	0.046	0.089	0.154	0.072	0.141
Obs	3,129	3,100	3,129	2,789	3,129	3,100	3,129	3,093
OLS	X	X			X	X		
Race, gender, FRPL		X		X		X		X
logit			X	X			X	X

NOTE: Dependent variables are indicators of dropouts or graduates; independent variables are interacted. Coefficients of interaction terms are not shown in the table, but variables are listed below the line to indicate the model used. Ninth grade attendance equals 1.
Source: Authors' analysis of KPS data.

Therefore, the relationship between attendance rates of elementary students and high school dropouts is strong and statistically significant, no matter which model is chosen. Using the estimates of Model B suggests that a 1 percentage point increase in attendance rates leads to nearly a 0.53 percent reduction in the number of dropouts. Similarly, using Model F a 1 percentage point increase in elementary attendance rates results in a 2.2 percent increase in the number of graduates. Table 37 illustrates the small but statistically significant difference in attendance rates for dropouts and graduates.

VII. DISCUSSION

The purpose of this evaluation is to determine whether CISK services assist students participating in CISK to perform better in school. The overarching purpose of CIS is to reduce the number of dropouts and increase the graduation rates of all students it touches. More specifically, CIS attempts to boost attendance of students in elementary school so that these interventions can be successful in stemming high school dropouts and increasing graduation rates.

The most significant finding in this report is that CISK interventions causally increase attendance rates for elementary students, and that this increase in attendance rates, in turn, reduces the number of dropouts and increases the number of graduates when those fifth graders move through high school. We have direct evidence that CISK services labeled as Tier II or Tier III activities improve the attendance rates of first through fifth graders. Following fifth graders through to ninth grade and then on to twelfth grade (for those who remain in school) provides an opportunity to examine the relationship between fifth grade attendance rates and the number of high school dropouts. In the same way, the longitudinal nature of the data enables us to examine the statistical relationship between fifth-grade attendance rates and the number of twelfth-grade graduates. We find both relationships to be statistically significant, despite the small difference in attendance rates between fifth graders who would go on to drop out in high school and those who would not. The difference in attendance rates between those who graduated from high school in twelfth grade and those who did not is also small.

The increase in elementary attendance rates because of CISK services is consistent with the stated purpose of CIS. This study provides a direct link between elementary attendance rates and statistics on high school dropouts. Policymakers have made a considerable number of a priori statements about how elementary attendance can reduce dropouts, but there is little in the CIS literature that tests this relationship. The results of this study demonstrate the strong relationship between elementary attendance rates and the number of high school dropouts. The same can be said for the effect of elementary attendance rates on the number of high school graduates.

Unfortunately, we did not find that CISK services reduced student behavioral issues for elementary students or for middle schoolers (reported in Appendix A), as measured by the number of unexcused absences. Measures other than unexcused absences may be used to proxy behavioral issues, including expulsions and encounters with law enforcement. However, we used unexcused absences because of its larger number of occurrences than expulsions and scuffles with law enforcement, as well as the lack of reliable data for encounters with law enforcement. Before ruling out any relationship that is statistically significant, we might want to find reliable data for these additional measures.

The MDRC evaluation of CIS services found no statistically significant relationships between such services and either traditional or nontraditional student outcomes. In contrast, our own results found that CISK Tier II or Tier III services had a direct effect on elementary attendance rates. We also found a strong statistically significant relationship between elementary attendance rates and both a reduction in the number of high school dropouts and an increase in the number of high school graduates. We did not, however, find a similar impact of CISK services on the attendance rates of middle schoolers. Nor did we find a consistent reduction in the number of unexcused absences, the measure we used to proxy behavioral issues.

There are several reasons why the results of the two evaluations may differ. As stated earlier, the two evaluations use different methodologies in constructing the treatment and comparison samples. MDRC constructs these groups using a random assignment approach. We exploit differences in the timing of the receipt of CISK services. The MDRC methodology is considered a better approach to ensure that the effects are exogenous to student outcomes, but there is no way of knowing how much better their approach is, or at what cost. We mentioned earlier the ethical issues that arise when individuals are denied beneficial services, which is what random assignment exacts on participants. We also noted the financial cost of using random assignment in terms of data collection, which is one reason why the MDRC evaluation stated it was not able to follow students through high school. In addition, MDRC did not include KPS, or any district in Michigan, in its evaluation. With our longitudinal data series, we *can* follow cohorts of students through high school, which is beneficial when linking the effects of CIS services to high school dropouts and graduates. That is why we were able to determine that elementary school attendance rates are strongly related to the number of high school dropouts and graduates for the same cohorts.

The effect of CISK services on elementary attendance rates is robust with respect to the various models used. All models yielded positive results—and for many models the results were statistically significant. The effects of CISK services on other student outcomes were far less consistent. Only attendance rates for elementary students passed the parallel-trend assumption test for all models. Some models showed an increase in math and reading test value-added for elementary students, but others did not. We considered a relationship between CISK services and a student outcome to be legitimate only when all models yielded results in the same expected direction, the parallel-trend assumption passed, and most estimates of lags were statistically significant.

Therefore, in our evaluation of CISK services, we conclude that CISK Tier II or Tier III services increase elementary attendance rates, and that this increase reduces the number of high school dropouts while it increases the number of high school graduates. The conclusion is consistent with the stated intent of CIS of targeting attendance in elementary schools to reduce dropouts in high school.

Appendix A

Effects of CISK Services on Middle School Student Outcomes

This appendix examines the effects of Tier II or Tier III services on KPS students in middle school. Under the KPS system, middle school students attend grades six through eight, and about 4.0 percent of KPS students in those three grades receive Tier II or Tier III CISK services. For middle school students, we look at the effect of CISK services on the same four outcome measures—the value added of NWEA tests in 1) math and 2) reading, 3) attendance rates, and 4) the number of unexcused absences—that are used for elementary students. However, we consider examining two models of the event study methodology, instead of examining differences in results from DID and panel event-study methodologies and comparing the two. We look at the first two models used for elementary school students; Model A includes only the outcome variable and the intervention variable, whereas Model B includes those two variables plus a year and building effect. Because we use the same number of leads and lags as for the five elementary grades considered, there is some overlap with elementary and high school grades. In addition, we use a sample of KPS middle schoolers that is analogous to Sample B for elementary students. Students in all middle-school grades are included in the data file each year.

As shown in Table A.2 and in Figures A.1–A.4, the results for middle schoolers are much like those for elementary students, except for attendance rates. This may be consistent with CIS philosophy. As stated at the beginning of this study, CIS targets attendance rates for elementary students and behavioral issues for middle schoolers. We found in previous sections of the report that CISK services increase attendance rates for elementary students. However, that is not the case for middle schoolers. None of the relevant coefficients for middle schoolers are statistically significant in either model. Using days of unexcused absences as a measure of behavioral issues, we did find some evidence that CISK services affect middle schoolers. We would expect CISK to reduce the number of unexcused absences, but the one coefficient that was statistically significant indicated more rather than fewer days of unexcused absences.

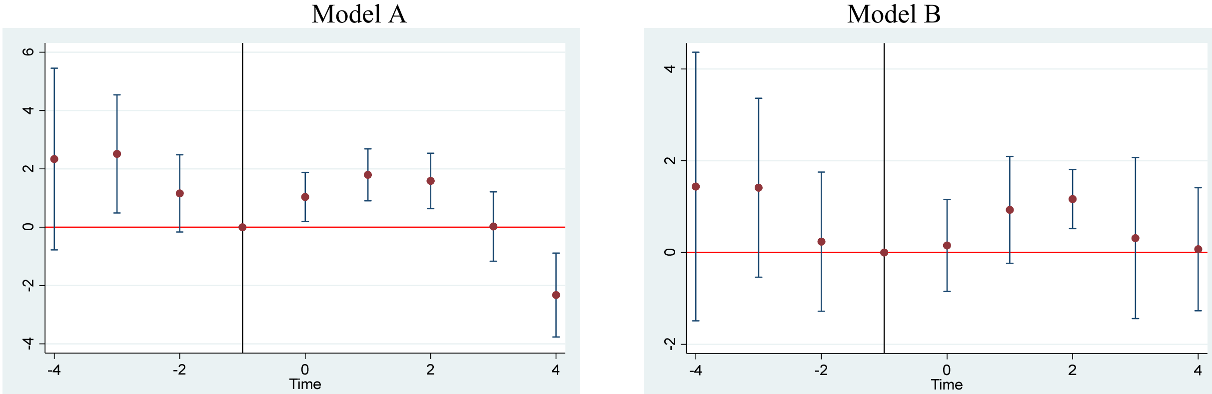
Table A.1 Estimated Effect of CISK Activities on Student Outcomes of Middle Schoolers for Larger Sample

	Model A				Model B			
	NWEA read VA	NWEA math VA	Attendance rate	Unexcused absences	NWEA read VA	NWEA math VA	Attendance rate	Unexcused absences
Lead 4	2.34	0.722	0.034	-2.67	1.44	−0.052	0.026	−1.60
	(1.47)	(0.50)	(2.27)	(−1.93)	(1.01)	(−0.04)	(3.11)	(−3.42)
Lead 3	2.51	−1.06	0.023	−1.40	1.41	−1.46	0.018	−0.893
	(2.43)	(−1.12)	(2.33)	(−1.54)	(1.48)	(−1.67)	(1.69)	(−1.32)
Lead 2	1.16	−0.450	0.014	0.055	0.237	0.008	0.009	0.456
	(1.72)	(0.73)	(2.24)	(0.09)	(0.32)	(0.01)	(1.52)	(1.28)
Lag 0	1.04	0.883	0.015	0.759	0.153	0.409	0.019	0.809
	(2.41)	(2.24)	(3.55)	(1.97)	(0.31)	(0.91)	(5.95)	(1.58)
Lag 1	1.80	0.714	0.010	0.200	0.929	0.343	0.018	0.054
	(3.95)	(1.71)	(2.40)	(0.49)	(1.63)	(0.89)	(5.03)	(0.16)
Lag 2	1.59	−0.043	0.003	1.18	1.16	−0.500	0.017	0.787
	(3.27)	(−0.10)	(0.65)	(2.68)	(3.70)	(−1.21)	(4.45)	(1.83)
Lag 3	0.026	−0.851	0.006	0.069	0.313	−0.712	0.033	−0.190
	(0.04)	(−1.56)	(1.01)	(0.13)	(0.37)	(−1.56)	(4.51)	(−0.42)
Lag 4	−2.32	−2.28	−0.017	1.50	0.071	−0.909	0.011	1.66
	(−3.17)	(−3.41)	(−2.35)	(2.27)	(0.11)	(−1.64)	(1.36)	(2.16)
Constant	7.22	9.98	0.911	7.45	11.03	12.72	0.931	7.60
	(89.76)	(135.7)	(1,320.29)	(116.31)	(33.11)	(28.48)	(277.76)	(48.57)
R2	0.003	0.001	0.001	0.0004	0.047	0.116	0.115	0.100
Obs	19,241	19,802	33,428	34,617	19,241	19,802	33,428	34,617

NOTE: Model A includes the dependent variable and the intervention variable; Model B includes the year effect and the building effect along with a cluster estimation of standard errors. *T*-statistics are found in the parentheses.
SOURCE: Authors' analysis of data from the Community Data System, which combines files from KPS and CISK.

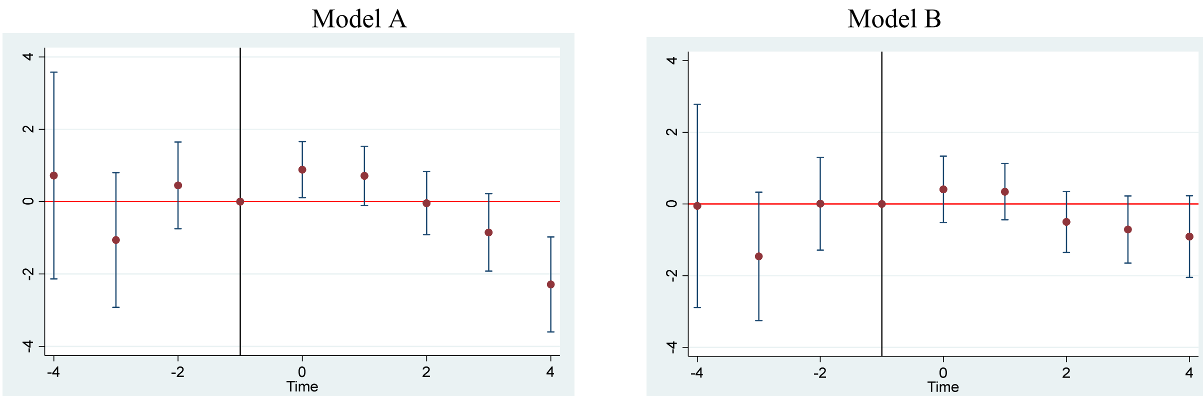
Figure A.1 Estimates of CISK Services on the Value-Added of NWEA Reading Test Taken by Middle Schoolers

Model A | Model B



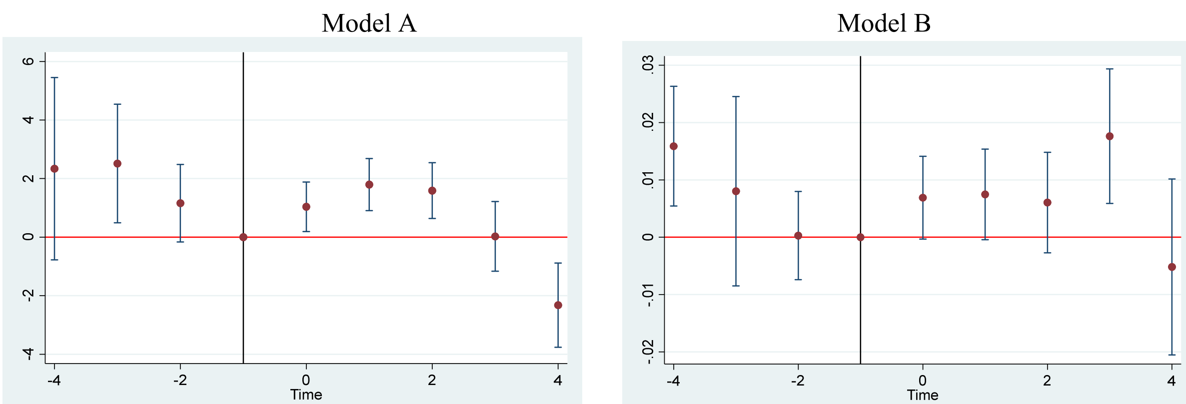
Source: Authors' analysis of KPS and CISK data.

Figure A.2 Estimates of CISK Services on the Value-Added of NWEA Math Test Taken by Middle Schoolers



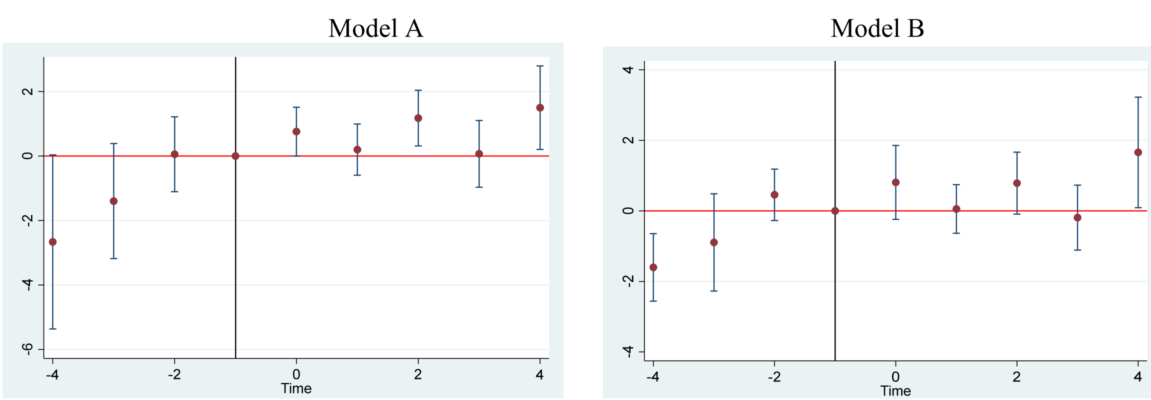
Source: Authors' analysis of KPS and CISK data.

Figure A.3 Estimates of CISK Services on Attendance Rates of Middle Schoolers



Source: Authors' analysis of KPS and CISK data.

Figure A.4 Estimates of CISK Services on the Number of Days of Unexcused Absences of Middle Schoolers



Source: Authors' analysis of KPS and CISK data.

References

Angrist, Joshua D., and Jorn-Steffen Pischke, *Mostly Harmless Econometrics: An Empiricist's Companion*, Princeton University Press, 2009.

Callaway, Brantly, and Pedro H. C. Sant'Anna. 2021. "Difference-in-Differences with Multiple Time Periods." *Journal of Econometrics* 225(2): 200–230.

Chetty, Raj, John N. Friedman, and Jonah E. Rockoff. 2014a. "Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates." *American Economic Review* 104(9): 2593--2632.

———. 2014b. "Measuring the Impacts of Teachers II: Teacher Value-Added and Student Outcomes in Adulthood." *American Economic Review* 104(9): 2633–2679.

Child Trends. 2014. *Making the Grade: Assessing the Evidence for Integrated Student Supports*. Publication No. 2014-07. Bethesda, MD: Child Trends.

Clarke, Damian, and Kathya Tapia Schythe. 2020. "Implementing the Panel Event Study." IZA Discussion Paper No. 13524. Bonn, Germany: Institute of Labor Economics.

Communities in Schools. 2022. *Our Mission; Our History*. Arlington, VA: Communities in Schools. https://www.communitiesinschools.org/about-us/mission-and-history/ (accessed April 20, 2022).

Dee, Thomas, Elizabeth Huffaker, Cheryl Phillips, and Eric Sagara. 2021. "The Revealed Preferences for School Reopening: Evidence from Public-School Disenrollment." NBER Working Paper No. 29156. Cambridge, MA: National Bureau of Economic Research.

Donald, Stephen G., and Kevin Lang. 2007. "Inference with Difference-in-Differences and Other Panel Data." *Review of Economics and Statistics* 89(2): 221–233.

Figlio, David N. 2015. *Experimental Evidence of the Effects of the Communities in Schools of Chicago Partnership Program on Student Achievement*. Communities in Schools of Chicago report. Chicago: Communities in Schools of Chicago.

Freyaldenhoven, Simon, Christian Hansen, Jorge Pérez Pérez, and Jesse M. Shapiro. 2021. "Visualization, Identification, and Estimation in the Linear Panel Event-Study Design." NBER Working Paper No. 29170. Cambridge, MA: National Bureau of Economic Research.

Freyaldenhoven, Simon, Christian Hansen, and Jesse M. Shapiro. 2019. "Pre-Event Trends in the Panel Event-Study Design. *American Economic Review* 109(9): 3307–3338.

Goodman, Joshua. 2019. "The Labor of Division: Returns to Compulsory High School Math Coursework." *Journal of Labor Economics* 37(4): 1141–1182.

Goodman-Bacon, Andrew. 2018. "Difference-in-Differences with Variation in Treatment Timing." NBER Working Paper No. 25018. Cambridge, MA: National Bureau of Economic Research.

ICF International. 2010. *Communities in Schools National Evaluation: Five Year Summary Report*. Fairfax, VA: ICF International.

Parise, Leigh M., William Corrin, Kelly Granito, Zeest Haider, Marie-Andrée Somers, and Oscar Cerna. 2017. *Two Years of Case Management: Final Findings from the Communities in Schools Random Assignment Evaluation*. New York: MDRC.

Somers, Marie-Andrée, and Zeest Haider. 2017. "Using Integrated Student Supports to Keep Kids in School: A Quasi-Experimental Evaluation of Communities in Schools." New York: MDRC

Stevenson, Betsey, and Justin Wolfers. 2006. "Bargaining in the Shadow of the Law: Divorce Laws and Family Distress." *Quarterly Journal of Economics* 121(1): 267–288.

Thum, Yeow Meng, and Megan Kuhfeld. 2020. *NWEA 2020 MAP Growth Achievement Status and Growth Norms for Students and Schools*. NWEA Research Report. Portland, OR: NWEA.

Warren, Mark R. 2005. "Communities and Schools: A New View of Urban Education Reform." *Harvard Educational Review* 75(2): 133–173.