When Resources Meet Relationships: The Returns to Personalized Supports for Low-Income Students*

Benjamin Goldman
Cornell, NBER, & Opportunity Insights

Jamie Gracie
Harvard University

DO NOT CITE OR CIRCULATE

August 2026

Abstract

Children from low-income families face barriers outside the classroom that impede learning and limit opportunities for economic mobility. We study Communities In Schools (CIS), a program that places coordinators in high-poverty schools to connect struggling students with personalized support. Coordinators match students to resources that address academic, behavioral, and basic needs—such as tutoring, mental health counseling, and housing stabilization. CIS is the largest program of its kind in the US, reaching 2 million students annually. Using the staggered rollout of CIS, we find that the program boosts test scores for struggling students. These gains persist, raising high-school completion and adult earnings. Improvements in test scores explain only about half of CIS's impact on high school graduation; the remainder reflects progress on non-cognitive outcomes such as attendance and behavior. Consistent with CIS's emphasis on personalization, coordinators tailor services to distinct student needs, yet students with different needs experience comparable long-run gains. Embedding a trusted adult with access to a network of programs and services in schools may offer a scalable way to improve economic mobility.

---

*The conclusions of this research do not necessarily reflect the opinion or official position of the Texas Education Research Center, the Texas Education Agency, the Texas Higher Education Coordinating Board, the Texas Workforce Commission, or the State of Texas. Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (Project 751987: CBDRB-FY23-CES014-028). We thank David Card, Raj Chetty, Sahil Chinoy, Heather Clawson, Kaveh Danesh, John Friedman, Claudia Goldin, Nathan Hendren, Joshua Hyman, Tom Kane, Larry Katz, Pat Kline, Martin Koenen, Kevin Leary, Matthew Notowodigdo, Amanda Pallais, Lynne Sacks, Doug Staiger, and Rob Watson for helpful feedback. Additionally, we extend our thanks to seminar participants at Aalto University, UC Berkeley, University of Connecticut, Cornell, Georgetown, IIES, Harvard, and Oxford. This research was funded by Opportunity Insights, The EdRedesign Lab at Harvard Graduate School of Education, the Stone Program at Harvard Kennedy School, the National Science Foundation (Graduate Research Fellowship, grant DGE-2140743), and Cornell University.

I Introduction

In the United States, children from low-income families grow up to earn far less than children from high-income families (Chetty et al., 2014; Solon, 1992). The gap opens long before they enter the labor market: they score lower on standardized tests, graduate from high school at lower rates, and go to college less often (Black, Devereux and Salvanes, 2005; Chetty et al., 2020). Given the evidence on the returns to education, improving educational outcomes is a promising way to narrow the earnings gap (Bhuller, Mogstad and Salvanes, 2017; Card, 1999; Zimmerman, 2014).

One approach is to improve the schools that low-income students attend. Efforts to raise teacher quality, reduce class sizes, expand after-school and tutoring programs, equalize school funding, and increase access to effective charter schools have all been shown to produce meaningful gains (Angrist, Pathak and Walters, 2013; Chetty, Friedman and Rockoff, 2014; Jackson, Johnson and Persico, 2016; Krueger, 1999; Nickow, Oreopoulos and Quan, 2020). Still, many low-income students face barriers that hinder their academic progress: they are more likely to miss school, face logistical or informational hurdles to participating in enrichment programs, and contend with unstable home environments that disrupt learning (Malkus, 2024; Robinson, Bisht and Loeb, 2025; Walters, 2018). Interventions that address academic and non-academic barriers together may improve outcomes both by complementing existing resources and by meeting students' needs where such resources are limited.

How can we help students in high-poverty schools clear the barriers that keep them from learning? Recent randomized trials outside of education point to a promising answer: pair the usual resources with personalized support from a trusted adult who coordinates services and advocates for the person receiving them (e.g., Bergman et al., 2024; Weiss et al., 2019). Most of the existing studies have focused on adults. Extending the same support to children in school may improve their outcomes directly and raise the returns to the school's other investments.

In this paper, we study Communities In Schools (CIS), the largest provider of integrated student supports in the United States, serving 2 million students across more than 3,500 schools each year—more than twice Head Start's annual enrollment (Office of Head Start, 2025). CIS places a full-time "site coordinator" in high-poverty schools to identify struggling students and connect them to personalized supports that address both academic and non-academic barriers to learning. A coordinator might arrange tutoring for one student and mental-health counseling for another, tailoring services to each child's circumstances. Rather than assembling these services alone, the coordinator draws on a network of community and government partners that the local CIS affiliate has built, pairing a consistent, trusted adult with access to a broad set of resources. We evaluate the program's effects on short-run academic outcomes (test scores, absences, and suspensions) and long-run outcomes (high school graduation, college attendance, and earnings).

Our data come from federal tax records covering nearly the full U.S. population and from K–12 education records in Texas, where nearly half of CIS schools are located.[1]

Our analysis yields five main findings. First, using the staggered rollout of CIS in Texas, we show that the program raises test scores among students who struggle academically, have poor attendance, or face disciplinary problems. Second, these gains translate into lasting benefits: higher rates of high school graduation and, ultimately, higher adult earnings, as shown in linked Census and IRS records. Third, these long-run effects are not driven by test scores alone: only about half of the impact on high school graduation is explained by short-run gains in test scores, with improvements in non-cognitive outcomes such as attendance and discipline accounting for much of the remainder. Fourth, personalization is central to the CIS model: coordinators give students with different needs different services, and these students improve on different short-run margins yet achieve similar long-run gains. Finally, we show that CIS's long-run earnings gains are large relative to the program's direct cost, and that targeted expansions could meaningfully raise high school graduation rates among at-risk students. Together, these results show that providing individualized supports in high-poverty schools can both improve student achievement in the short run and translate into durable gains in education and earnings.

CIS operates through a national network of independent local nonprofit "affiliates" that implement a common model of integrated student supports. Each affiliate maintains partnerships with community organizations and government agencies, giving schools access to academic and non-academic services well beyond what they could provide on their own (CIS, 2025*b*). Site coordinators connect students to these resources while remaining employees of the local affiliate rather than the school district. As a result, adding CIS to a school requires no changes to its staffing, budget, or governance, making the program straightforward to scale. Expansion is paced by affiliate capacity: directors add schools with persistent need as resources allow, rather than in response to demand from individual schools or year-to-year fluctuations in local need (CIS, 2025*c*). Consistent with the putative role of supply-side constraints in governing expansion, schools that adopt CIS do not appear to have been on an improving trajectory before entry.

Within a CIS school, all students can interact with the coordinator, but the most intensive support comes through "case management." Teachers, counselors, and administrators refer students with academic, behavioral, attendance, basic-needs, or social-emotional concerns, who then meet regularly with the coordinator to get the resources they need. In our Texas sample, 12.4% of students receive case management in a given year, and the median case-managed student meets the coordinator 29 times—nearly once a week during the school year. Coordinators

[1] Texas is unique in providing substantial state funding for CIS. In 2022, the state allocated more than $38 million to CIS programs, drawing on both the state budget and federal Temporary Assistance for Needy Families funds (Texas Education Agency, 2022).

connect these students to academic and non-academic supports. Academic services include tutoring, test preparation, and help with study habits and course planning. Non-academic services cover social-emotional needs (small-group counseling, goal-setting, and mentoring), basic needs (food, transportation, clothing, and emergency housing), and physical and mental health care (McShane, 2019; CIS, 2025*a*). Services are delivered directly by coordinators, by school-based staff, or through public programs and community partners, with the mix varying across affiliates. Coordinators also run school-wide initiatives—attendance campaigns, climate-building activities, and family-engagement events—that strengthen attendance culture and the broader learning environment for all students.

We begin by showing that case management assignments can be predicted with a relatively high degree of accuracy using data from before the CIS program begins. This allows us to construct a natural control group of students in non-CIS schools who would likely have received case management, while avoiding bias from conditioning on case management itself (e.g., a parental divorce may trigger case management and independently lower test scores). Our baseline estimates focus on Texas middle schools with CIS programs, both because standardized test coverage is substantially higher in middle school and because we use elementary school test scores, attendance, suspensions, and student demographics (e.g., free and reduced-price lunch, or FRPL status) to predict case management assignments for middle-school students. A gradient-boosted decision-tree algorithm produces strong out-of-sample predictions: 27.3% of students in the top quartile of predicted case management actually receive services, compared to 3.6% in the bottom quartile. Likely case managed students are more likely to be FRPL-eligible, have disciplinary records, be chronically absent, and score lower on standardized tests before CIS exposure. Models restricted to a single variable, such as test scores, perform substantially worse. Throughout, we refer to students in the top and bottom quartiles of predicted case management as *high-risk* and *low-risk*, reflecting case management's stated goal of serving students at risk of dropping out in the future.

In our baseline specification, we estimate an event study matching each CIS school to a demographically and academically similar "never-treated" control school. The match is constructed using only data from more than five years before CIS begins, excluding all observations in the event-study window and allowing us to assess parallel trends using data not used in the matching procedure. We estimate that three years of exposure to CIS increases math scores for high-risk students—relative to their high-risk counterparts in control schools—by 0.18 standard deviations (SD), with a standard error (SE) of 0.04. For low-risk students, the corresponding estimate is a statistically insignificant 0.02 SD (SE = 0.04), indicating that CIS's academic benefits are concentrated among students with prior academic, attendance, or discipline challenges. We do not interpret this heterogeneity as arising solely from case management. CIS can also raise

outcomes through school-wide initiatives and through spillovers within and across groups of students. These channels operate separately from case management, and we do not try to separate their effects from it. We instead read the pattern as evidence that CIS is most effective for struggling students and less so for those already on track.

We use several strategies to test the robustness of our baseline findings on academic outcomes. Results are similar when we draw comparison groups from outside the matched control, including "later treated" CIS schools. We also relax the parallel trends assumption by allowing trends to vary at the race by sex level. As another check, we construct a control group of schools within the same commuting zone as the CIS school, ensuring that our estimates are not driven by city-wide changes coinciding with CIS entry. We find no evidence of changes in student demographics, teacher quality, or teacher tenure around the time CIS enters a school. We also show that results are robust to assigning children to a middle school based on their first elementary school, ruling out the possibility that higher-ability students sort into CIS schools after program entry.

Beyond improving students' performance in school, CIS aims to equip them with the resources they need to succeed in adulthood. Keeping students on track to a diploma is central to this goal, and it is a pressing challenge for the students CIS serves: roughly two in five high-risk students in our sample do not graduate from high school.

Consistent with this goal, three years of CIS exposure in middle school raises high school graduation for high-risk students by 3.4 percentage points (pp), or 5.2% (SE = 1.2 pp), closing 14% of the initial gap between high- and low-risk students. These gains extend beyond high school: although estimated imprecisely, we also find a 2.9 pp (9.2%; SE = 2.0 pp) increase in college attendance, driven entirely by two-year college enrollment—the most common postsecondary path among high-risk students who attend college. For low-risk students, we find no effect on high school graduation, consistent with our short-run results.

We complement our analysis of academic outcomes in Texas with linked Census and IRS records to estimate the effect of CIS on earnings. Because most Texas programs began after 2010, many exposed students are still too young to have reliable earnings measures. To address this, we use the staggered rollout of CIS programs outside Texas to estimate event studies. We find that three years of exposure raises earnings at age 27 by 1.5 percentile ranks (4.3%; SE = 0.5), corresponding to approximately $\$1,550$. Since we cannot distinguish high- and low-risk students in this sample, this estimate reflects the average effect across the full school population.[2] The gains come mainly from moving students out of the bottom quartile of the income distribution and into the second, with no effect at higher income levels. Across both school and adult outcomes, CIS

[2]We find similarly sized effects in Texas. Because our Texas panel is too short to observe age-27 earnings directly for a sufficient number of cohorts, we predict them from college attendance and earnings measured at earlier ages (between ages 18 and 22). See Appendix Figure A.9 for details.

works by keeping students out of the left tail.

In the final section, we ask which of CIS's short-run gains explain its long-run effect on high school graduation, and whether short-run outcomes can serve as reliable surrogates when long-run data are not yet available. We find that test scores alone explain only about half of CIS's impact on high school graduation, so focusing solely on cognitive achievement would understate the program's long-run effect. In contrast, a prediction based on the full set of short-run outcomes closely matches the observed long-run impact: CIS raises high school graduation by 3.1 pp (SE = 1.0 pp), while the effect predicted from test scores, attendance, and disciplinary incidents is 2.4 pp (SE = 0.6 pp). Engagement and behavior, not achievement alone, carry much of CIS's long-run effect. The prediction is similarly accurate for both high- and low-risk students, suggesting that rich administrative measures of contemporaneous student progress can provide a useful window into long-run program effectiveness.

We next turn to *personalization*, a cornerstone of the CIS model. Site coordinators tailor services to each student's needs. Tailoring can give each student the mix that helps her most, or it can leave some students worse off than giving everyone the same mix would have. We ask whether tailoring leaves any group systematically behind. Using detailed data on services received in case management, we show that students with academic needs receive more academic and cognitive support and show the largest short-run test score gains, while students with primarily non-academic needs show the largest gains in attendance and behavior. Each group improves most along the dimension aligned with its needs. Yet the effects of CIS on high school graduation are nearly identical for both groups, with no evidence that any single service mix is uniformly more effective. Although this is not a formal test of optimal service assignment, the evidence is consistent with coordinators matching services to students' needs in ways that produce comparable long-run gains across different types of students.

Our paper contributes to the literature on improving outcomes in high-poverty schools by evaluating a coordinator-led model that connects struggling students to academic and non-academic supports. Prior studies have largely focused on expanding resources, improving curricula, or introducing new programs—high-dosage tutoring, teacher quality initiatives, behavior interventions, class-size reductions, "no excuses" charter schools, and school finance reforms (e.g., Alan and Kubilay, 2025; Guryan et al., 2023; Chetty, Friedman and Rockoff, 2014; Angrist et al., 2019; Angrist, Pathak and Walters, 2013; Jackson, Johnson and Persico, 2016). We provide large-scale evidence that a comprehensive individualized support model raises test scores, attendance, and behavior for high-risk students, with gains that persist through high school graduation and early-career earnings.[3] These long-run earnings gains are large relative to the

[3]Prior work on CIS sits outside of the economics literature and is predominantly descriptive, typically focusing on a single affiliate or district group, and generally finds positive but imprecise effects on dropout and graduation

program's direct cost. Our results also speak to a growing literature on personalization in education. In contrast to prior work—which has focused on instructional or technology-driven approaches (e.g., Kim et al., 2025; Muralidharan, Singh and Ganimian, 2019)—CIS personalizes access to academic and non-academic supports outside the classroom. Taken together with evidence from similar interventions in other settings (Dee et al., 2026; Lavecchia, Oreopoulos and Brown, 2020), our results show that integrated student supports deliver durable gains for low-income students.

A related literature finds that school counselors can raise educational attainment and college-going among low-income students, largely by providing information about postsecondary options and assistance with applications, course selection, and financial aid (Mulhern, 2023; Sparks and Mulhern, 2026; Carrell and Hoekstra, 2014). A CIS site coordinator plays a different role, drawing on the local CIS affiliate's network of community partners to address a broad range of non-academic barriers: basic needs such as food and housing, physical and mental health, and social-emotional development. Coordinators work alongside the school's own counselors and teachers, who often refer students to them. They also carry a smaller caseload, serving roughly a third as many students as a typical school counselor (Parise et al., 2017; American School Counselor Association, 2024). The counseling literature finds that lower student-to-counselor ratios improve outcomes on their own (Sparks and Mulhern, 2026; Carrell and Hoekstra, 2014), so the coordinator's much smaller caseload is likely an important ingredient in CIS's effects. Our results complement this body of work: a broader, more intensive support role operating alongside school counseling can deliver lasting gains for the highest-need students.

This study also contributes to a growing literature on the role of social connections in shaping economic opportunity. Social networks are an important determinant of economic mobility (Chetty et al., 2022; Putnam, 2016), shape how local shocks persist across generations (Chetty et al., 2026), and influence human capital investments (Cattan, Salvanes and Tominey, 2025). A distinctive feature of CIS is its reliance on site coordinators—caring adults who identify students' needs, connect them with supports, and keep them engaged in school. Our findings add to a growing body of experimental evidence showing that social ties both directly improve outcomes and enhance the returns to traditional investments such as college access programs, emergency rent assistance,

(e.g., Porowski and Passa, 2011; Somers and Haider, 2017). The randomized evidence is confined to the short run: a single-site trial in Chicago found modest but significant test-score gains (Figlio, 2015), while a within-school randomized evaluation of case management—following approximately 2,000 students for one to two years—found null effects on academic outcomes (Corrin et al., 2015; Parise et al., 2017). We provide the first evidence on the long-run effects of the national CIS model, linking student records to IRS and Census data to follow millions of students from test scores, attendance, and behavior to high school graduation, college, and earnings. We estimate the total effect of attending a CIS school—including school-wide programming and peer spillovers—rather than case management alone. Our event studies show that CIS's effects build gradually over several years, suggesting that short-run studies capture only the beginning of the program's full impact.

housing vouchers, and job training when paired with mentorship and relational support (Barr and Castleman, 2025; Bergman et al., 2024; Caspi and Rafkin, 2025; Evans et al., 2020; Katz et al., 2022; Weiss et al., 2019).[4] A sustained relationship with a caring adult, for instance, can strengthen children's non-cognitive skills and increase educational attainment (Kosse et al., 2020; Falk, Kosse and Pinger, 2026). CIS combines these two mechanisms: a trusted adult who also brokers access to a broad network of academic and non-academic resources within K–12 schools. Social connections matter not only for adults navigating the labor market. They can be built deliberately in childhood, and they pay off decades later.

Finally, we contribute to a growing literature examining how reliably short- and medium-run impacts—particularly on test scores—predict the long-run returns to interventions targeting children. In the absence of long-run outcomes, a common practice is to project long-run effects from test score gains (e.g., Ganimian, Muralidharan and Walters, 2024; Guryan et al., 2023; Hanushek, 2011; Kline and Walters, 2016).[5] We show that test scores alone cannot account for CIS's long-run impacts: measures of engagement and behavior are also needed to capture the non-cognitive dimensions through which the program operates. This finding adds to evidence that test scores can understate the lasting impacts of interventions, particularly those that improve non-cognitive skills (e.g., Anders, Barr and Smith, 2023; Chetty et al., 2011; Jackson, 2018; Gray-Lobe, Pathak and Walters, 2023). For programs that operate mainly through achievement, such projections can perform well. In Boston's charter schools, for example, impacts on high-stakes exams are a reliable guide to four-year college enrollment (Angrist et al., 2016). In other settings, measures of student engagement and school climate also predict schools' longer-run effects above and beyond test scores (Angrist et al., 2025). More broadly, our results illustrate the promise of statistical surrogates for tracking program impacts in real time while highlighting the importance of selecting surrogate measures that align with the mechanisms through which an intervention affects long-run outcomes (Athey et al., 2025; Prentice, 1989).

II Background on Communities In Schools

Communities In Schools (CIS) is the largest and most established provider of "integrated student support" services in the United States. CIS was founded in New York in the 1970s. Today it operates in more than 3,500 schools and serves 2 million students across 27 states and the District of Columbia (CIS, 2023). CIS aims to reduce dropout rates and improve academic outcomes

[4] One exception is among formerly incarcerated individuals, where similar case-management approaches after release show little impact on post-release outcomes (Doleac, 2019).

[5] The scaling factor from Chetty, Friedman and Rockoff (2014) is frequently used for this purpose. See Appendix Table A.IV in Kline and Walters (2016) for a review of observational and experimental estimates linking test scores to earnings.

by addressing both academic and non-academic barriers to learning. The key component of the CIS model is the "site coordinator" placed in each school. Site coordinators work with school staff to identify students in need, assess the root causes of disengagement, and connect them to tailored supports. Importantly, the site coordinator is a dedicated CIS employee, not a teacher or administrator, allowing school personnel to focus on their own roles.

The barriers CIS targets sit mostly outside the classroom: food insecurity, unstable housing, untreated mental health needs, and no one at home to help with schoolwork. Coordinators supply some of the response themselves, such as mentoring and academic encouragement, but their central task is brokering: connecting the student to school, community, and government services that already exist.

CIS is organized as a decentralized network of independent local affiliates, each serving a specific geographic area, typically a city (e.g., Dallas or Chicago). Each affiliate is a separately incorporated nonprofit with its own board of directors, supported by state offices and a national organization that sets common standards. When an affiliate establishes itself in a community, it assesses local needs and develops partnerships with school districts and community organizations that provide tutoring, mentoring, health and mental-health care, food assistance, and housing support (Attea and Tuck, 2005; CIS, 2025*c*). The coordinator enters a school with the affiliate's partner network already in place, giving students access to a broad range of services that the school itself could not otherwise provide (CIS, 2025*a*,*b*).

This ability to broker a broad ecosystem of services distinguishes the CIS coordinator from the school counselor, whose role centers primarily on course scheduling and college and career advising (Mulhern, 2023; Johnston, Meer and Serra, 2026). The two typically coexist within a school and play complementary roles—counselors are among those who refer students to the coordinator. The coordinator also serves roughly a third as many students as a typical counselor, and so can sustain the frequent, individualized contact that defines case management (Parise et al., 2017; American School Counselor Association, 2024).

Affiliates finance this work through a combination of philanthropy, school-district contracts, and state and local government funding (CIS of Houston, 2024; CIS, 2025*c*). Expansion into new schools is typically paced by affiliate capacity, not by requests from individual schools or year-to-year variation in school outcomes. Capacity here means whether an affiliate has secured the funding and staffing to place a coordinator in an additional school. As a result, affiliates typically expand into schools with persistent, well-established need as capacity allows.

When CIS enters a school, all students become eligible to work with the site coordinator. At the most targeted level, students may be enrolled in "case management," which involves regular check-ins with the coordinator. There is no strict rule governing referrals. In practice, students are often identified based on observable risk factors such as chronic absenteeism, behavioral issues,

low academic performance, or signs of unmet basic needs (Parise and Hossain, 2017). Teachers, counselors, and administrators frequently refer students, though coordinators may also identify students proactively using administrative records.

A referred student first sits for a needs assessment that identifies the academic, behavioral, and non-academic barriers in the way. The coordinator then builds a support plan with the student, the family, and school staff, and connects the student to services: tutoring, credit recovery, and test preparation inside the classroom; food assistance, transportation, medical and mental health care, crisis housing, and family support outside it. Some of this the coordinator delivers personally, such as mentoring, a social-emotional learning group, or a weekly attendance check-in. The rest comes from school staff or a community partner. Throughout, the coordinator tracks progress and revises the plan as needs change, and so becomes the one adult in the building who sees a student's whole situation at once.

Case management is a high-touch intervention: the median case-managed student meets with the coordinator 29 times per year—nearly once per week during the school year. In addition to formal case management, CIS also implements school-wide initiatives such as attendance drives, behavioral-support campaigns, and family-engagement events intended to improve school climate and support all students.

Although CIS operates nationally, our analysis of academic outcomes focuses on Texas, which accounts for nearly half of all CIS schools nationwide and where the program has expanded substantially over the last three decades. Unlike many other states, CIS is written into the state education charter and receives dedicated public funding from both the Texas Education Agency (TEA) and federal Temporary Assistance for Needy Families (TANF) funds (Texas Education Agency, 2022).[6] In FY 2022, state funding for CIS totaled $33.5 million, supplemented by an additional $4.9 million in TANF funding. Figure I shows the share of students enrolled in a CIS school over time. The program expanded rapidly in recent decades: in 1994, only 1.6% of Texas public school students attended a school with CIS, compared to 12.7% by 2019.

Consistent with its mission, CIS operates in schools and neighborhoods with high poverty rates. Figure II Panel *A* maps the locations of CIS schools in Texas as of 2019, with counties shaded by their poverty rates in 2000. CIS programs are heavily concentrated in the western and southern regions of the state—areas with the highest poverty levels—and are also clustered in major urban centers in the east.

County-level poverty rates do not tell the full story. In Figure II Panel *B*, we compare the characteristics of neighborhoods directly served by CIS schools to those without CIS, using school

[6]Because CIS receives substantial public funding in Texas, the state imposes additional eligibility guidelines. Students may be enrolled in case management only if they are classified as "at-risk" by TEA, eligible for free or reduced-price lunch, engaging in delinquent activities, or experiencing family conflict or crisis (Texas Education Agency, 2021). In practice, however, these designations are widespread in CIS schools.

catchment zones to map neighborhoods to schools. Students in CIS schools are substantially less likely to be White and more likely to qualify for free or reduced-price lunch. The neighborhoods surrounding these schools also tend to be more urban, with higher poverty rates and lower levels of educational attainment. The starkest difference is in economic mobility: children from low-income families in CIS neighborhoods have rates of upward mobility 0.35 SD lower, on average, than children in non-CIS neighborhoods. CIS's targeting does not appear to simply reflect which cities the program operates in. The same panel shows that the within-commuting-zone differences are very similar to the raw ones, so within a single labor market CIS still reaches the tracts with the lowest incomes and the least upward mobility.

CIS is more common in high-poverty neighborhoods, but most neighborhoods, including many poor ones, still do not have a CIS school. Figure II Panel *C* plots CIS coverage against tract poverty. Among the least poor tracts, about 10% are served by CIS. Among the poorest, coverage rises to about 25%. CIS is well targeted, but even in the poorest tracts it reaches only one neighborhood in four, leaving substantial room for expansion.

For further background, a three-minute video at benjamindgoldman.com/cis illustrates the students CIS serves and how site coordinators work with their local affiliate to deliver services in partner schools (Posner, Watson and Shanti, 2026).

III Data

Three sources of data support the analysis. Texas administrative education records follow students from elementary school into the labor market. Administrative records from CIS of Texas log every interaction between a case-managed student and a coordinator. Linked Census and IRS records carry the earnings analysis beyond Texas, to children who grew up in neighborhoods zoned to a CIS school anywhere in the country.

III.A Texas Education Research Center

Our primary data come from administrative academic, postsecondary, and labor market records maintained by the Texas Education Research Center (ERC) at the University of Texas at Austin. These data link individual-level records from the Texas Education Agency (TEA), the Texas Higher Education Coordinating Board (THECB), and the Texas Workforce Commission (TWC). Critically, the ability to track students across these datasets allows us to follow a single individual from elementary school through high school, into college, and into the labor market. A more detailed description of the data construction can be found in Appendix A.

We begin with student-level data from TEA on all students enrolled in Texas public schools

between 1994 and 2019. The dataset includes more than 109 million student-year observations for approximately 16 million unique students. For each student, we observe both time-invariant characteristics (birth cohort, race/ethnicity, and sex) and time-varying information (campus attended, grade level, FRPL status, and indicators for being classified as "at-risk" under TEA guidelines).[7]

We also identify whether each student ever attends a school served by the CIS program. To do this, we worked directly with Communities In Schools to obtain a list of all schools with a CIS program as of the 2019–2020 school year, including the year the program began at each school. Using this information, we construct an indicator for whether a student ever attends a CIS school and track the timing of CIS entry at each site. Of the 16 million students in our dataset, approximately 3.3 million ever attend a school that has adopted CIS by 2019.

We construct a range of in-school outcomes for each student. Our primary cognitive outcome is performance on standardized state math assessments, administered annually in grades 3-8.[8] Because standardized tests are not administered universally in high school and our empirical strategy requires observing students prior to their exposure to CIS, we focus our analysis of test score outcomes on middle school. We also examine behavioral outcomes, including attendance and discipline. For attendance, we observe the number of days present in each grading period of the school year (six periods, each six weeks long). For discipline, we construct indicators for whether a student was ever suspended (in-school or out-of-school) during the year, as well as a continuous measure of disciplinary incidents based on severity and frequency[9].

We follow students over time to measure long-run educational outcomes. For each student, we observe whether they graduate from high school by 2022. We also track in-state college enrollment using records from the Texas Higher Education Coordinating Board (THECB) through 2019.[10] We define indicators for whether a student ever attends any postsecondary institution between ages 18 and 22, and separately for two-year and four-year colleges.

We measure labor market outcomes using quarterly earnings data from the Texas Unemployment Insurance (UI) system, which capture wages for UI-covered employment but

[7]The at-risk indicator flags students at risk of dropping out of school, as defined by TEA. There are 13 criteria, including poor academic performance, interactions with the criminal justice system, parenthood/pregnancy, and experiencing homelessness. See Texas Education Agency (2013).

[8]The specific assessments vary over time (e.g., TAAS (1990-2002), TAKS (2003-2012), STAAR (2012-)), so we standardize scores within grade and year. Eventually, the STAAR exam switches from a grade-based exam to a series of end-of-course exams, so when analyzing test scores we restrict to exams taken in 2016 or earlier to avoid this change.

[9]For each suspension incident we observe the type of suspension (in-school or out-of-school) as well as the number of days the student was required to serve. We also observe the reason for the disciplinary action (e.g., fighting at school).

[10]The ERC also maintains National Student Clearinghouse records, which capture out-of-state enrollment, but only for more recent cohorts. We focus on in-state enrollment using THECB data to maximize coverage across the full panel.

exclude self-employed and informal workers. We impute zero earnings in quarters without reported income and restrict the sample to individuals who ever appear in the UI data between ages 18 and 25, covering roughly 82% of students. We then construct earnings percentile ranks by age and cohort. However, because most students in our Texas sample are too young for us to observe long-run earnings, we use these data primarily to validate results in the national sample and rely on linked Census and IRS records to measure earnings (see Section III.C).

Table I provides summary statistics from the ERC data for three groups: all students statewide, students who ever attend a school with a CIS program as of 2019, and students who attend a middle school with a CIS program in 2019. Students who ever attend a school that receives a CIS program by 2019 are on average somewhat more disadvantaged than the typical Texas student, but not dramatically so. They are less likely to be White (26.2% vs. 36.0%) and more likely to qualify for free or reduced-price lunch (52.2% vs. 47.0%). On average, they score about 0.13 SD lower on standardized math tests. Long-run outcomes are broadly similar: high school graduation rates are nearly identical, and average earnings at age 27 differ by only $1,000 ($32,400 vs. $33,700 statewide). However, these average differences mask substantial heterogeneity in student need within schools, which we explore in Section IV. Students who are case-managed are more likely to come from lower-income backgrounds and have markedly worse outcomes than their peers in the same school.

III.B CIS Student-Level Data

We use student-level data from Communities In Schools of Texas, available in years 2014-2019, and link them to Texas ERC administrative records using student identifiers. These data allow us to observe individual students' participation in CIS and the specific services they receive. There are two key components: (1) intake records for students who are formally case-managed by CIS, which include baseline needs determinations, and (2) logs of each interaction between students and CIS coordinators, detailing the date, duration, and type of service provided. We do not observe information about the coordinator such as a coordinator ID or demographic information.

The intake records contain one row per student-year for each case-managed student. Each row gives the date a parent consented to case management and the broad category of need that prompted the referral: academics, attendance, behavior, social services, or any combination of them. The records also name the specific needs the coordinator expects to target, such as academic readiness, social skills, self-esteem, classroom conduct, and family conflict.

The interaction data contain one row per student-interaction with the coordinator. For each interaction, we observe the date, the type of service provided, the organizational provider (whether the service was delivered by CIS staff or an external partner), and the duration of the interaction.

For example, a coordinator might spend 10 minutes providing food assistance through a backpack program, or one hour leading a workshop with a small group of students.

Combining these two data sources with the broader ERC administrative data, we can characterize both the types of students who receive case management from CIS and the nature of their interactions with coordinators. Case-managed students are overwhelmingly from racial and ethnic minority groups (only 12.8% are White), and largely low-income (71.0% qualify for free or reduced-price lunch). Academically, these students tend to struggle: the average score is more than half a standard deviation below the state average. Case-managed students also face higher disciplinary rates. In sixth grade, 26.5% receive an in-school suspension–more than twice the statewide average of 12.9% (see Appendix Table A.1 for details).

Consistent with these student characteristics, academics and behavior are the most common reasons for referral to CIS case management: 83.4% of students are referred for academic concerns and 65.5% for behavioral issues (not mutually exclusive). This focus is reflected in the specific needs that coordinators aim to address. The most frequently targeted specific need is "standardized test readiness," which appears in 41.0% of case-managed student-years. Behavioral supports are also prevalent: 36.0% of students are targeted for social skills, 32.3% for self-esteem, and 15.3% for issues related to classroom conduct. Appendix Table A.2 provides a full list of targeted needs.

These data also confirm that CIS provides services targeting both cognitive and non-cognitive domains. Analyzing the ten most common services provided in each year from 2015 to 2017, we find that time spent is roughly evenly split between cognitive and non-cognitive supports (see Appendix Figure A.1). The most common services include academic and social skills development, tutoring, character building, and attendance interventions. Notably, food assistance is also among the most frequently provided services, potentially influencing both academic performance and behavioral engagement.

III.C National Census-IRS Data

In Section V.D, we estimate the impact of CIS on long-run earnings outcomes using national linked Census and IRS data. Our analysis relies on a combination of the 2000 and 2010 Census short forms and data from federal income tax returns in 1979, 1984, 1989, 1994, 1995, and 1998-2019. We follow Chetty et al. (2026) to construct our analysis sample.[11] We focus on children born in the 1978–1992 birth cohorts who appear in the 2020 Numident file and are successfully linked to a parent via dependent claiming on 1040 tax forms.[12]

We measure both parent and child income using IRS tax records. Parent income is defined

[11] See Section II of Chetty et al. (2026) for additional details on data sources and sample construction.

[12] Valid parents are adults who appear in the 2020 Numident and would have been between ages 15–50 at the time of the child's birth. The first such individual(s) is (are) assigned as the child's invariant parent(s).

as the average of total household income between child ages 13–17, imputing zeros for any year without a tax filing. We exclude families with negative mean income over this five-year period, which typically indicates sufficient wealth to realize substantial capital losses. All dollar values are inflation adjusted to 2025 dollars. We rank parents relative to others with children in the same birth cohort to construct percentile ranks. For children, we measure household income at age 27 (including spousal income if present) and convert it to percentile ranks relative to all other children nationally in the same cohort.

Although we do not observe school enrollment in these data, we infer neighborhood of residence based on the Census tract from which a parent first files a 1040 tax return claiming the child. We assign each child to this invariant neighborhood and crosswalk Census tracts to school catchment zones using the procedure described in Section V.D.

Our final sample includes 689,200 children who grew up in neighborhoods that ever received a CIS program during our analysis window. Of those, approximately 392,000 children would have been exposed to the program, based on their age at the time of implementation in their assigned school. We return to these data after detailing our empirical methodology and estimating the effect of CIS on short-run student outcomes in the Texas data.

IV Empirical Methodology

If CIS programs had been randomly assigned, we could identify causal effects by simply comparing outcomes in schools that did and did not receive the program. In practice, adoption was not randomized: local CIS affiliates select schools based on persistent need as well as funding and staffing constraints that determine the timing of expansions. We therefore use a difference-in-differences approach, comparing changes in outcomes within CIS schools following program entry to changes over time in a comparison group of similar schools that either never adopt the program or have not yet done so.

CIS operates at the school level, but its most intensive supports, such as case management, reach only a subset of students, as we describe in Section II. We observe actual case-management status for students in later years but do not use it directly for estimation. A negative, contemporaneous shock can both raise the probability that a student becomes case-managed and independently affect downstream outcomes. Conditioning on the observed label would therefore confound the effect of CIS with the effect of the shock that triggered case management. We do not compare case-managed students with other students. Instead, we construct a proxy for case-managed status from pre-treatment student characteristics.

As a starting point, we assess how well individual pre-treatment variables predict case management status using simple, univariate comparisons. For each student who attends a

CIS middle school between 2014 and 2019, we construct an indicator for ever receiving case management during middle school, using the CIS microdata. In Figure III Panel *A*, we show that 19.1% of students in the bottom 5% of the elementary school standardized math score distribution are ever case managed in middle school, compared to 4.5% of students in the top 5% of the distribution. Similarly, students with the highest number of elementary school absences are nearly twice as likely to be case-managed, compared to students with near-perfect attendance. These univariate predictors show clear targeting patterns, but explain only a modest share of the variation in case management status on their own.

To improve predictive accuracy, we train a gradient-boosted decision-tree model using the full set of elementary school predictors that we observe (Ke et al., 2017). Inputs include standardized math scores, attendance, disciplinary history, grade retention, and demographic variables such as race and ethnicity, sex, and FRPL status. The model is trained using five-fold cross-validation and evaluated on a held-out test set (see Appendix B for additional details). In Figure III Panel *C*, we show the out-of-sample accuracy of the model. For each percentile of the prediction distribution, we plot the fraction of students in the hold out data who were ever case-managed. The model has an AUC of 0.71. If we randomly draw one case-managed student and one non-case-managed student, the model assigns the higher score to the case-managed student 71% of the time. By including the full set of elementary predictors, we achieve greater predictive power than any individual variable: among the 5% of students with the lowest predictions, only 1.1% are case managed, while in the top 5% of predictions, 37.4% are. Figure III, Panel *D*, shows the relative importance of each predictor based on model "gain," a measure of the improvement in model fit attributed to splits on that variable. Test scores are among the most important predictors, but other features—including attendance, grade retention, and prior disciplinary incidents—also contribute meaningfully.

We then apply the trained model to the full panel of students, in every year and every school, whether or not the school had adopted CIS. Every student therefore carries a predicted case-management risk score. Throughout the paper we call students in the top quartile of the score distribution "high-risk" and students in the bottom quartile "low-risk" (Figure III, Panel *C*). The score marks high-risk students in treated schools before CIS arrives, and it marks observationally similar high-risk students in control schools, who are never exposed to the program.

Table II reports summary statistics for high- and low-risk students in both CIS and non-CIS schools. We find large differences between high- and low-risk students across both groups of schools. For example, 74.9% (77.6%) of high-risk students are eligible for FRPL in a given year, compared to 9.2% (7.1%) of low-risk students in CIS (non-CIS) schools. Disparities in outcomes are similarly large: high-risk students score 1.4 SD lower than low-risk students in 6th grade and earn half as much at age 27 ($22,000 vs. $46,000). Although outcomes for these groups are similar in CIS and non-CIS schools, the concentration of high-risk students is substantially higher

in CIS schools. High-risk students outnumber low-risk students by a factor of 1.67 in CIS schools, compared to 0.88 in non-CIS schools.

To estimate the effect of CIS on student outcomes, we run an event study comparing outcomes in treated schools before and after adoption with outcomes in control schools that never receive CIS. We measure outcomes at the student-year level (a 6th grade test score, for example) and date event time from the program's first year. We estimate treatment effects over an eleven-year window centered on adoption. Because we use elementary school characteristics to predict case-management risk, we focus our analysis on programs that start in middle schools.

We match each treated school to one control school with similar pre-treatment characteristics. For each treated school we measure covariates six years before its CIS program starts, at event time $t = -6$, so the match rests on data outside the event study window. For a candidate control school we measure the same covariates at $t = -6$ relative to the start year of the treated school it would match, which lines the two up in calendar time: a control school paired with a school that adopts CIS in 2012 is measured in 2006. We estimate a linear probability model of CIS adoption on school-level demographics and outcomes in the donor pool of untreated schools, predict the probability of treatment for every school, and match each treated school to the untreated school with the closest predicted value.[13]

After matching, we stack the treated and control school pairs across implementation cohorts and estimate the following event study specification:

$$y_{i,t} = \sum_{\tau \in [-5,5]; \tau \neq -1} \beta_\tau \cdot (\mathbf{1}[t - T(s) = \tau] \cdot CIS_s) + \lambda_{s,T(s),r,m} + \gamma_{t,T(s)} + \varepsilon_{i,t}, \tag{1}$$

where $y_{i,t}$ is an outcome for student i in year t. Student i attends school s, and $T(s)$ is that school's implementation cohort: the year school s adopts CIS, or, for a never-treated control school, the year its matched treated school adopts. The treatment indicator CIS_s equals one in treated schools and zero in their matched controls, so the coefficients β_τ measure the gap in outcomes between the two groups at relative time τ, with $\tau = -1$ the reference period. The fixed effects $\lambda_{s,T(s),r,m}$ absorb level differences across schools by implementation cohort, student i's race r, and sex m. The $\gamma_{t,T(s)}$ fixed effects absorb year-specific shocks within each implementation cohort, tracking how outcomes evolve in the control schools.

In many cases, we also report results from a more parsimonious difference-in-differences

[13]The full set of predictors includes the share identified as at risk of dropping out (per TEA); the share of White students; the fraction of students in the top and bottom quartiles of the predicted case management distribution; the absence rate; and the predicted high school graduation rate, described in Section VI. We obtain similar results when expanding the set of control schools to the nearest 5 matches or restricting to unique control matches.

specification:

$$
\begin{aligned}
y_{i,t} = & \beta_{post,0-2}(\mathbf{1}[0 \leq \tau \leq 2] \cdot CIS_s) \\
& + \beta_{post,3-5}(\mathbf{1}[3 \leq \tau \leq 5] \cdot CIS_s) \\
& + \tilde{\lambda}_{s,T(s),r,m} + \tilde{\gamma}_{t,T(s)} + \varepsilon_{i,t}.
\end{aligned} \tag{2}
$$

This specification includes the same school and year fixed effects as Equation 1, but instead of estimating separate coefficients for each relative year, it averages treatment effects over two post-treatment periods: years 1–2 and years 3–5. We focus primarily on $\beta_{post,3-5}$, which captures the average effect several years after implementation. In our baseline specification, we cluster standard errors at the school-by-birth-cohort level to allow for within-unit dependence over time. We also report standard errors clustered at the school level.

The key identifying assumption in this design is that, absent the introduction of CIS, outcomes in the treated schools would have followed the same path as those in the matched control schools. This assumption may be violated if the timing of CIS coincides with other reforms or resource shifts at the school (e.g., leadership changes, increases in funding) that also impact student outcomes. It could also be threatened if the composition of students changes at the time of CIS adoption, for example due to differential sorting. In the next section, after presenting our baseline results, we assess the robustness of these findings to a range of alternative specifications designed to address these concerns.

V Effect of CIS on Student Outcomes

We begin by documenting the impact of CIS adoption on students' exposure to the program. For each student-year, we calculate cumulative years spent in a middle school with a CIS program. We estimate Equation 1 separately for high- and low-risk students, comparing changes in exposure in treated versus matched control schools. Figure IV presents the event study results. Before program adoption, there are no differences in exposure, as expected.[14] Following CIS entry, exposure rises rapidly and then levels off, consistent with students progressing through middle school grades in CIS schools. In the absence of student mobility across schools or grade retention, we would expect exposure to equal one year at $\tau = 0$, increase to approximately 1.67 at $\tau = 1$, and level off at 2.0 in

[14] Although the CIS program does not begin in treated schools until $t = 0$ and control schools are never treated, pre-period exposure is not mechanically 0. Students could have attended a different middle school with a CIS program before enrolling in either the treatment or control school. In practice, however, at $t = -1$, average exposure is less than 0.03 years. In Appendix Figure A.5, we also show there is essentially no difference in exposure to CIS in elementary school between these groups.

subsequent years.[15] Observed exposure closely tracks this pattern. At $\tau = 1$, students average 1.45 (high-risk) and 1.56 (low-risk) years of exposure, roughly 90% of the anticipated exposure level.

In the years three to five after adoption—the main period we use for estimating more parsimonious DID effects throughout our analysis—students average 1.70 years of exposure (high-risk) and 1.75 years (low-risk). These estimates are statistically indistinguishable (t-statistic on difference = –1.38; see Table III), confirming that high- and low-risk students experience similar exposure to CIS at a school-level once the program is in place.

V.A Impacts on Test Scores

Does that exposure raise test scores? We estimate Equation 1 for high-risk students in middle schools, with results in Figure V Panel *A*. Coefficients in the pre-period center on zero, supporting the parallel trends assumption. Control schools were matched on covariates measured at $\tau = -6$, outside the event-study window, so the pre-period serves as an out-of-sample test of parallel trends against the matched controls. The adoption process in Section II predicts that treated and control schools should not exhibit differential pre-treatment trends: affiliates expand into persistently high-need schools as funding and staffing permit, not in response to trends in school performance.

Following CIS adoption, test scores begin to improve in treated schools relative to their matched controls, before stabilizing in later years. The pattern mirrors the exposure to the program shown in Figure IV. We estimate a difference-in-differences effect of 0.11 SD (SE = 0.02) for high-risk students. When we scale the effect using an instrumental variables (IV) specification based on years of exposure, we find that three years of CIS exposure increases test scores by 0.18 SD (SE = 0.04).

Off-the-shelf standard errors ignore one source of uncertainty in our setting: we estimate risk status with a predictive model rather than observing it. The t-statistic in Figure V Panel *A* is 4.67, so standard errors would have to be about 140% larger before the effect lost significance at the 5% level. A clustered bootstrap that re-estimates the full decision-tree model in every draw would handle the problem formally, but it is computationally infeasible in the ERC environment. Instead we show that the results survive without the model. We define high-risk students as those in the bottom of a combined index averaging elementary math scores and attendance, and low-risk students by the corresponding cutoff at the top. Because the cuts are low-dimensional and rest on observed characteristics rather than the output of a flexible model, they carry less estimation uncertainty. The alternate definition delivers treatment effects similar to our baseline, with t-statistics above 3 for high-risk students (see Appendix Table A.3). In addition, every baseline

[15] If students never move and there are an equal number of students in 6-8th grade in a school, then in any given year at $\tau \geq 3$, 6th graders will have one year of exposure, 7th graders 2, and 8th graders 3, averaging to 2 years of exposure.

estimate for high-risk students that is significant at the 5% level stays significant under the more conservative school-clustered standard errors (Appendix Table A.4).

How do the effects of CIS on test scores for high-risk students compare to the effects for their low-risk peers? In Figure V Panel *B*, we repeat the analysis for low-risk students. As with high-risk students, we observe coefficients close to zero in the pre-period, supporting the parallel trends assumption. However, following CIS adoption, we find no evidence of improvement in test scores for low-risk students; the estimated treatment effect is 0.01 SD (SE = 0.02). In Table III, we report the DID coefficients from a fully interacted joint specification, estimating the treatment effects for high- and low-risk students simultaneously. We can reject that treatment effects for high- and low-risk students are the same (t-statistic = 3.23).

To assess whether the treatment effects decline gradually across the risk distribution or are concentrated entirely among high-risk students, we repeat this analysis for the middle two quartiles of the predicted risk distribution. We find a monotonic pattern: treatment effects are largest for high-risk students, somewhat smaller for students in the third quartile, smaller still for those in the second quartile, and near-zero for low-risk students (see Appendix Figure A.2). The gains therefore fall off smoothly as need falls, rather than dropping away at the high-risk cutoff.

High-risk students are more likely to receive case management, and CIS meaningfully improves outcomes for this group. However, we do not interpret the difference in treatment effects across groups as the causal effect of case management alone because CIS can also improve outcomes through channels such as school-wide initiatives and spillovers within and across groups of students. For the same reason, we do not rescale the reduced-form estimates by case-management take-up, which would attribute the entire effect of CIS to case management. Low-risk students may benefit less in part because they are less likely to receive case management and in part because they have fewer unmet needs. We therefore read the gap between the two groups as evidence about who benefits from CIS, not as an estimate of what case management alone does.

A natural question is whether the gains for high-risk students vary across subgroups. We re-estimate Equation 2 separately by sex and ethnicity within the high-risk sample. We find that the test score effects of CIS are similar for male and female students (0.127 vs. 0.091 SD) and for Hispanic and non-Hispanic students (0.106 vs. 0.117 SD; see Appendix Table A.5). In neither case can we reject equality of the effects across groups (t-stat=1.46, t-stat=0.33). We find little evidence that the program's academic benefits are concentrated within any particular demographic group; instead they vary by differences in student need.

We perform several exercises to test the robustness of these results. We first allow for a more flexible set of controls, augmenting Equation 1 so that the year-specific shocks in γ can also vary flexibly by race/ethnicity and sex. This specification allows different demographic groups

to experience distinct time trends in test scores by defining the counterfactual using students in control schools from the same demographic group. The results are shown in Figure VI Panels *A* (high-risk) and *B* (low-risk). Treatment effects for high-risk students remain positive, while those for low-risk students remain smaller and statistically insignificant.

Next, we compare earlier-treated schools to later-treated schools instead of matched control schools, and the results hold. For each implementation cohort $T(s)$ and relative time period τ, we define the set of control schools as those with later implementation cohorts ($T(s') > T(s) + \tau$). In order to compare to schools that received CIS around the same period as the treated school, we restrict comparison schools that were treated in a 5 year period ($T(s') - T(s) \in [3,8]$).[16] We estimate a regression analogous to our baseline specification, using the same λ and γ fixed effects.

The results are shown in Figure VI, Panels *C* (high-risk) and *D* (low-risk). This approach yields similar treatment effects: for high-risk students, we estimate a 0.09 SD increase in test scores (SE = 0.01), compared to 0.11 (0.02) in the baseline; for low-risk students, the estimate is 0.01 (0.02), similar to the baseline effect of 0.01 (0.02). Unlike our baseline design, which uses never-treated schools as controls, this approach exploits only the quasi-random timing of CIS adoption across eventually treated schools.

A further concern is that other policy changes may coincide in timing with the introduction of CIS in a given commuting zone, driving the effects we attribute to the program. We therefore restrict the comparison to schools within the same commuting zone. Specifically, we interact the γ year fixed effects with commuting zone indicators and compare earlier- and later-treated schools within the same city over time. Doing so ensures that both treatment and control schools are subject to the same commuting zone-level policies, funding environments, and initiatives. We continue to define the control group dynamically as schools with later implementation dates, as in the previous specification. Figure VI, Panels *E* and *F*, show stable results. Here we can reject a null effect for low-risk students (t-statistic = 2.2), but the effect for high-risk students is nearly twice as large (0.07 vs. 0.04 SD), as in the previous specifications. This within-city design provides additional support for the claim that the effects we observe are driven by the CIS program rather than differential local trends that coincide with the timing of CIS entry.

A related concern is that the composition of a school's students or staff changes around the time CIS arrives, and that the change, rather than the program itself, drives the improvements we observe. We test for changes in observable school-wide characteristics at the time of adoption. Figure VII presents the results. In Panel *A*, we plot an event study of the fraction of students who are White, finding no evidence of any changes in racial composition around program entry.

[16] Because control schools eventually receive treatment, the panel is not balanced: control schools drop out of the comparison group in a given relative year once they begin treatment themselves. However, by comparing to schools that received CIS up to 8 years later, we ensure that the later relative time coefficients in the event studies are estimated using multiple later treated cohorts.

In Panel *B*, we extend this analysis to a broader set of student, teacher, and staff characteristics, reporting difference-in-differences estimates in standard deviation units for comparability. Across nearly all characteristics—including student demographics, teacher experience, and staff size—the estimated effects are small and statistically indistinguishable from zero. The closest exception is average teacher pay, which falls by 0.1 SD (approximately $900 per year) and has a t-statistic of 1.86. We find no evidence that the presence of CIS substitutes for other non-teaching staff members at a school, such as counselors, social workers, and school psychologists. These checks indicate that the estimates reflect the effects of CIS itself rather than changes in student composition or school characteristics. They are also consistent with the adoption process described in Section II, in which CIS expands into schools with persistent need rather than in response to changing school circumstances.

Balance on observable characteristics does not rule out sorting on unobservables. Higher-ability, high-risk students may still sort into CIS schools at entry, so we construct an alternative assignment measure that predicts middle school enrollment based on students' elementary school rather than their actual middle school.[17] We then re-estimate treatment effects using this predicted school assignment, holding treatment and control matches fixed. As shown in Figure VI Panels *G* and *H*, the results remain similar: we continue to find gains for high-risk students and no effects for low-risk students. Since childhood moves somewhat reduce the effect on CIS exposure in this exercise, we also report IV estimates based on three years of exposure for both actual and anticipated school assignments (Appendix Figure A.3). The similarity of the estimates across approaches helps alleviate concerns about endogenous sorting into CIS schools.

Finally, we show that these patterns are not specific to math or driven by changes in test-taking behavior. Reading scores follow the same pattern, with smaller gains for high-risk students and none for low-risk students. After CIS enters a school, reading scores rise by 0.05 SD for high-risk students (SE = 0.02) and do not move for low-risk students (–0.01; SE = 0.02). We also test whether CIS affects the likelihood that students take the standardized exams. For each student-year, we construct an indicator for missing a standardized math score and estimate treatment effects on this outcome. High-risk students in treated schools become slightly less likely to be missing test scores (–0.4 pp), but the change is not statistically significant (t = –0.8). We similarly find no detectable effect for low-risk students. These results indicate that the achievement gains we document are not driven by CIS changing the composition of test-takers in treated schools (see Appendix Figure A.4).

Taken together, these results provide consistent evidence that CIS improves academic outcomes for high-risk students. Across all specifications, we find positive effects for high-risk

[17] We construct a transition matrix from grade g to grade $g+1$ separately by birth cohort and assign students to schools based on modal transitions from their first observed elementary school.

students and no or modest gains for their low-risk peers. For high-risk students, the gain in math scores is about a tenth of a standard deviation.

V.B Impacts on Student Behavior

In addition to academics, a primary focus of the CIS coordinator is improving students' behavior. As described in Section III.B, two-thirds of students referred to case management are flagged as having behavioral needs. Those students may receive targeted supports such as mentoring, referrals to mental health counseling, and conflict resolution support. In addition to case management, coordinators target improvements in attendance culture for all students as one of their school-wide goals.

To measure the effect of CIS on student behavior, we define a composite measure of student engagement—days spent in class, meaning days neither absent nor removed for disciplinary reasons. We first examine whether CIS reduces chronic disengagement, defined as being out of the classroom for 18 or more days in a school year—approximately 10% of the school year, following the conventional definition of chronic absenteeism Department of Education (2026). We estimate effects separately for high- and low-risk students. Figure VIII presents the corresponding event studies. Before CIS, 15.4% of high-risk students were chronically disengaged. After it arrived, the share fell by 1.8 pp (SE = 0.6 pp). The decline was smaller among low-risk students, 0.8 pp (SE = 0.3 pp), who started from a base of 3.1%.

We also measure the effect of CIS on changes in engagement measured in total days in the classroom. On average, high-risk students are out of class nearly twice as much as low-risk students (9.4 vs. 5.1 days). After the introduction of CIS, students in both groups spend more time in the classroom. We estimate an increase of 0.49 days per year for high-risk students (SE = 0.16) and 0.39 days for low-risk students (SE = 0.11; see Appendix Table A.6). The effects differ by risk group. Among low-risk students, the improvement reflects an increase in near-perfect attendance, despite modest effects on chronic disengagement: the probability of near-perfect attendance increased by 2.1 pp (SE = 0.9 pp; see Appendix Table A.6). The effects on student engagement throughout the risk distribution are consistent with coordinators' school-wide focus on improving attendance.

Because our measure of engagement combines absences and discipline, we separate absences from disciplinary removals. On average, the effects are driven by changes in attendance. The increase of 0.49 days in class for high-risk students comes from an increase in attendance of 0.45 days (SE = 0.14) and a decrease in days missed due to discipline of 0.04 days (SE = 0.06; see Appendix Table A.6). For boys, who are more likely to have disciplinary records, a larger share of the reduction comes from discipline. For high-risk boys, an increase of 0.31 days in class reflects

0.13 additional days of attendance and a 0.18-day reduction in disciplinary days, though the effects are imprecise (see Appendix Table A.6). We also find that CIS decreases low-level disciplinary infractions for boys stemming from fighting, as opposed to changes in more serious infractions or changes along the extensive margin of having any disciplinary record (see Appendix Table A.6). CIS therefore returns days to the classroom, and it returns the most to the students who were losing the most. The clearest gain is the drop in chronic disengagement among high-risk students.

V.C Impacts on Long-Run Outcomes in Texas

In addition to improving academic performance, CIS aims to increase high school graduation and economic opportunity, particularly for students at risk of dropping out. In this section, we examine whether the gains in test scores for high-risk students translate into improvements in students' long-run outcomes. We focus on high school graduation as a key milestone, followed by postsecondary enrollment, and labor market outcomes.

Unlike test scores and other contemporaneous outcomes, long-run outcomes are only observed once per student. As a result, our empirical design must shift from capturing year-specific exposure to capturing a student's cumulative exposure to the CIS program. In Section V.A, we defined a single implementation cohort per school, $T(s)$. Here, we instead define a grade-specific cohort, $T(s,g)$: the first year in which students in school s and grade g are ever exposed to CIS during any part of middle school.

As an example, consider a middle school that adopts a CIS program in 2005. The students who were in 8th grade in 2004 complete middle school before the program begins and are therefore never treated. In contrast, the 7th graders in 2004 are exposed to CIS during their 8th grade year, and the 6th graders in 2003 are exposed only in 8th grade, while the 6th graders in 2004 are exposed in both 7th and 8th grade. In this case, we define $T(s,8) = 2005$, $T(s,7) = 2004$, and $T(s,6) = 2003$. This cohort-specific adjustment allows us to estimate the following grade-specific version of Equation 1 that accounts for cumulative exposure to CIS:

$$y_i = \sum_{\tau \in [-5,5]; \tau \neq -1} \beta_\tau \cdot (\mathbf{1}[t - T(s,g) = \tau] \cdot CIS_s) + \lambda_{s,T(s,g),r,m,g} + \gamma_{t,T(s,g),g} + \varepsilon_{i,g}, \tag{3}$$

where y_i is a long-run outcome for student i, who attends school s and is observed in grade g in year t. This approach is analogous to estimating treatment effects using a single-grade sample (e.g., restricting to students observed in 8th grade), and stacking separate single-grade specifications into a pooled regression framework.

We begin by estimating the effect of CIS adoption on cumulative exposure to CIS in elementary, middle, and high school, as well as total exposure across all grades. Our identifying

variation comes from programs that begin in middle schools, and the differences in exposure concentrate there, as they should. In later years ($\tau \geq 3$), students in treatment schools gain only 0.12 additional years of elementary school exposure, just 5% of the treatment–control difference in middle school. Nearby elementary schools adopting CIS in later years account for most of that gap. High school exposure does not differ between treatment and control students. Middle school exposure drives our estimated effects (see Appendix Figure A.5). To account for the small elementary differences, our IV estimates use exposure across all grades.

Next, we estimate the effect of CIS on high school graduation, separately for high- and low-risk students. Results for high-risk students are shown in Figure IX Panel *A*. As with test scores, we find positive treatment effects: pre-period coefficients are near zero, and graduation rates rise following CIS adoption. We estimate a difference-in-differences effect of 3.1 pp (SE = 1.0 pp).[18] Rescaling for the effect of 3 additional years of total exposure, we find that CIS increases graduation rates by 3.4 pp (see Table III). This represents a 5.2% increase over the high-risk baseline graduation rate of 65%, and closes 14% of the pre-treatment gap in graduation rates between high- and low-risk students.

As with test scores, we find no corresponding increase in graduation rates for low-risk students. Figure IX Panel *B* shows that the pre- and post-period coefficients are similar; the apparent decline in the post-period reflects the elevated $\tau = -1$ coefficient rather than a persistent decline. We estimate a small and statistically insignificant decline of 1.1 pp (SE = 0.7 pp) off a high baseline graduation rate of 89.3%. While we cannot reject a null or small positive effect for low-risk students, we can reject impacts as large as those observed for high-risk students (see Table III). These graduation gains are similar across demographic groups, with comparable effects for male and female and for Hispanic and non-Hispanic students (see Appendix Table A.5). Like with test scores, we find that the effects on high school graduation are monotonically increasing throughout the risk distribution, showing that CIS has the largest impacts for the students most likely to interact with the program (see Appendix Figure A.6).

Together, these results show that CIS improves high school graduation rates—delivering on one of the program's central goals. The effects are concentrated among the students most likely to be served by the program, mirroring the pattern observed for test scores. By providing high-risk students with targeted support during middle school, CIS meaningfully alters their educational trajectory, leading to lasting improvements even after students leave the program's direct reach.

We next test whether the gains in graduation rates translate into increased postsecondary enrollment in Texas. We focus on 2-year college attendance, since students on the margin of graduating high school are more likely to attend 2-year rather than 4-year institutions. We estimate a 2.7 pp increase in the share of high-risk students who ever enroll in a 2-year college between

[18] We define graduation as ever appearing as a high school graduate, but our results are not sensitive to this definition.

ages 18 and 22, but the standard error is large (1.8 pp), and the effect is not statistically significant. In contrast, we find no increase for low-risk students (see Appendix Figure A.7). Consistent with 2-year college being the primary margin of response, we find little evidence of effects on other types of postsecondary enrollment. For high-risk students, the estimated effect on attending any institution in the Texas Higher Education Coordinating Board (THECB) data is nearly identical to the effect on 2-year enrollment (2.9 pp vs. 2.7 pp; see Appendix Figure A.7 Panel *C*). [19]

CIS increases educational attainment for high-risk students, both by raising high school graduation rates and through suggestive increases in 2-year college enrollment. To assess whether these educational gains translate into improved economic outcomes, we next examine impacts on earnings in early adulthood. Estimating earnings effects in our current Texas-based design requires a long panel: we must observe students in elementary school to predict risk and follow them for multiple decades to measure earnings at a meaningful age. There are only three birth cohorts for which we can measure in-school data in first grade (used for the risk prediction) and earnings at age 27.[20] To overcome this limitation, we turn to a second design based on CIS programs outside of Texas, using national data to measure labor market outcomes.

V.D Impacts on Earnings in National Sample

We estimate the impact of CIS on earnings using linked Census and IRS administrative data covering nearly the entire U.S. population. The new data and setting force three changes in approach. First, we study CIS programs outside of Texas, which allows us to test whether the program's impacts generalize beyond the Texas context. Second, the national data do not record which school a child attends, so we compare cohorts within treated and control neighborhoods rather than within treated and control schools, exploiting differences in CIS exposure from the timing of enrollment relative to program adoption. Third, the national data also lack detailed early-life measures, so we do not separate high-risk from low-risk students here, though we explore heterogeneity by parent income.

Because we do not observe schools, we rely on addresses to map between neighborhoods and school catchment zones. We assign children to neighborhoods based on the first Census tract in which they are claimed by their parents.[21] We then map each tract to the set of schools that contain

[19] 31.5% of high-risk students attend college at some point between ages 18 and 22; among them, nearly all attend at least one 2-year college (97%). Only 9% of high-risk students ever attend at least one 4-year college in the pre-period, and we find no impact of CIS on 4-year college attendance. Low-risk students attend 4-year colleges at much higher rates (45%) but also show no treatment effect (DID = –0.06 pp).

[20] For example, a student in first grade in 1995—the first year of our education data—turns 27 in 2016. We can measure data over the entire window for the first-graders in 1995,1996, and 1997.

[21] This is analogous to the approach we take in Appendix Figure A.3, assigning students to schools based on their first elementary school, and helps to limit concerns of endogenous sorting into schools and neighborhoods with CIS programs.

the tract in their catchment zones following the approach in Chetty et al. (2025).[22]

Unlike in Texas, we do not restrict attention to middle schools. We are no longer constrained by the availability of universal standardized tests or by the need to use elementary data to assign high- and low-risk status. Instead, we use programs at all levels of school and map CIS exposure to birth cohorts. For each tract ever served by CIS, we identify the first birth cohort exposed to CIS, based on the timing of the program's introduction. For example, if a tract contains both an elementary school that adopted CIS in 2002 and a middle school that adopted in 2000, we define the first treated cohort as the 8th graders enrolled in the middle school in 2000—primarily students born in 1987. We denote this first exposed cohort in tract n as the tract's implementation cohort, $C(n)$, analogous to our approach in Section IV.

We implement a staggered adoption design, comparing outcomes for children in tracts exposed to CIS earlier versus later. For each implementation cohort $C(n)$, we define the control group as the set of tracts where CIS is introduced more than five birth cohorts later—that is, all n' such that $C(n') > C(n) + 5$.[23] For these later-treated control tracts, we assign a pseudo-implementation cohort $\tilde{C}(n') = C(n)$, matching the implementation cohort of the treated group they are compared to. We then stack the data across implementation cohorts following the approach of Sun and Abraham (2021), fully interacting all specifications with indicators for the implementation cohort. This structure allows later-treated tracts to serve as controls for multiple earlier implementation cohorts. We cluster standard errors at the high school catchment level.

We estimate:

$$y_i = \sum_{\tau \in [-5,5]; \tau \neq -1} \beta_\tau \cdot (\mathbf{1}[c - \tilde{C}(n) = \tau] \cdot CIS_n) + \phi_{n,r,p,\tilde{C}(n)} + \delta_{c,r,p,\tilde{C}(n)} + \varepsilon_i, \tag{4}$$

where y_i is an adult outcome for individual i, who grew up in Census tract n, was born in cohort c, belongs to race/ethnicity group r, and whose parents are in income quartile p. The term $\tilde{C}(n)$ denotes the implementation cohort for the treated tract and the pseudo-implementation cohort for the matched control tract.

This structure closely parallels our baseline specification in Equation 1, adapted to a cross-cohort design. The fixed effects ϕ absorb neighborhood-by-demographic variation within each implementation cohort, analogous to the school-by-demographic fixed effects in Texas. The δ terms control for cohort-by-demographic shocks within implementation cohorts, analogous to

[22] We use the same crosswalk between school catchment zones and Census tracts as in Chetty et al. (2025), provided by Peter Bergman and originally from Maponics (2017).

[23] This restriction ensures that the sample remains balanced across all relative time periods in the event study window. In Section V.A and Figure VI, we do not impose this restriction. However, applying it here does not drive the results: the unbalanced panel yields similar conclusions. Similarly, our results are not driven by the choice of control group. We show that our results are robust to using a matching procedure similar to the one used in the Texas sample (see Appendix Figure A.11).

year fixed effects. The coefficients β_τ trace the path of estimated treatment effects from five cohorts before to five cohorts after the CIS program begins in treated tracts.

We begin by estimating the effect of CIS entry on the total number of years spent in a school with a CIS program, using Equation 4. After assigning students to treatment status based on their initial Census tract, we define total exposure using a student's observed Census tract in each subsequent year. Figure X, Panel *A* plots the event study coefficients. The pattern is similar to the one shown in Figure IV: exposure begins to rise after the program is introduced and then levels off. To interpret the magnitudes, consider the coefficient at event time $\tau = 0$, corresponding to the first cohort exposed to CIS. If no students moved away from their initial Census tract, this coefficient would mechanically equal 1.[24] Instead, we estimate a coefficient of 0.51, reflecting migration during childhood. Unlike our Texas middle school setting, the national design does not cap exposure at three years. A child can accumulate more, depending on how many of the elementary, middle, and high schools in the neighborhood adopt CIS. Children born 3–5 years after the implementation cohort have the longest potential exposure, and we estimate that they experience an average of 2 additional years of CIS exposure relative to children in later-treated control tracts.

Does this exposure translate into improved outcomes later in life? Figure X Panel *B* plots the effect of CIS on household income rank in the national distribution at age 27. The estimated coefficients in the pre-period are close to zero and statistically insignificant, suggesting that treatment and control tracts were on parallel trajectories before CIS was introduced. In the years following implementation, children in the treated tracts begin to earn more than their peers in later-treated control tracts. Estimating the difference-in-differences analog of Equation 4, we find that CIS improves adult income ranks by 1.14 ranks (SE = 0.37) for students born 3-5 years after the implementation cohort. This corresponds to a 3.2% increase in income—about $1,100 annually—relative to a pre-period average of $35,700.[25] Rescaling by the estimated effect on years of exposure, we estimate that 3 years of exposure to CIS programs increases earnings at age 27 by 1.54 percentile ranks (SE = 0.45), equivalent to a $1,550 increase (4.3%) in annual earnings, measured in 2025 dollars.

The earnings results come from CIS programs outside of Texas, but we can examine whether the same gains appear in Texas. Our Texas panel is too short to observe earnings ranks at age 27 for enough cohorts, so we predict them using earlier observed outcomes. We estimate a

[24] We measure exposure assuming that students attend their local schools, though this is not always the case. Chetty et al. (2025) use the same tract-to-school crosswalk and show in North Carolina that it reliably predicts actual school attendance. Deviations from local attendance would lead us to overstate exposure, making our estimates an upper bound for exposure and any IV estimates a lower bound for the treatment effect.

[25] We do not find evidence that these effects are driven by changes in marriage rates. We estimate very similar effects when using individual income ranks (see Appendix Figure A.11).

gradient-boosted decision-tree model, similar to the one we used to predict case management in Section IV, to forecast income rank at age 27 from earnings ranks and college attendance between ages 18 and 22 (Athey et al., 2025; Chetty, Deming and Friedman, 2025). See Appendix Figure A.9 for details. Using the predictions as the dependent variable, we re-estimate Equation 3. The results closely mirror those from Figure X (see Appendix Figure A.8 for a direct comparison), suggesting that the earnings gains from CIS are similar both inside and outside of Texas.

In the Texas sample, we found large differences in treatment effects for children who were more or less likely to be intensely served by CIS. Although we cannot construct the same predictions of high- and low-risk in the national sample, we examine heterogeneity by parent income as a proxy for student need. We estimate difference-in-differences treatment effects separately for children whose parents were in the top quartile (lower need) and bottom quartile (higher need) of the national income distribution. The results broadly mirror the patterns observed in Texas: we find a 1.14-rank increase in adult income (SE = 0.51) for children from low-income families, and no detectable effect for students from high-income families (–0.03 ranks; SE = 0.84), though the estimates are imprecise (see Appendix Figure A.10).[26] Mirroring our earlier findings, these results suggest that CIS exposure during childhood continues to benefit children with greater needs well into adulthood.

The earnings gains for low-income children are driven primarily by a lower likelihood of left-tail outcomes (non-employment or very low earnings) and a higher likelihood of holding a modestly paid job, rather than by increased chances of reaching the top of the income distribution. This pattern is consistent with CIS's mission to help students stay in school and avoid adverse outcomes such as dropping out. We construct four binary outcome variables indicating whether each child is in a given adult income quartile at age 27 and estimate difference-in-differences treatment effects on the probability of falling into each quartile. Figure XI shows that exposure to CIS reduces the likelihood of being in the bottom quartile by 2.3 pp (SE = 1.1 pp) and increases the likelihood of being in the second quartile by 1.7 pp (SE = 1.0 pp). We find no significant effects on the probability of being in the top half of the adult income distribution. The bottom quartile comprises children earning less than $20,000 annually, but 80% of these children are not working. We cannot reject that the entire effect on bottom-quartile incidence is driven by changes in employment, as proxied by the presence of W-2 earnings. We estimate an effect of CIS on employment of 1.8 pp (SE = 0.5 pp, see Appendix Figure A.11).

These patterns are consistent with the education effects documented in the Texas sample. CIS increases high school graduation rates for students in need and modestly boosts 2-year

[26] Unlike in Texas, where treatment effects decline monotonically with predicted risk, we observe somewhat larger effects among children from middle-income families. However, we have limited power to compare effects across finer income groups, and parental income is likely an imperfect proxy for the risk measures that generate sharper separation of treatment effects in the Texas sample.

college attendance, but does not shift students into four-year college pathways. For context, the earnings effects we document of $1,550 are approximately 17% of the returns to a four-year degree (Zimmerman, 2014). Using the lifetime earnings projection method from Hendren and Sprung-Keyser (2020), we estimate that three years of CIS exposure during childhood raise the present value of expected lifetime earnings by about $48,800 per student. The gain is large relative to the program's direct cost of roughly $3,750 per student, or three years of case management at about $1,260 a year (Texas Legislative Budget Board, 2016).[27] We weigh that gain against the direct cost alone. We do not observe what the other services CIS connects students to cost, nor whether CIS students crowd others out of those services, so we cannot compute the marginal value of public funds for the program as a whole.[28] Such interactions with other publicly funded programs can substantially alter a program's estimated rate of return (Kline and Walters, 2016).

VI Explaining Long-Run Impacts and Personalization

In the short run CIS raises test scores and improves student engagement. Years later it lifts high school graduation and earnings. How much of the long-run gain is already captured by the short-run outcomes, and can they serve as reliable surrogates before long-run data are available? A separate question concerns how CIS delivers its services: site coordinators personalize them to each student's needs, and we ask whether that personalization leaves any group behind.

From Short-Run Outcomes to Long-Run Impacts

How much of the long-run improvement can test-score gains alone explain? A common approach to estimating the long-run impact of educational programs is to "scale up" short-run test-score impacts, drawing on an extensive body of estimates linking test scores to later outcomes (e.g., Ganimian, Muralidharan and Walters, 2024; Guryan et al., 2023; Hanushek, 2011; Kline and Walters, 2016; You, Fu and Chou, 2025; Chetty, Friedman and Rockoff, 2014). This approach effectively treats test scores as a statistical surrogate for long-run effects (Prentice, 1989; Athey et al., 2025), allowing analysts to forecast returns when long-run data are not yet available. Whether such projections fully capture a program's benefits depends on the extent to which its

[27] We show that the earnings gains we estimate persist beyond an individual's late 20's. Measuring treatment effects in 2019, when these cohorts are in their mid- to late-thirties yields similar effects in percentile ranks (see Appendix Figure A.11).

[28] From the government's perspective, the earnings gains imply roughly $9,750 in additional federal tax revenue per student. The program would pay for itself, with a marginal value of public funds approaching infinity, as long as the additional services CIS students consume cost the government less than about $6,000 per student. Absent data on those costs, we cannot say whether the threshold is met. The threshold is conservative: $1,260 is the cost of case-managing one student, and case-managed students are probably the most expensive to serve, yet we apply that cost to every student.

impacts operate through measured achievement—a question that likely varies across interventions and is often not possible to evaluate directly. CIS provides an unusual opportunity to assess this assumption, because we observe both short-run and long-run outcomes for the same students.

A priori, the link between short-run test-score gains and long-run impacts could run in either direction. Early test-score gains sometimes fade out (see Cascio and Staiger, 2012, for a review), which would lead test-score-based projections to overstate long-run effects. At the same time, programs can raise long-run attainment through improvements in engagement, motivation, and behavior that test scores only partially capture. In that case, projections based on test scores alone would understate the true impact (e.g., Anders, Barr and Smith, 2023; Chetty et al., 2011; Cortes, Goodman and Nomi, 2015; Gray-Lobe, Pathak and Walters, 2023; Jackson, 2018). Which pattern holds is an empirical question. We answer it for CIS by constructing predicted long-run outcomes from students' short-run measures and comparing the treatment effect on these predictions to the actual long-run effect.

We focus on high school graduation because our approach requires observing students from elementary school through a long-run outcome, and graduation is available for enough cohorts to make this analysis feasible. We predict each student's probability of graduating from high school using contemporaneous measures, estimate the effect of CIS on these predictions using our standard event-study framework, and compare the predicted effect with the program's actual impact on graduation. We generate these predictions using a gradient-boosted decision-tree model, similar to the model used to predict case management in Section IV. We estimate separate models for grades 6–8 using all students statewide, not only those in CIS schools, so that the resulting weights reflect how strongly each measure predicts graduation in the broader student population (see Appendix Figure A.12 for details).

We begin with the predictor most often used in such projections: test scores. We form two test-score-based predictions. The first uses the observational relationship between test scores and graduation, estimated by regressing graduation on test scores by grade in statewide data; the second allows for a flexible, possibly non-linear relationship by training the decision-tree model on test scores alone. Both understate the program's long-run effect. As shown on the right-hand side of Figure XII Panel *A*, the predicted impacts are far smaller than the true 3.1 pp effect of CIS on high school graduation (SE = 1.0 pp); the decision-tree version is only 46% as large.

Adding behavioral measures closes most of this gap. We expand the predictor set to include not only standardized math scores but also attendance and disciplinary outcomes (in- and out-of-school suspensions and the number of infractions), so the prediction reflects both academic performance and behavioral measures that may capture non-cognitive traits relevant for long-run success. Figure XII Panel *B* plots the event study for predicted graduation alongside the true effect: predicted graduation rises after CIS implementation in step with true graduation. The

corresponding difference-in-differences in Panel *A* yields a predicted effect of 2.4 pp (SE = 0.6 pp), nearly 80% of the true impact. Test scores and the cognitive skills they reflect explain a meaningful share of CIS's long-run effect, while behavioral measures account for much of the remainder. For low-risk students the prediction again tracks the true effect. We estimate a predicted effect on graduation of 0.3 pp (SE = 0.5 pp), consistent with the modest changes to test scores and behavior for these students. Explaining the program's long-run gains requires looking beyond test scores to broader dimensions of student development.

These results have a practical implication for efforts to forecast long-run impacts from short-run data. For CIS, test scores alone are an incomplete statistical surrogate for long-run attainment, whereas the full set of contemporaneous academic and behavioral measures performs well. The lesson is not that test-score-based projections are misguided, but that the choice of surrogates should be matched to the channels through which a given program works. Test-score-based projections are well-suited to programs whose long-run benefits operate primarily through achievement (e.g., Angrist et al., 2016). More broadly, our findings suggest that rich short-run administrative data hold promise for forecasting long-run program impacts in real time, without waiting years or decades for follow-up outcomes.

Personalization in the CIS Model

Site coordinators personalize support to each student's needs, deciding whether to prioritize academic tutoring, mentoring, counseling, or attendance interventions. Personalization could produce service portfolios well matched to each student's challenges, or it could leave some students worse off than a different mix would. We ask whether students with different needs receive different services, whether they improve most on the dimensions where they were struggling, and whether they go on to achieve similar long-run gains. We cannot test whether the match is optimal. That would require experimental variation in service assignment and estimates of how the effects of each service vary across students.

We first document how coordinators decide which services students receive. Figure XIII Panel *A* plots the relationship between students' prior academic performance and the likelihood of being targeted for academic or behavioral support, controlling for the predicted probability of case management.[29] Students with lower elementary-school math scores are substantially more likely to be targeted for academic support: 94% of students in the bottom ventile of prior math scores are flagged for academics compared to 57% in the top ventile. Conversely, students with relatively strong academic performance—who appear in case management for other reasons—are

[29]Site coordinators assign students in the case-management database any combination of flags for academics, attendance, behavior, and social services. We define a student as flagged for academics (behavior) if their targeted areas include academics (behavior).

more likely to be targeted for behavioral support. Students in the top ventile of math scores are more than 10 pp more likely to be flagged for behavioral issues than those in the bottom ventile.

These distinctions translate into meaningful differences in the set of services students ultimately receive. In Figure XIII Panel *B*, we plot the share of services these students receive that are academic in nature.[30] Students with lower average standardized math scores in grades 3–5 receive a higher fraction of academic services than peers with stronger prior performance, even after conditioning on predicted case-management risk. The pattern holds within schools and years: coordinators personalize services across students in the same cohort. Guided by this relationship, we classify high-risk students with below-median elementary-school academic performance as those with primarily academic needs, and high-risk students with above-median performance as those with primarily non-academic (e.g., behavioral or engagement-related) needs. We then compare the short- and long-run impacts of CIS across these two groups.

CIS has a larger impact on standardized test scores for students with primarily academic needs. We estimate a modified version of the difference-in-differences specification in Equation 2 for the high-risk sample, including a linear control for predicted case-management risk, an indicator for below-median elementary-school academic performance, and interactions of these controls with the post-period indicators. The resulting coefficients for each group are shown in Figure XIII Panel *C*. The estimated effect on math scores for students with academic needs is nearly twice as large as that for students with primarily non-academic needs (0.103 vs. 0.056 SD; t-statistic on difference = 3.7). The pattern reverses for non-academic outcomes: students with primarily non-academic needs show larger short-run gains in attendance (reduced absences and suspensions), gaining roughly 0.66 additional days in class versus roughly 0.48 for students with academic needs (t-statistic on difference = 1.9). Personalization thus works in both directions, with each group improving most on the outcome its services were designed to target.

Despite the short-run differences, the long-run effects are nearly identical: graduation rises 2.8 pp for students with academic needs and 2.9 pp for students with non-academic needs (Figure XIII Panel *D*). Students with different needs receive different services and improve most where they were struggling, yet they graduate at the same higher rate.

[30] The set of service codes in the case-management microdata varies slightly by year. As an example, in 2018–2019 we classify as academic services: Academic Readiness Interventions, Parent Conferences with Teachers, Academic Skills Development and Support, Credit Recovery Assistance, Academic Monitoring, Homework Assistance, Language (ESL/ELL) Support, Tutoring, Student/Teacher Conferences, Support for Students in Special Education, School Supplies, and General – Academic Enhancement and Support.

VII Conclusion

Gaps in outcomes between children from low-income families and their more affluent peers emerge in school and persist through educational attainment and the labor market (Chetty et al., 2014; Solon, 1992; Black, Devereux and Salvanes, 2005; Chetty et al., 2020). We show that providing personalized, coordinated supports in high-poverty schools can meaningfully improve outcomes in the classroom and over the long run. Communities In Schools (CIS), the largest integrated student support program in the country, raises test scores, boosts high school graduation, and improves early-career earnings for students most at risk of falling behind. These impacts arise through both cognitive and non-cognitive channels, not through academic performance alone. Coordinators personalize services to each student's situation, yet students with different needs achieve comparable long-run gains.

Traditional resources—housing vouchers, cash transfers, job training, college-access programs—tend to deliver larger and more durable gains when paired with a dedicated person who provides sustained, personalized support and helps individuals navigate barriers. This finding recurs across many domains in recent randomized trials (Bergman et al., 2024; Caspi and Rafkin, 2025; Espinosa et al., 2024; Finkelstein and Notowidigdo, 2019; Katz et al., 2022; Weiss et al., 2019). CIS shows that this relational model can work at scale and early in life. For children in high-poverty schools, coordinated, personalized support improves educational attainment and early-career earnings in lasting ways. The largest gains accrue to the students who need it most.

Most high-poverty neighborhoods still lack a CIS school, leaving substantial room for expansion. CIS has already demonstrated its capacity to scale, serving 2 million students nationwide. Because independent local nonprofit affiliates deploy site coordinators, CIS can expand without requiring changes to schools' existing staffing, budgets, or governance. The concentration of CIS's benefits among the most disadvantaged students also provides a framework for targeting expansion.

As a back-of-the-envelope illustration, we simulate a targeted rollout that prioritizes schools with the largest predicted effects on high school graduation. We compute those effects from each school's distribution of students across risk quartiles and the estimated treatment effect on graduation for each quartile (Figure XIV Panel *A*; see also Appendix Figure A.6). CIS programs currently reach about 20% of middle school students in Texas, an allocation that captures roughly 24% of the total possible gains in high school graduation. Expanding coverage to 40% by serving the schools with the next-highest predicted impacts would capture roughly %. The calculation assumes that effects estimated at currently served schools carry over to schools at the margin of expansion. Our design does not identify marginal treatment effects directly, but Panel *B* compares schools reached earlier and later in an affiliate's rollout (Cohodes, Setren and Walters, 2021).

Returns fall somewhat as the program expands, as the simulation implies, though we cannot rule out equal treatment effects at earlier and later sites. The benefits of CIS are not confined to its earliest sites. Roughly three-quarters of the achievable gains in graduation remain unrealized, and they lie in the schools with the most students at risk of falling behind.

References

Alan, Sule, and Elif Kubilay. 2025. "Empowering adolescents to transform schools: Lessons from a behavioral targeting." *American Economic Review*, 115(2): 365–407.

American School Counselor Association. 2024. "Student-to-School-Counselor Ratio 2023–2024." *ASCA, analysis of NCES Common Core of Data*, Recommended ratio 250:1; national average approximately 376:1.

Anders, John, Andrew C Barr, and Alexander A Smith. 2023. "The effect of early childhood education on adult criminality: Evidence from the 1960s through 1990s." *American Economic Journal: Economic Policy*, 15(1): 37–69.

Angrist, Joshua D., Parag A. Pathak, and Christopher R. Walters. 2013. "Explaining charter school effectiveness." *American Economic Journal: Applied Economics*, 5(4): 1–27.

Angrist, Joshua D., Peter Hull, Russell Legate-Yang, Parag A. Pathak, and Christopher R. Walters. 2025. "Putting School Surveys to the Test." National Bureau of Economic Research Working Paper 33622.

Angrist, Joshua D., Sarah R. Cohodes, Susan M. Dynarski, Parag A. Pathak, and Christopher R. Walters. 2016. "Stand and Deliver: Effects of Boston's Charter High Schools on College Preparation, Entry, and Choice." *Journal of Labor Economics*, 34(2): 275–318.

Angrist, Joshua D, Victor Lavy, Jetson Leder-Luis, and Adi Shany. 2019. "Maimonides rule redux." *American Economic Review: Insights*, 1(3): 309–324.

Athey, Susan, Raj Chetty, Guido W Imbens, and Hyunseung Kang. 2025. "The surrogate index: Combining short-term proxies to estimate long-term treatment effects more rapidly and precisely." *Review of Economic Studies*, rdaf087.

Attea, Kate, and Alan Tuck. 2005. "Communities In Schools: Propelling a National Network to the Next Level." The Bridgespan Group. Accessed December 3, 2025.

Barr, Andrew, and Benjamin Castleman. 2025. "Increasing Degree Attainment among Low-Income Students: The Role of Intensive Advising and College Quality." *American Economic Review*, 115(11): 4075–4103.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F Katz, and Christopher Palmer. 2024. "Creating moves to opportunity: Experimental evidence on barriers to neighborhood choice." *American Economic Review*, 114(5): 1281–1337.

Bhuller, Manudeep, Magne Mogstad, and Kjell G Salvanes. 2017. "Life-cycle earnings, education premiums, and internal rates of return." *Journal of Labor Economics*, 35(4): 993–1030.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2005. "Why the apple doesn't fall far: Understanding intergenerational transmission of human capital." *American Economic Review*, 95(1): 437–449.

Card, David. 1999. "The causal effect of education on earnings." *Handbook of Labor Economics*, 3: 1801–1863.

Carrell, Scott E., and Mark Hoekstra. 2014. "Are School Counselors an Effective Education Input?" *Economics Letters*, 125(1): 66–69.

Cascio, Elizabeth U, and Douglas O Staiger. 2012. "Knowledge, tests, and fadeout in educational interventions." National Bureau of Economic Research Working Paper 18038.

Caspi, Aviv, and Charlie Rafkin. 2025. "Legal assistance for evictions."

Cattan, Sarah, Kjell G Salvanes, and Emma Tominey. 2025. "First-Generation Elite: The Role of School Social Networks." *American Economic Review*, 115(12): 4369–4403.

Chetty, Raj, David J Deming, and John N Friedman. 2025. "Diversifying society's leaders? The determinants and causal effects of admission to highly selective private colleges." *The Quarterly Journal of Economics*, qjaf050.

Chetty, Raj, John N Friedman, and Jonah E Rockoff. 2014. "Measuring the impacts of teachers II: Teacher value-added and student outcomes in adulthood." *American Economic Review*, 104(9): 2633–2679.

Chetty, Raj, John N Friedman, Emmanuel Saez, Nicholas Turner, and Danny Yagan. 2020. "Income segregation and intergenerational mobility across colleges in the United States." *The Quarterly Journal of Economics*, 135(3): 1567–1633.

Chetty, Raj, John N Friedman, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2025. "The opportunity atlas: Mapping the childhood roots of social mobility." *The American Economic Review*, forthcoming.

Chetty, Raj, John N Friedman, Nathaniel Hilger, Emmanuel Saez, Diane Whitmore Schanzenbach, and Danny Yagan. 2011. "How does your kindergarten classroom affect your earnings? Evidence from Project STAR." *The Quarterly Journal of Economics*, 126(4): 1593–1660.

Chetty, Raj, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, et al. 2022. "Social Capital I: Measurement and Associations with Economic Mobility." *Nature*, 608(7921): 108–121.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the land of opportunity? The geography of intergenerational mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623.

Chetty, Raj, Will Dobbie, Benjamin Goldman, Sonya R Porter, and Crystal S Yang. 2026. "Changing opportunity: Sociological mechanisms underlying growing class gaps and shrinking race gaps in economic mobility." *The Quarterly Journal of Economics*, qjaf057.

CIS. 2023. "Communities in Schools: Databook 2023-24." *https://www.communitiesinschools.org/articles/article/2023-2024-data-book-and-state-profiles/*.

CIS. 2025*a*. "Our Approach." *https://www.communitiesinschools.org/our-approach/*, Accessed December 3, 2025.

CIS. 2025*b*. "Site Coordinators." *https://www.communitiesinschools.org/site-coordinators/*, Accessed December 3, 2025.

CIS. 2025*c*. "Start a Local Affiliate." *https://www.communitiesinschools.org/network/starting-affiliate/*, Accessed December 3, 2025.

CIS of Houston. 2024. "Annual Financial and Compliance Report for the Year Ended August 31, 2023." *https://projects.propublica.org/nonprofits/download-audit?download=true&filename=2023-08-GSAFAC-0000030536*.

Cohodes, Sarah R., Elizabeth M. Setren, and Christopher R. Walters. 2021. "Can Successful Schools Replicate? Scaling Up Boston's Charter School Sector." *American Economic Journal: Economic Policy*, 13(1): 138–167.

Corrin, William, Leigh Parise, Oscar Cerna, Zeest Haider, and Marie-Andrée Somers. 2015. "Case management for students at risk of dropping out: Implementation and interim impact findings from the communities in schools evaluation." *Available at SSRN 2609366*.

Cortes, Kalena E, Joshua S Goodman, and Takako Nomi. 2015. "Intensive math instruction and educational attainment: Long-run impacts of double-dose algebra." *Journal of Human Resources*, 50(1): 108–158.

Dee, Thomas, Chloe R Gibbs, Rebecca Margolit-Chan, Lillian Morrell, and James X Sullivan. 2026. "Unpacking Treatment Effect Heterogeneity in a Large-Scale Dropout-Prevention Experiment." National Bureau of Economic Research Working Paper 35558.

Department of Education. 2026. "Chronic Absenteeism." *https://www.ed.gov/teaching-and-administration/supporting-students/chronic-absenteeism*, Accessed July 17, 2026.

Doleac, Jennifer L. 2019. "Wrap-around services don't improve prisoner reentry outcomes." *Journal of Policy Analysis and Management*, 38(2): 508–514.

Espinosa, Javier, William N Evans, David C Phillips, and Tim Spilde. 2024. "How Do Holistic Wrap-Around Anti-Poverty Programs Affect Employment and Individualized Outcomes?" National Bureau of Economic Research Working Paper 32911.

Evans, William N, Melissa S Kearney, Brendan Perry, and James X Sullivan. 2020. "Increasing community college completion rates among low-income students: Evidence from a randomized controlled trial evaluation of a case-management intervention." *Journal of Policy Analysis and Management*, 39(4): 930–965.

Falk, Armin, Fabian Kosse, and Pia Pinger. 2026. "Mentoring and schooling decisions: Causal evidence." *Journal of Political Economy*, 134(1): 366–396.

Figlio, David. 2015. "Experimental evidence of the effects of the Communities In Schools of Chicago partnership program on student achievement."

Finkelstein, Amy, and Matthew J Notowidigdo. 2019. "Take-up and targeting: Experimental evidence from SNAP." *The Quarterly Journal of Economics*, 134(3): 1505–1556.

Ganimian, Alejandro J., Karthik Muralidharan, and Christopher R. Walters. 2024. "Augmenting State Capacity for Child Development: Experimental Evidence from India." *Journal of Political Economy*, 132(5): 1565–1602.

Gray-Lobe, Guthrie, Parag A Pathak, and Christopher R Walters. 2023. "The long-term effects of universal preschool in Boston." *The Quarterly Journal of Economics*, 138(1): 363–411.

Guryan, Jonathan, Jens Ludwig, Monica P Bhatt, Philip J Cook, Jonathan MV Davis, Kenneth Dodge, George Farkas, Roland G Fryer Jr, Susan Mayer, Harold Pollack, et al.

2023. “Not too late: Improving academic outcomes among adolescents.” *American Economic Review*, 113(3): 738–765.

Hanushek, Eric A. 2011. “The economic value of higher teacher quality.” *Economics of Education Review*, 30(3): 466–479.

He, Haibo, and Edwardo A. Garcia. 2009. “Learning from Imbalanced Data.” *IEEE Trans. on Knowl. and Data Eng.*, 21(9): 1263–1284.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. “A Unified Welfare Analysis of Government Policies.” *The Quarterly Journal of Economics*, 135(3): 1209–1318.

Jackson, C Kirabo. 2018. “What do test scores miss? The importance of teacher effects on non–test score outcomes.” *Journal of Political Economy*, 126(5): 2072–2107.

Jackson, C Kirabo, Rucker C Johnson, and Claudia Persico. 2016. “The effects of school spending on educational and economic outcomes: Evidence from school finance reforms.” *The Quarterly Journal of Economics*, 131(1): 157–218.

Johnston, Kennedy, Jonathan Meer, and Danila Serra. 2026. “Shaping Major Choice: The Role of High School Counselors.” National Bureau of Economic Research NBER Working Paper 35365.

Katz, Lawrence F., Jonathan Roth, Richard Hendra, and Kelsey Schaberg. 2022. “Why Do Sectoral Employment Programs Work? Lessons from WorkAdvance.” *Journal of Labor Economics*, 40(S1): S249–S291.

Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, and Tie-Yan Liu. 2017. “LightGBM: A Highly Efficient Gradient Boosting Decision Tree.” *Advances in Neural Information Processing Systems 30 (NIPS 2017)*, 3146–3154.

Kim, Donggwan, Dmitry Mitrofanov, Qingsong Wen, and Tianlong Xu. 2025. “Generative AI Can Improve Performance and Engagement without Harming Learning.”

Kline, Patrick, and Christopher R. Walters. 2016. “Evaluating Public Programs with Close Substitutes: The Case of Head Start.” *The Quarterly Journal of Economics*, 131(4): 1795–1848.

Kosse, Fabian, Thomas Deckers, Pia Pinger, Hannah Schildberg-Hörisch, and Armin Falk. 2020. “The formation of prosociality: causal evidence on the role of social environment.” *Journal of Political Economy*, 128(2): 434–467.

Krueger, Alan B. 1999. "Experimental estimates of education production functions." *The quarterly journal of economics*, 114(2): 497–532.

Lavecchia, Adam M, Philip Oreopoulos, and Robert S Brown. 2020. "Long-run effects from comprehensive student support: Evidence from pathways to education." *American Economic Review: Insights*, 2(2): 209–224.

Malkus, Nat. 2024. "Long COVID for Public Schools: Chronic Absenteeism Before and After the Pandemic."

Maponics. 2017. "Maponics School Boundaries."

McShane, Michael Q. 2019. "Supporting Students Outside the Classroom: Can Wraparound Services Improve Academic Performance?" *Education Next*, 19(3): 38–45.

Mulhern, Christine. 2023. "Beyond Teachers: Estimating Individual School Counselors' Effects on Educational Attainment." *American Economic Review*, 113(11): 2846–2893.

Muralidharan, Karthik, Abhijeet Singh, and Alejandro J Ganimian. 2019. "Disrupting education? Experimental evidence on technology-aided instruction in India." *American Economic Review*, 109(4): 1426–1460.

Nickow, Andre, Philip Oreopoulos, and Vincent Quan. 2020. "The impressive effects of tutoring on prek-12 learning: A systematic review and meta-analysis of the experimental evidence." National Bureau of Economic Research Working Paper 27476.

Office of Head Start. 2025. "Head Start Program Facts: Fiscal Year 2024." https://headstart.gov/program-data/article/head-start-program-facts-fiscal-year-2024, Accessed: 2025-11-18.

Parise, Leigh, and Farhana Hossain. 2017. "Using Research Evidence to Strengthen Support for At-Risk Students: A Case Study of Communities in Schools (Issue Focus Infographic)." MDRC, New York, NY.

Parise, Leigh, 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." *MDRC, Project: Communities In Schools*.

Porowski, Allan, and Aikaterini Passa. 2011. "The effect of communities in schools on high school dropout and graduation rates: Results from a multiyear, school-level quasi-experimental study." *Journal of Education for Students Placed at Risk*, 16(1): 24–37.

Posner, Joe, James Watson, and Nishtha Shanti. 2026. "This Top Secret Kindness Club Fights Poverty." *https://www.youtube.com/watch?v=u8k3q-azV8w*, DREAM ON, episode 1. Produced at Fire-Works in collaboration with the EdRedesign Lab, Harvard Graduate School of Education. Released May 5, 2026.

Prentice, Ross L. 1989. "Surrogate endpoints in clinical trials: definition and operational criteria." *Statistics in Medicine*, 8(4): 431–440.

Putnam, Robert D. 2016. *Our Kids: The American Dream in Crisis.* Simon and Schuster.

Robinson, Carly D, Biraj Bisht, and Susanna Loeb. 2025. "The inequity of opt-in educational resources and an intervention to increase equitable access." *Educational Researcher*, 54(6): 328–338.

Solon, Gary. 1992. "Intergenerational income mobility in the United States." *The American Economic Review*, 393–408.

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

Sparks, Daniel, and Christine Mulhern. 2026. "Expanding School Counseling: The Impacts of California Funding Changes." *Economics of Education Review*, 110: 102754.

Stevens, David W. 2007. "Employment that is not covered by state unemployment insurance laws." Center for Economic Studies, US Census Bureau.

Sun, Liyang, and Sarah Abraham. 2021. "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects." *Journal of Econometrics*, 225(2): 175–199.

Texas Education Agency. 2013. "PEIMS Data Standards: At-Risk Indicator Code (E0919)." *http://ritter.tea.state.tx.us/peims/standards/1314/e0919.html*, Accessed: 2025-09-17.

Texas Education Agency. 2021. "Commissioner's Rules Concerning the Communities In Schools Program, Chapter 89, Subchapter EE." *https://tea.texas.gov/about-tea/laws-and-rules/commissioner-rules-tac/coe-tac-currently-in-effect/ch089ee.pdf*, Texas Administrative Code, Title 19, Part 2. Adopted July 13, 2021.

Texas Education Agency. 2022. "Communities In Schools." *https://tea.texas.gov/texas-schools/support-for-at-risk-schools-and-students/communities-in-schools*, Accessed: 2025-01-29.

Texas Legislative Budget Board. 2016. "Overview of Performance and Effectiveness Information for Texas Education Agency Non-Foundation School Programs." Texas Legislative Budget Board Presentation ID 3330, Austin, TX. Presented to the House Appropriations Committee, Subcommittee on Article III. See slide 11: "Average Expenditure Per CIS Participant" (state and local costs per case-managed student).

U.S. Department of Education, National Center for Education Statistics. 2022. "Table 309.30. Residence and migration of all first-time degree/certificate-seeking undergraduates in 4-year degree-granting postsecondary institutions who graduated from high school in the previous 12 months, by state or jurisdiction: Fall 2020." U.S. Department of Education. Digest of Education Statistics 2021, January 2022. Data source: Integrated Postsecondary Education Data System (IPEDS), Spring 2021, Fall Enrollment component. Available at https://nces.ed.gov/programs/digest/d21/tables/dt21_309.30.asp.

Walters, Christopher R. 2018. "The Demand for Effective Charter Schools." *Journal of Political Economy*, 126(6): 2179–2223.

Weiss, Michael J, Alyssa Ratledge, Colleen Sommo, and Himani Gupta. 2019. "Supporting community college students from start to degree completion: Long-term evidence from a randomized trial of CUNY's ASAP." *American Economic Journal: Applied Economics*, 11(3): 253–297.

You, Shijun, Wei Fu, and Shin-Yi Chou. 2025. "Local Problems, Local Solutions: Evidence from the Drug-Free Communities Program on Youth Development." National Bureau of Economic Research Working Paper 34182.

Zimmerman, Seth D. 2014. "The Returns to College Admission for Academically Marginal Students." *Journal of Labor Economics*, 32(4): 711–754. Publisher: The University of Chicago Press.

TABLE I: Summary Statistics

	All Students (1)	Ever in School with CIS Program (2)	Ever in MS with CIS Program (3)
A. Student Characteristics			
Fraction Male	0.520	0.514	0.495
Fraction White	0.360	0.262	0.239
Fraction Hispanic	0.454	0.551	0.591
Fraction Black	0.142	0.157	0.143
Fraction Other Race/Ethnicity	0.044	0.030	0.028
Fraction Student-Years FRPL	0.470	0.522	0.515
Fraction Student-Years Dropout-Risk	0.450	0.497	0.511
B. Student Outcomes			
Standardized Math Score in 6th Grade	0.004	-0.133	-0.156
Absence Rate in 6th Grade	0.037	0.039	0.040
Fraction Ever Receive In-School Suspension in 6th Grade	0.129	0.169	0.169
Fraction Ever Receive Out-of-School Suspension in 6th Grade	0.061	0.087	0.090
Fraction Ever Graduate from HS	0.738	0.732	0.727
Fraction Ever Enroll in TX College Ages 18-22	0.501	0.477	0.479
Fraction Appear in Earnings Data	0.820	0.846	0.833
Average Earnings at Age 27 ($)	33707	32356	32548
Employment Rate at Age 27	0.605	0.624	0.609
Average Individual Earnings Rank at Age 27	0.497	0.490	0.488
C. School Characteristics			
Average Number of Students	991	1368	903
Average Number of Teachers	65	91	61
Fraction Teachers Male	0.229	0.309	0.312
Fraction Teachers White	0.670	0.610	0.582
Fraction Teachers Hispanic	0.217	0.294	0.306
Fraction Teachers with Advanced Degree	0.266	0.280	0.258
Average Teacher Tenure (yrs)	7.643	7.584	7.428
Average Teacher Pay ($)	70078	71675	70678
Number of Student-Years (1,000s)	109172	27222	15547
Number of Students (1,000s)	16165	3332	1639
Number of Students in HS Grad Sample (1,000s)	9714	2248	1110
Number of Students in College Sample (1,000s)	8311	1907	911
Number of Students in Earnings Sample (1,000s)	5960	1351	614
Number of Schools	10682	613	208
Number of School-Years	189748	12941	4365

Notes: This table reports summary statistics for three groups: (1) all students in our data, (2) students who ever attend a school with a CIS program in 2019, and (3) students who ever attend a middle school with a CIS program in 2019. All statistics use the maximum available data on students enrolled in grades 1-12 between 1994 and 2019. Panel A summarizes student demographic characteristics. Panel B reports student outcomes. High school graduation is measured among students who are age 19 or older by 2019; the corresponding sample size and those for subsequent outcomes are listed at the bottom of the table. College enrollment is measured among students age 22 or older by 2019. The fraction of students in the earnings data refers to the share with non-zero earnings at any point between ages 18 and 25. The employment rate is the fraction with non-zero earnings at age 27, conditional on being in the earnings data. All earnings outcomes are measured among students who are age 27 or older by 2019. Panel C presents summary statistics on school characteristics. "Teachers with advanced degrees" refers to those with a master's degree or higher. "Teacher tenure" reflects years of continuous employment at the current school. All monetary values are expressed in 2025 dollars.

TABLE II: Characteristics of Low- and High-Risk Students

	Low-Risk in CIS School (1)	High-Risk in CIS School (2)	Low-Risk in Non-CIS School (3)	High-Risk in Non-CIS School (4)
A. Student Characteristics				
Fraction Male	0.553	0.377	0.550	0.370
Fraction White	0.551	0.149	0.671	0.209
Fraction Hispanic	0.348	0.618	0.189	0.524
Fraction Black	0.031	0.222	0.033	0.254
Fraction Other Race/Ethnicity	0.070	0.012	0.106	0.013
Fraction Student-Years FRPL	0.092	0.749	0.071	0.776
Fraction Student-Years Dropout-Risk	0.163	0.717	0.120	0.694
B. Student Outcomes				
Standardized Math Score in 6th Grade	0.695	-0.743	0.773	-0.679
Absence Rate in 6th Grade	0.028	0.048	0.026	0.044
Fraction Ever Receive In-School Suspension in 6th Grade	0.060	0.244	0.046	0.225
Fraction Ever Receive Out-of-School Suspension in 6th Grade	0.020	0.132	0.013	0.113
Fraction Ever Graduate from HS	0.863	0.589	0.868	0.607
Fraction Ever Enroll in TX College Ages 18-22	0.721	0.315	0.744	0.321
Fraction Appear in Earnings Data	0.871	0.810	0.873	0.797
Average Earnings at Age 27 ($)	45546	21529	47143	21960
Employment Rate at Age 27	0.643	0.588	0.641	0.576
Average Individual Earnings Rank at Age 27	0.559	0.421	0.561	0.422
Number of Student-Years (1,000s)	3678	6073	20028	17512
Number of Students (1,000s)	318	533	1755	1544
Number of Students in HS Grad Sample (1,000s)	203	325	1091	988
Number of Students in College Sample (1,000s)	164	251	874	772
Number of Students in Earnings Sample (1,000s)	100	140	537	428

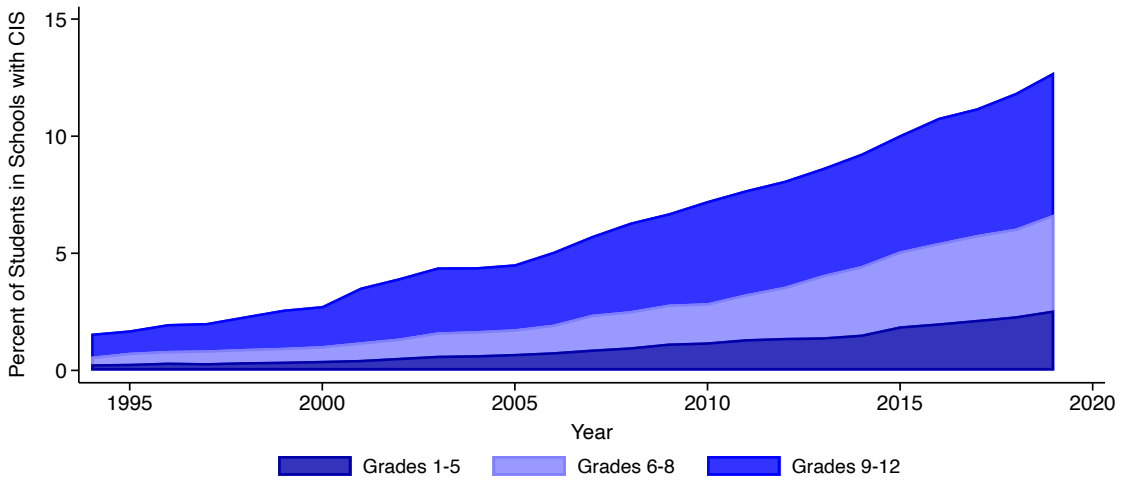
Notes: This table reproduces the summary statistics in Table I, separately for students who are high- versus low-risk and who attend CIS versus non-CIS schools. High-risk students are those in the top quartile of the case management prediction model described in Section IV and shown in Figure III Panel C. See the notes to Table I for additional details on the construction of these statistics.

TABLE III: Effect of CIS for Low- vs. High-Risk Students

	High-Risk	Low-Risk	Difference
	(1)	(2)	(3)
Years of CIS Exposure			
$\beta_{post,3-5}$	1.699	1.745	-0.046
	(0.025)	(0.038)	(0.034)
Pre-Period Mean	0.0	0.0	
Observations (1,000s)	455	329	784
Standardized Math Scores			
$\beta_{post,3-5}$	0.106	0.011	0.094
	(0.023)	(0.024)	(0.029)
3 Years of Exposure	0.183	0.020	
	(0.039)	(0.042)	
Pre-Period Mean	-0.7	0.7	
Observations (1,000s)	417	286	703
True HS Graduation			
$\beta_{post,3-5}$	3.088	-1.082	4.170
	(1.038)	(0.676)	(1.189)
3 Years of Exposure	3.445	-1.103	
	(1.158)	(0.696)	
Pre-Period Mean	64.5	89.3	
Observations (1,000s)	429	316	745
2-Year College Attendance			
$\beta_{post,3-5}$	2.654	-3.162	5.815
	(1.795)	(1.858)	(2.387)
3 Years of Exposure	2.881	-2.915	
	(1.931)	(1.681)	
Pre-Period Mean	29.2	52.7	
Observations (1,000s)	161	111	272
Predicted HS Graduation			
$\beta_{post,3-5}$	2.363	0.301	2.062
	(0.551)	(0.403)	(0.617)
Pre-Period Mean	53.9	83.2	
Observations (1,000s)	455	329	784

Notes: This table presents DID coefficients on the effect of CIS on high- and low-risk students, respectively. For each outcome in the horizontal panels we present the DID coefficient $\beta_{post,3-5}$ from equation 2 (or its long-run analog) for high-risk students in column 1 and low-risk students in column (2). In column 3 we present the difference estimated from a modified version of equation 2 (or its long-run analog) estimated on the sample of low- and high-risk students where we fully interact the regression with an indicator for being in the high-risk group. In addition to the DID coefficients, we also present IV coefficients scaled to three years of exposure to the program. Finally, we include the values of the mean of each outcome in the pre-period for reference. See Appendix Table A.4 for the analogous table with standard errors clustered at the school level.

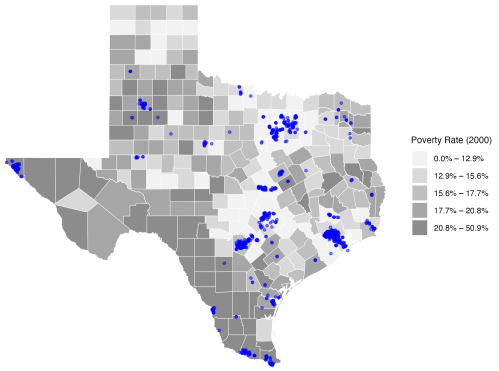
FIGURE I: CIS Program Rollout Across Texas



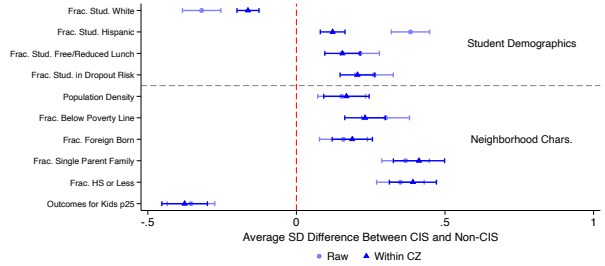
Notes: This figure plots the share of all Texas students enrolled in schools with a CIS program from 1994 to 2019. For each year, we identify the schools with an active CIS program and calculate the number of students enrolled in those schools divided by the total statewide student enrollment. We decompose this total by school level—elementary (grades 1–5), middle (grades 6–8), and high school (grades 9–12)—based on each school's predominant grade span. The shaded regions show the fraction of all Texas students enrolled in CIS elementary, middle, and high schools, respectively, so that the vertical height of each band reflects the share of entire the statewide student population served by CIS at that level and the total height reflects the fraction of all Texas public school students in a school with a CIS program in that year.

FIGURE II: CIS Location Characteristics

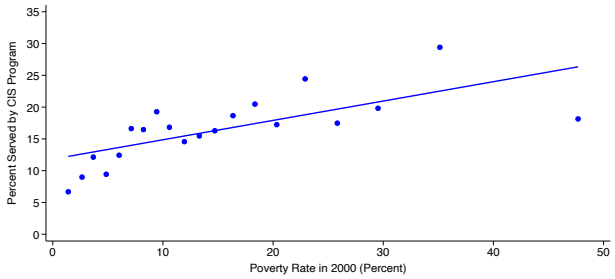
A: Map of CIS Schools in Texas and County-Level Poverty Rates



B: School and Neighborhood Correlates of CIS Programs



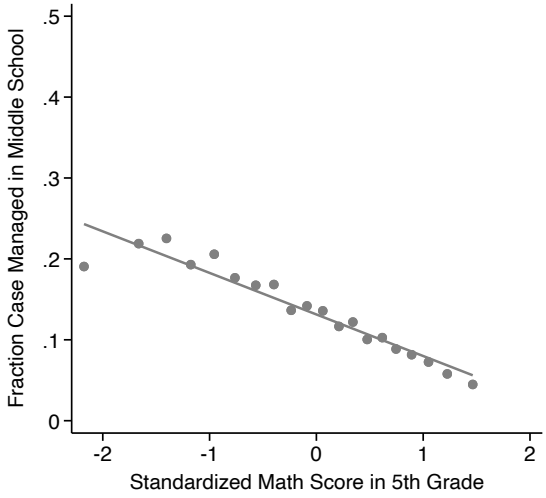
C: Relationship Between CIS Presence and Census Tract-Level Poverty Rates



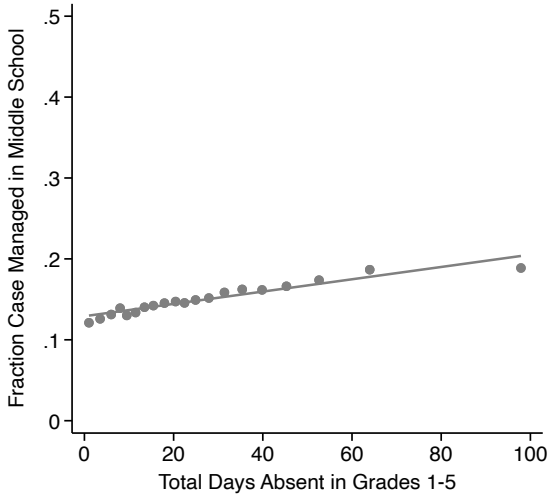
Notes: This figure presents descriptive evidence on the locations and characteristics of CIS programs in Texas. Panel A overlays the locations of CIS programs as of 2019 onto a county-level map of Texas, with counties shaded according to poverty rates from the 2000 Census. Panel B plots differences in school and neighborhood characteristics between schools with and without CIS programs. The light blue series (circles) shows raw differences in standard deviation units, while the dark blue series (triangles) restricts comparisons to schools within the same commuting zone. Student demographic characteristics are measured in the using TEA administrative data described in Section IVA. Neighborhood characteristics are drawn from the 2000 Census tract-level data and assigned to schools using school catchment zones following Chetty et al. (2025). Panel C presents a binned scatterplot of the relationship between 2000 Census tract poverty rates and CIS program coverage in 2019. To construct the figure, we divide Census tracts into 20 equal population-sized bins by poverty rate and plot the fraction of tracts in each bin that are served by a CIS program against the average poverty rate in the bin, along with a line of best fit.

FIGURE III: Predicting CIS Case Management

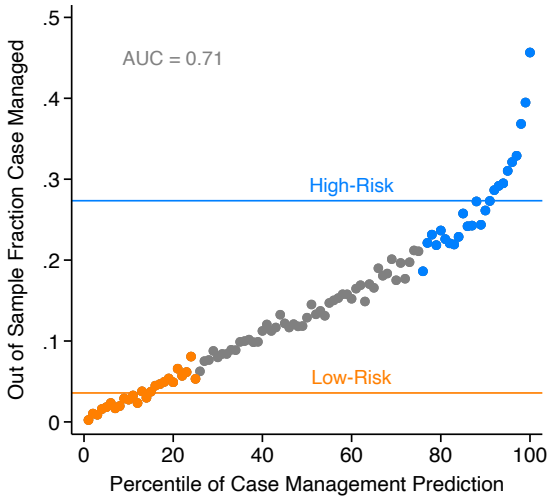
A: Relationship with Test Scores



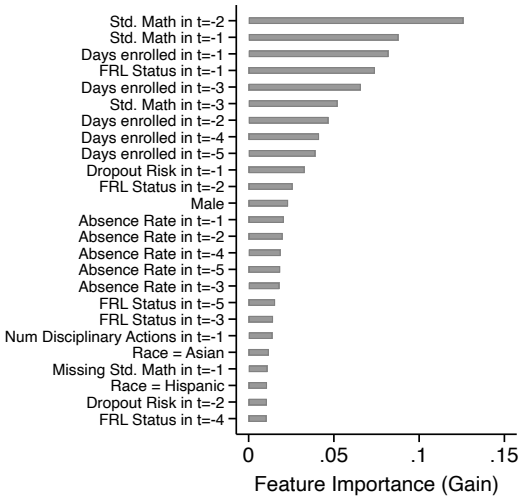
B: Relationship with Attendance



C: Out-of-Sample Accuracy of Case Management Prediction

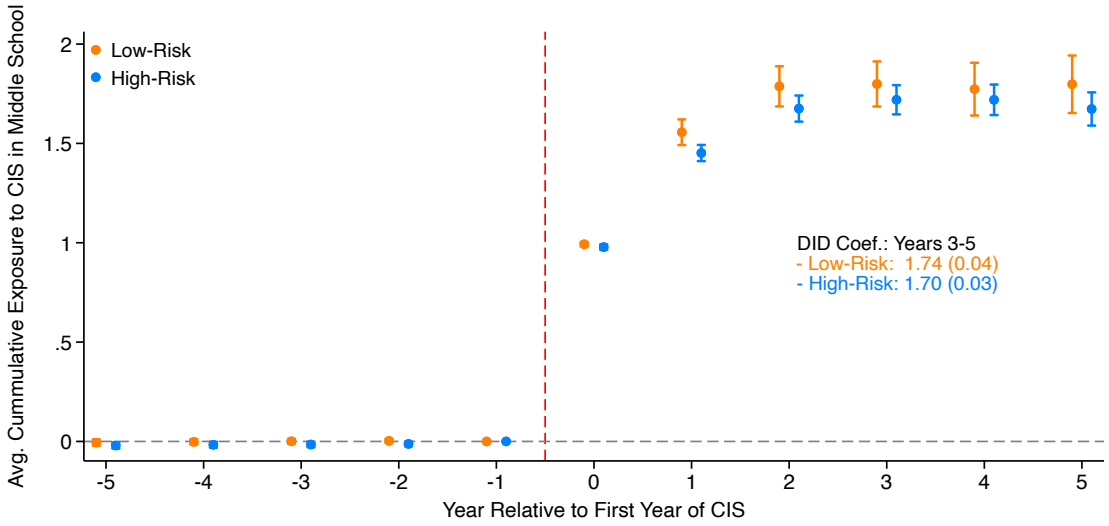


D: Feature Importance – Model Gain



Notes: This figure summarizes the empirical prediction of CIS case management in middle school using pre-treatment student characteristics measured in elementary school. The plots are constructed using a 30% hold-out test set not included in model training. Panel A shows a binned scatterplot of the relationship between standardized 5th grade math scores and the probability of being case managed. Panel B shows the analogous relationship using cumulative absences in grades 1–5. Panel C reports a binned scatterplot of out-of-sample rates of case management by predicted percentile of predicted case management status, based on a gradient boosting model trained on student elementary school outcomes. We classify the top quartile of the predicted distribution as high-risk and the bottom as low-risk and refer to these groups throughout. Panel D presents the top 25 predictors in the GBM ranked by model gain. Predictor variables are measured in years relative to when students first enroll in middle school, ranging from one year prior ($t = -1$) to five years prior ($t = -5$). The model is estimated using data from middle schools with active CIS programs between 2014 and 2019. The dependent variable is an indicator for ever being case managed, as defined in Section III. B, and predictors include student demographics, test scores, attendance, and disciplinary history. See Section IV and Appendix B for additional details on the prediction model.

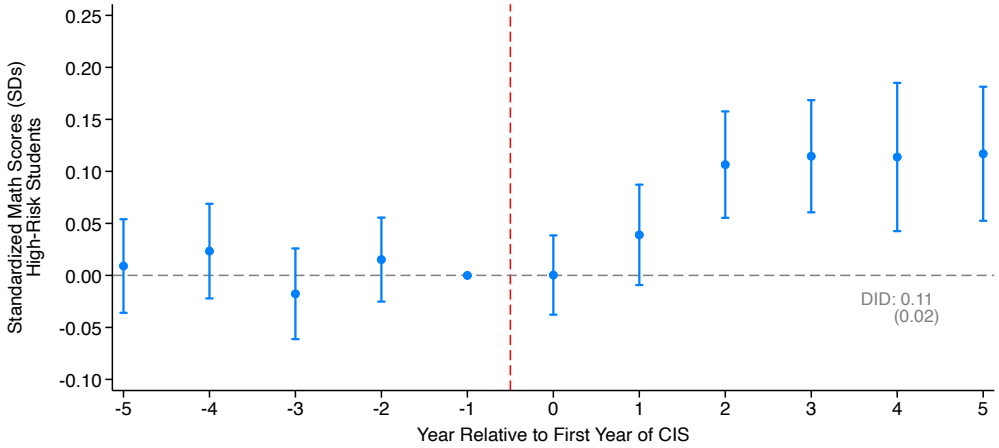
FIGURE IV: Effect of CIS Adoption on Exposure to CIS in Middle School



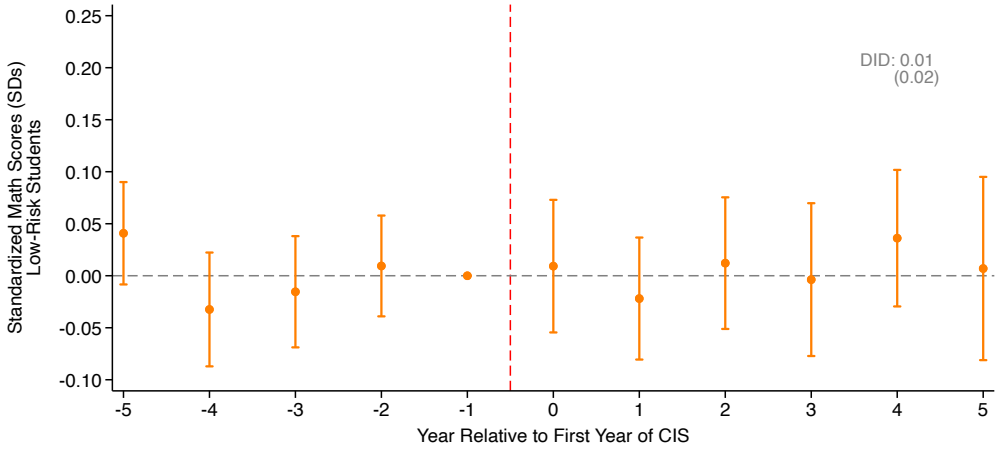
Notes: This figure plots the relationship between time relative to the start of a CIS program and cumulative exposure to CIS in middle school, separately for high-risk (blue) and low-risk (orange) students. See Section IV and the notes to Figure III for details on the assignment to groups. For each group, we estimate Equation 1, comparing treated schools to matched control schools, where the dependent variable is the number of years a student has been exposed to CIS in middle school as of year t. The figure presents event study coefficients estimated using data on students in middle school between 1998 to 2016. In addition to the event study, we estimate a simplified difference-in-differences specification (Equation 2) and report the average difference in CIS exposure between treated and control students over relative years 3 to 5, compared to the pre-treatment period ($\beta_{post,3-5}$). In Table III we report the DID coefficients from a single, fully interacted regression including both high- and low-risk students. Standard errors are clustered at the school-by-birth-cohort level.

FIGURE V: Effect of CIS Adoption on Standardized Math Scores

A: High-Risk Students



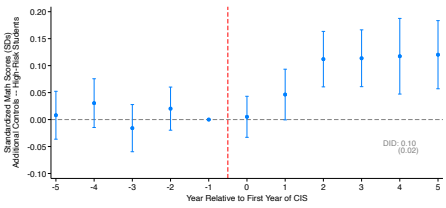
B: Low-Risk Students



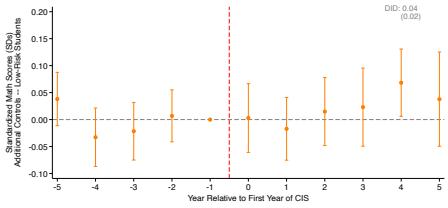
Notes:This figure presents estimates of the effect of CIS adoption on standardized math scores in middle school, separately for high-risk (Panel A) and low-risk (Panel B) students. See Section IV and the notes to Figure III for details on the assignment to groups. For each group, we estimate Equation 1, comparing treated schools to matched control schools, where the outcome is the standardized math score of students in year t, standardized within grade and calendar year. The figure shows event study coefficients estimated using data on middle school standardized test taken between 1998 and 2016. In addition to the event study, we estimate a simplified difference-in-differences specification (Equation 2) and report the average effect on test scores over relative years 3 to 5, compared to the pre-treatment period ($\beta_{post,3-5}$). Difference-in-difference coefficients for the other risk quartiles (medium-high and medium-low) as well as students with missing predictions are reported in Appendix Figure A.2. Standard errors are clustered at the school-by-birth-cohort level.

FIGURE VI: Robustness of Baseline Treatment Effects for Std. Math Scores

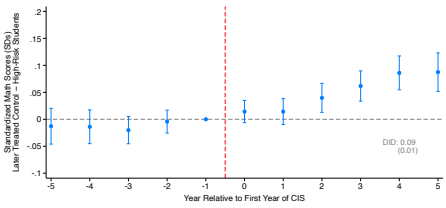
A: Flexible Time Controls by Race and Sex
High-Risk



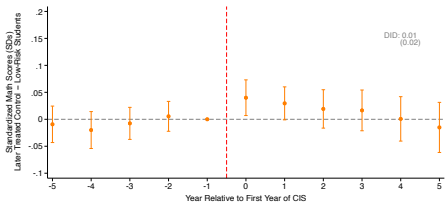
B: Flexible Time Controls by Race and Sex
Low-Risk



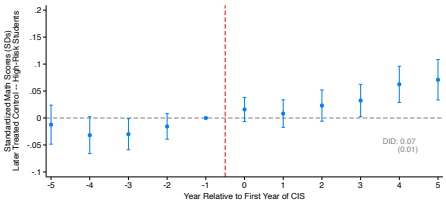
C: Alternate Design: Later Treated Schools
High-Risk



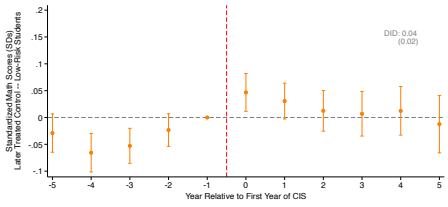
D: Alternate Design: Later Treated Schools
Low-Risk



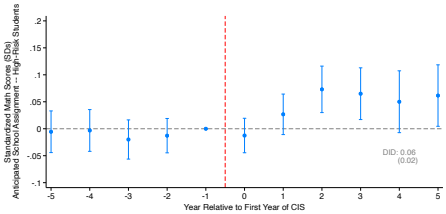
E: Within-CZ Comparison
High-Risk



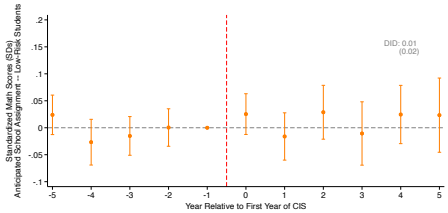
F: Within-CZ Comparison
Low-Risk



G: Assign to First School
High-Risk



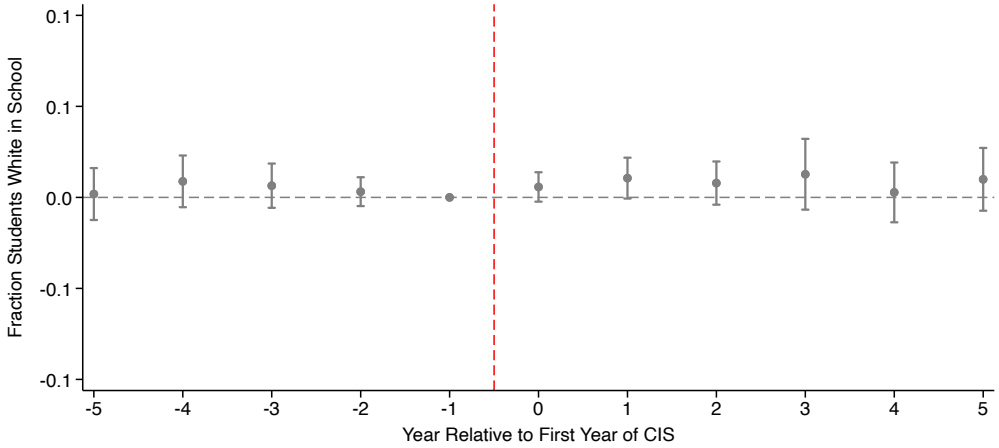
H: Assign to First School
Low-Risk



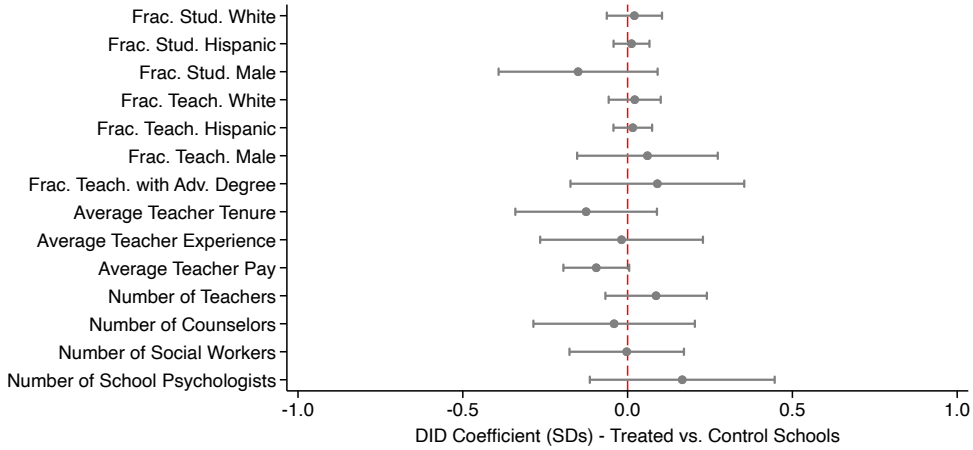
Notes: This figure presents a series of robustness checks for the baseline results shown in Figure V. Panels A and B modify the specification in Equation 1 to interact the γ year fixed effects with race/ethnicity and sex indicators, separately for high- and low-risk students. Panels C (high-risk) and D (low-risk) are estimated using a staggered adoption design in which later-treated CIS schools serve as the control group for earlier-treated schools. For each implementation cohort $T(s)$ and relative time τ, the control group is defined as the set of schools s' such that $T(s') > T(s) + \tau$. See Section V.A for further details. Panels E (high-risk) and F (low-risk) extend this staggered design by interacting the γ year fixed effects with commuting zone fixed effects, ensuring comparisons occur within commuting zones. Panels G (high-risk) and H (low-risk) return to the matched-control design but assign students to a predicted middle school based on their first elementary school enrollment. To do so, we construct a school-to-school transition matrix by grade and birth cohort and assign students to the modal transition path. We then re-estimate Equation 1 among students predicted to appear in the treatment and control schools. See Section V.A for additional details and Appendix Figure A.3 for the IV estimates accounting for differences in CIS exposure due to moves during childhood.

FIGURE VII: Change in Covariates After CIS Program Begins

A: Fraction of Students White



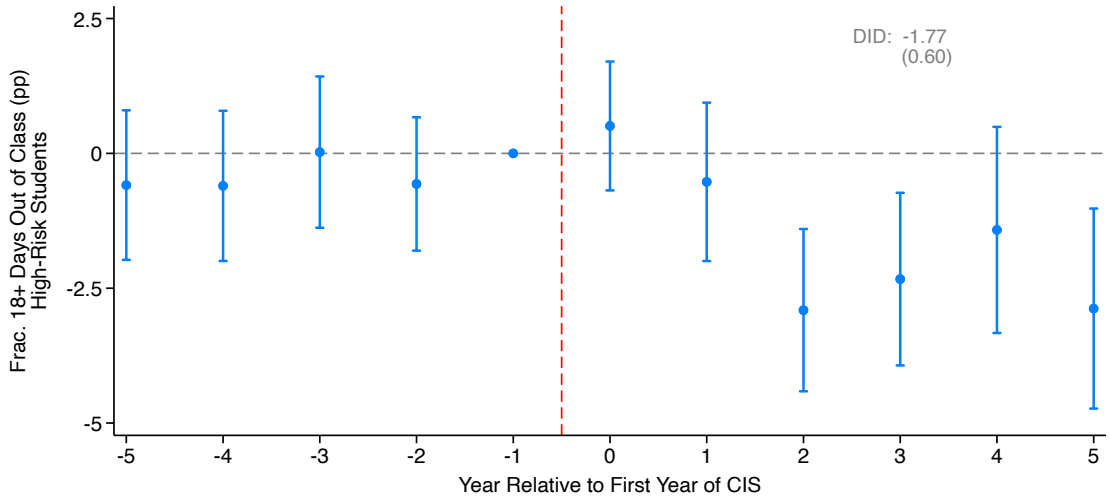
B: DID Coefficients for Student and Teacher Characteristics



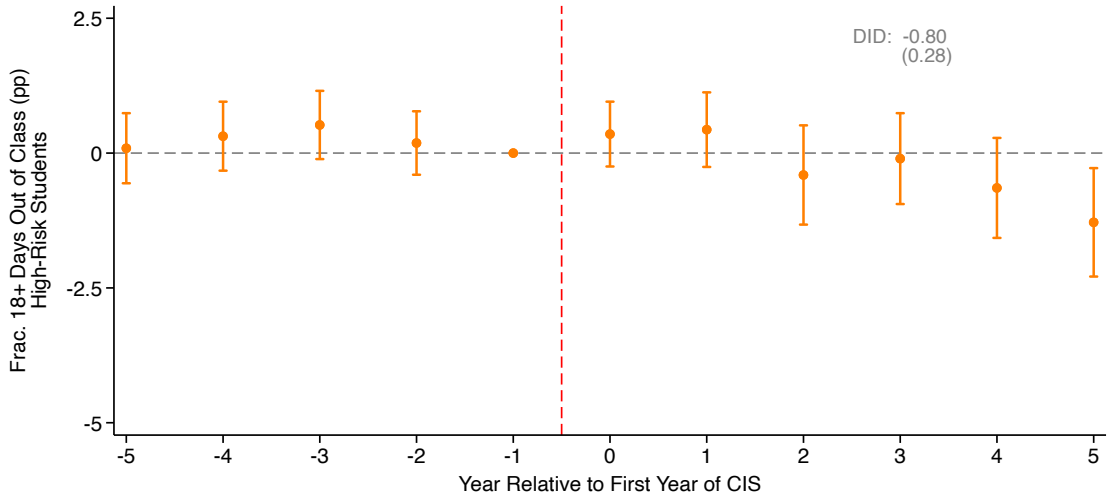
Notes: This figure assesses whether student and teacher characteristics change in treatment schools relative to matched control schools following the start of a CIS program. Panel A estimates an event study version of Equation 1, omitting race/ethnicity and sex interactions since the dependent variable—the fraction of students who are White—is not group-specific. Panel B presents results from a difference-in-differences specification (Equation 2), again excluding demographic interactions, applied to a range of student,teacher, and staff covariates. Coefficients are scaled in standard deviation units for comparability, and we report estimates of $\beta_{post,3-5}$, representing average post-treatment differences in years 3 to 5. Standard errors are clustered at the school level.

FIGURE VIII: Effect of CIS on Chronic Disengagement

A: High-Risk Students



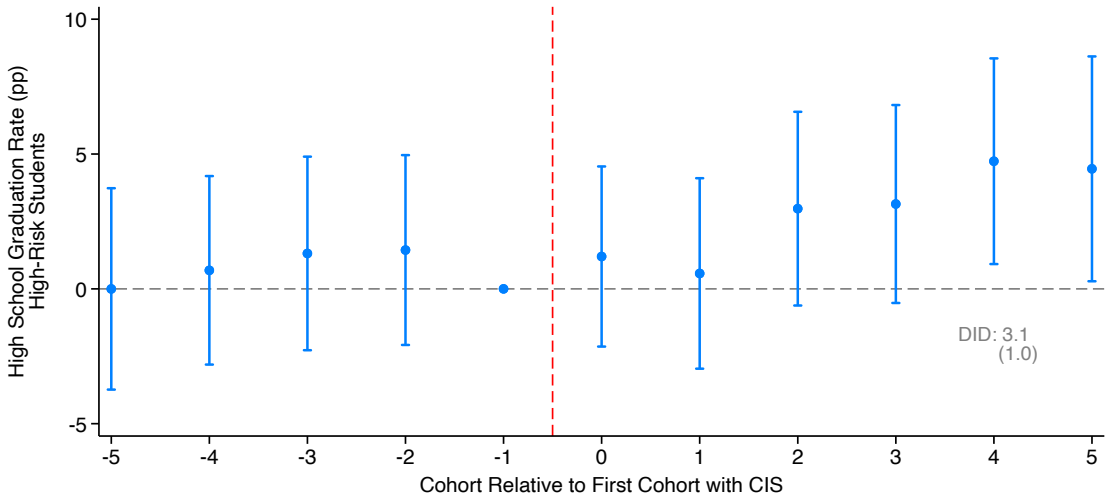
B: Low-Risk Students



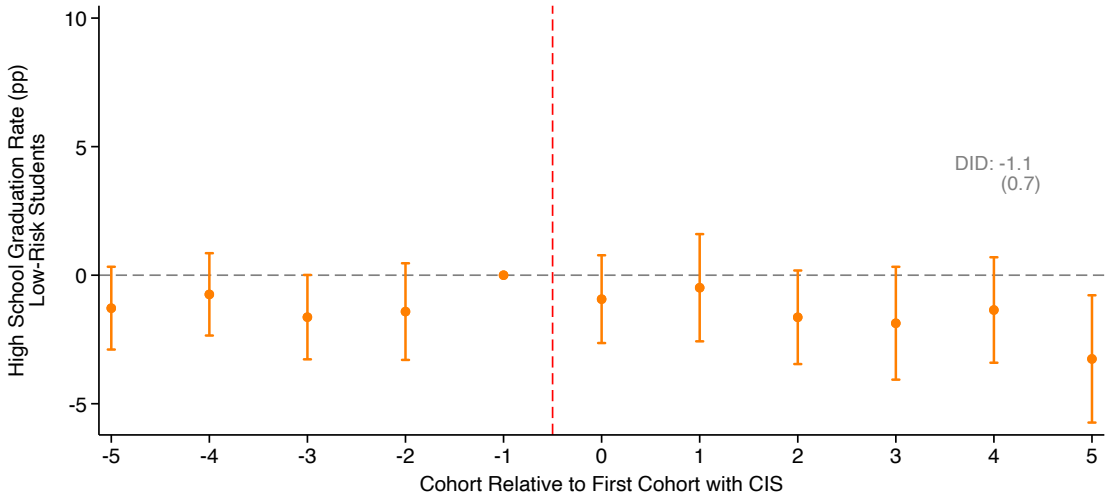
Notes: This figure presents estimates of the effect of CIS adoption on a measure of chronic disengagement, separately for high-risk (Panel A) and low-risk (Panel B) students. We define an indicator equal to one if a student misses 18 or more days of the school year due to either absences or time out of the classroom for discipline and estimate Equation 1 separately for each group using data on middle school students between 1998 and 2016. Difference-in-differences coefficients are presented on the figure and can be found in Appendix Table A.6, along with effects for other behavioral outcomes. Standard errors are clustered at the school-by-birth-cohort level.

FIGURE IX: Effect of CIS on High School Graduation

A: High-Risk



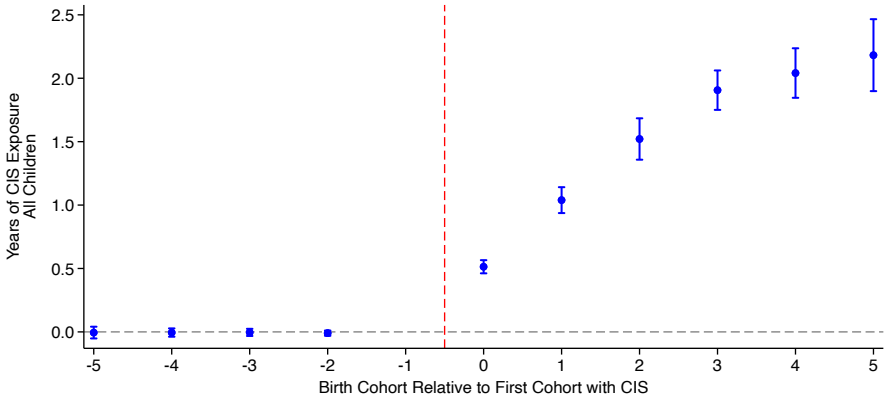
B: Low-Risk



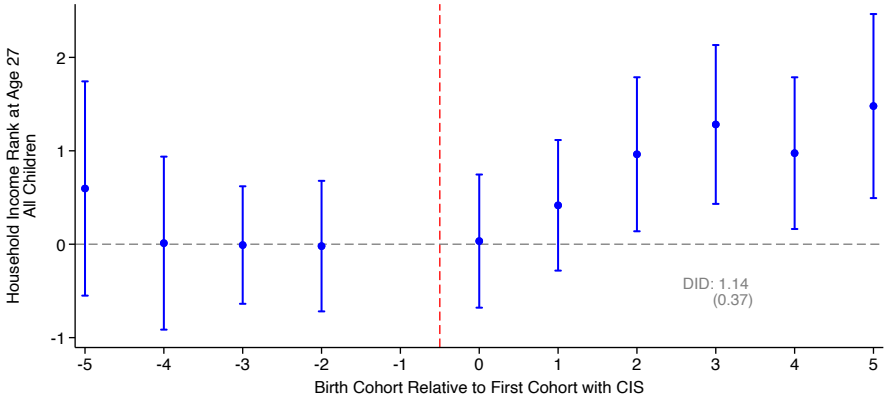
Notes: This figure presents event study coefficients estimating the effect of CIS on high school graduation rates. Panels A and B show effects on high school graduation rates for high- and low-risk students, respectively, using data on students in middle schools between 1998 to 2016, and estimated using Equation 3. High school graduation is defined as having graduated by 2022. Standard errors are clustered at the school-by-birth-cohort level.

FIGURE X: Effect of CIS on Program Exposure and Household Income Rank at Age 27
Census-IRS Sample: All Children

A: Years of CIS Exposure

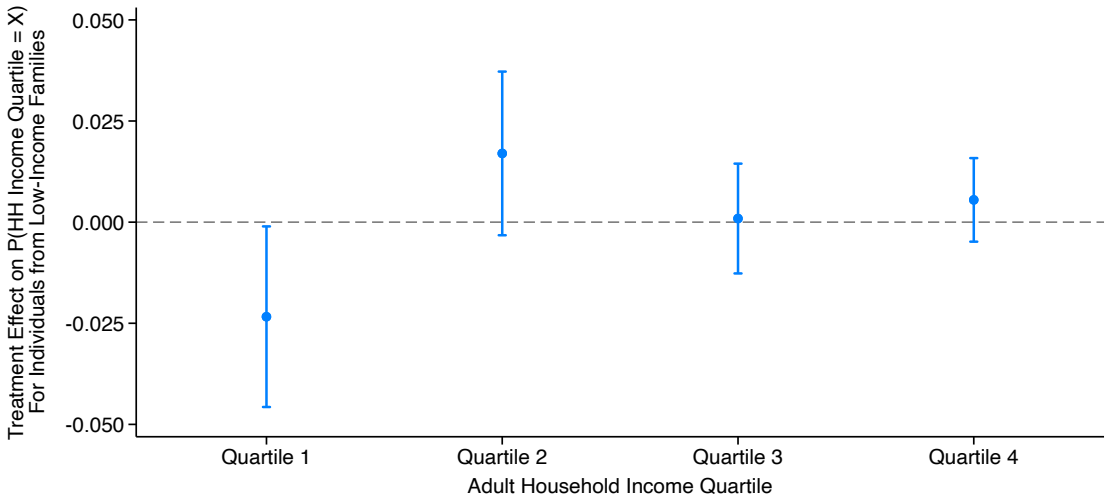


B: Household Income Rank at Age 27



Notes:This figure presents the effects of CIS on program exposure and adult household income rank using the Census-IRS data described in Section III C. In both panels, we plot event study coefficients estimated from Equation 4. In Panel A, the dependent variable is the average number of years of exposure to a CIS program. In Panel B, the dependent variable is adult household income rank measured at age 27. Standard errors are clustered at the high school catchment zone boundary level.

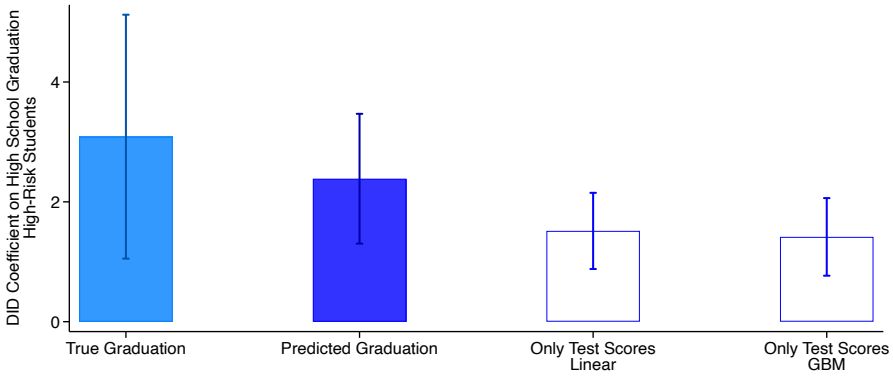
FIGURE XI: Effect of CIS Exposure on Adult Income Quartiles
Census-IRS Sample: Low-Income Children



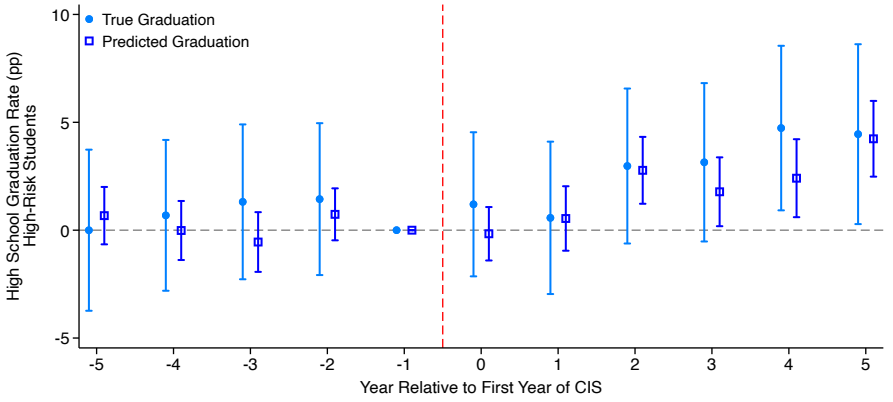
Notes: This figure shows the effect of CIS on the probability that low-income students fall into different quartiles of the income distribution at age 27. It is estimated using Census-IRS data, described in Section III.C. We show difference in difference coefficients from four separate regressions, where in each the dependent variable is an indicator equal to 1 if a student is in quartile X at age 27. Coefficients capture average effects for students born 3–5 years after the implementation cohort, estimated using the difference-in-differences analog of Equation 4. Standard errors are clustered at the high school catchment zone level.

FIGURE XII: Comparing True vs. Predicted Long-Run Outcomes

A: DID High School Graduation
High-Risk Students



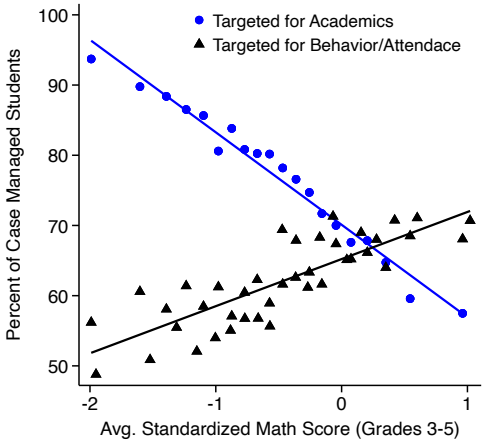
B: High School Graduation
High-Risk Students



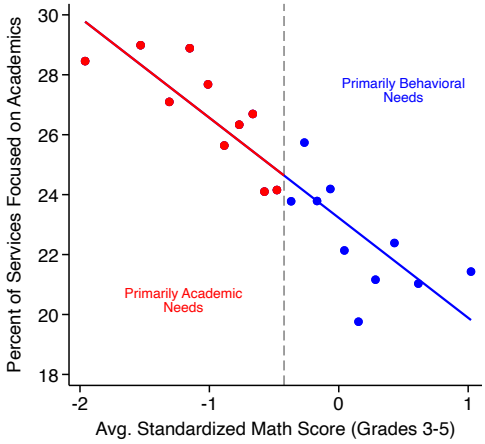
Notes: This figure compares the true effect of CIS on high school graduation for high-risk students to predicted effects based on contemporaneous outcomes. Panel A reports difference-in-differences estimates from each specification. Hollow bars show two alternative predictions based only on test scores. In the first, we use a linear regression of graduation on middle school test scores, estimated using statewide data from 1998–2004. In the second, we fit a flexible gradient boosting model on the same data to capture possible nonlinear relationships. In Panel B, we present event study coefficients estimated using the matched control design. The dark purple series (circles) replicates the result shown in Figure IX Panel *A*. The lighter series (squares) shows coefficients from Equation 1, where the dependent variable is predicted high school graduation based on a gradient boosting model trained on contemporaneous academic and behavioral outcomes. See Appendix Figure A.12 for details on the prediction model. Standard errors are clustered at the school-by-birth-cohort level.

FIGURE XIII: The Role of Personalization

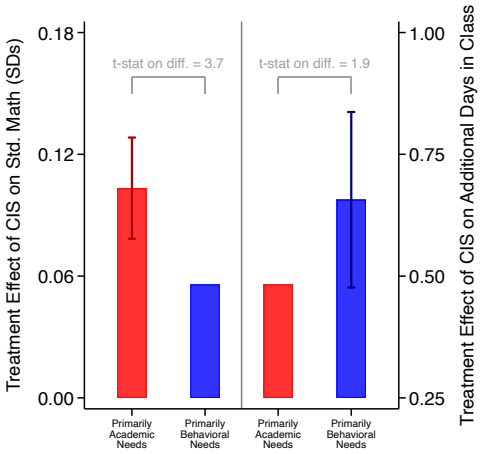
A: Targeted Areas of Support Among Case-Managed Students



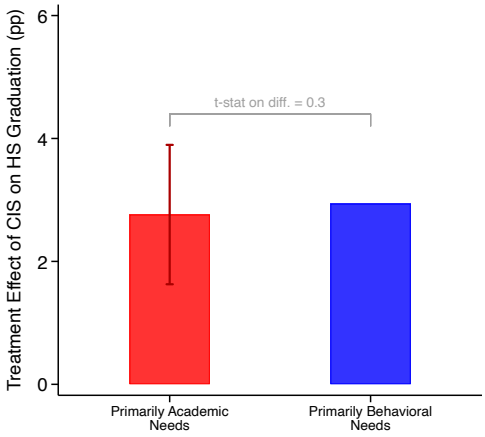
B: Percent of Services Academic Among Case-Managed Students



C: Effect of CIS on Short-Run Outcomes By ES Math Performance



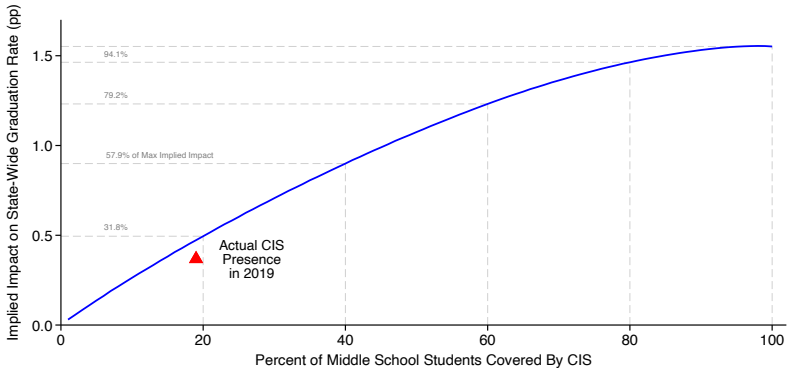
D: Effect of CIS on HS Graduation By ES Math Performance



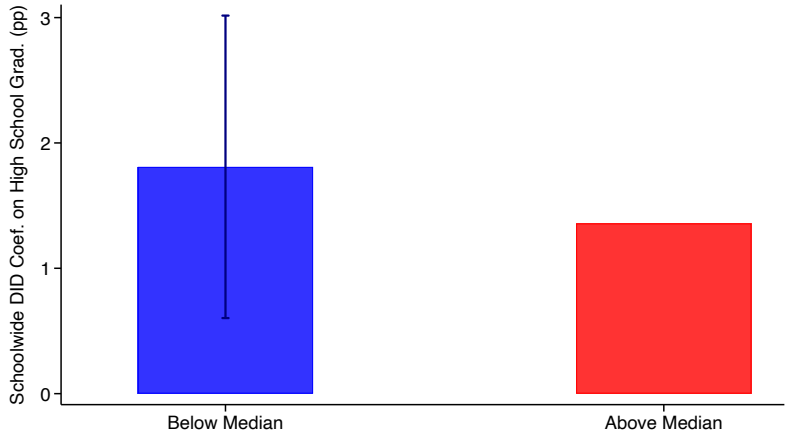
Notes: This figure presents results on the personalization of CIS services. Panel A shows the relationship between the fraction of students in case-management flagged for academics and behavior and their average standardized math scores in grades 3-5. Panel B plots the relationship between the fraction of academic services a student receives and their average standardized math scores in grades 3–5. In both panels, the sample includes case-managed sixth graders between 2016 and 2019 from the CIS microdata, excluding the 20 percent of students with the fewest navigator interactions. In Panel A we remove school-year observations where 100% or 0% of students are flagged for a particular category. The figures present a binned scatterplot controlling for students' predicted probability of case management and year fixed effects. Students can be flagged for any combination Academics, Behavior, Attendance, and/or Social Services (see Section III.A for details and Appendix Table A.1 for the full breakdown). We define a student as flagged for academics (behavior/attendance) if academics (behavior/attendance) is included in their list of targeted areas of support. Academic services are defined in Section VI. Panels C (short-run outcomes) and D (high school graduation) show difference-in-differences (DID) estimates from a modified version of Equation 2. The specification is estimated at the individual level among high-risk students and includes controls for predicted case-management risk, an indicator for below-median elementary-school test scores, and interactions between this indicator and the two post-period × treatment indicators. The blue bars (right) report the estimated treatment effects for students with above-median prior test scores, while the red bars (left) report the corresponding effects for students with below-median scores. Vertical lines show 95 percent confidence intervals for the difference between the two group-specific treatment effects. Standard errors are clustered at the school-by-birth-cohort level.

FIGURE XIV: CIS Expansion Counterfactual

A: Frontier Allocation by Predicted Impact on High School Graduation



B: Difference in School-Level Treatment Effects for Above vs. Below Median Years Since First School in Affiliate



Notes: This figure presents the results of a simulation estimating the effects of a hypothetical expansion of CIS programs. Motivated by the results in Figure IX and Figure A.6, where gains are concentrated among higher-risk students, Panel A simulates a counterfactual rollout in which CIS is expanded to schools in descending order of predicted effect, based on student risk concentration. The horizontal axis plots the share of all middle school students covered by CIS; the vertical axis shows the corresponding implied increase in the statewide high school graduation rate (equal to the treatment effect of CIS for each risk group, multiplied by the share of students at the school in each group. The triangle marks the actual 2019 CIS allocation, which covers approximately 20% of middle school students. The vertical gap between the actual allocation and the blue frontier curve reflects that CIS programs are not strictly allocated to the highest-need schools (as measured by the risk composition in the school). In dashed lines we show the percent of the maximum implied impact every 20 percentiles. Panel B asks whether the treatment effects of CIS attenuate as the program expands within an area. For each CIS program, we calculate the difference in the start date of that program relative to the first program in that affiliate (e.g. relative to the first CIS program in Dallas). We split schools at the median of this measure and estimate school-wide treatment effects on high school graduation. We estimate the difference-in-differences analog of Equation 3, allowing λ and γ to vary flexibly in each risk group. We then interact the DID coefficients with an indicator for being above the median in years since first affiliate program. We also include fixed effects of the interaction between the above median indicator and $T(s)$, the stack identifier, to allow for level differences for above/below median schools in each implementation cohort. We present the DID coefficients for each group and report the t-statistic on the difference between above and below median schools. Standard errors are clustered at the school by birth cohort level.

Supplementary Appendix

A Additional Details on Data Construction

Our analysis primarily uses administrative data from the Texas Education Agency (TEA), supplemented with postsecondary enrollment data from the Texas Higher Education Coordinating Board (THECB) and quarterly earnings records from the Texas Workforce Commission (TWC).

We begin by assembling a student-level enrollment panel starting in 1994. Each year's enrollment file reports the student's campus of attendance; basic demographics (sex, race/ethnicity, and age as of September 1); and indicators for free- or reduced-price lunch eligibility and "at risk of dropping out" status. After constructing the panel, we clean invariant demographics (sex, birth cohort, and race/ethnicity) to enforce within-student consistency. When a student's reported value varies across years—which is rare, with fewer than 1% of students having inconsistent sex or cohort information and about 4% having inconsistent race/ethnicity—we assign the modal value.[1] Because enrollment data are collected at a single point in the fall semester, it is also uncommon for a student to appear in more than one campus within a year: fewer than 1% of students are listed at multiple schools in the same academic year. In such cases, we randomly assign the student to a single campus for that year. The resulting enrollment panel serves as the spine for all data linkages used in the analysis.

Our primary contemporaneous outcome is students' performance on standardized tests in grades 3–8. Over our study period, Texas administered three different assessments: the Texas Assessment of Academic Skills (TAAS, 1990–2002), the Texas Assessment of Knowledge and Skills (TAKS, 2003–2012), and the State of Texas Assessments of Academic Readiness (STAAR, 2012–present). To ensure comparability across regimes, we standardize all scores within exam–grade–year cells and include year fixed effects in all specifications.

The administrative testing files sometimes contain multiple records for the same student in the same exam–grade–year because Texas reports scores from the initial administration as well as make-up tests and retakes. When multiple scores are present, we assign the student the maximum score. This choice reflects the institutional purpose of retakes—students may demonstrate mastery following an absence or a low initial score—and reduces noise arising from within-student variation. For instance, in the 2009 8th-grade TAKS math exam, 20% of students have multiple records, some of which contain scores of 0 indicating a retake. When students have multiple non-zero or non-missing scores, they tend to be close in value. The within-student variance among these duplicates (4.1 points) is less than half the overall variance in that exam–grade–year (9.1

[1] Under the Education Research Center's FERPA review process, we are prohibited from reporting percentages below 1%, regardless of sample size.

points). Using the maximum score therefore provides a consistent measure of each student's demonstrated achievement without materially altering the score distribution.

In addition to test scores, we use detailed attendance records reported every six weeks. For each student–campus–grade–period cell, the data include the number of instructional days the student was enrolled and the number of days absent. It is extremely uncommon for a student to have multiple records within the same student–campus–grade–period; when this occurs, we retain the record with the highest number of instructional days and break any remaining ties randomly.

A small share of cases, however, reflect students appearing at multiple campuses within the same six-week period. In 2015, 1.2% of student–period cells contain more than one observation, and 5.8% of students experience at least one such period. In these cases, we assign the student the average number of days enrolled and absent across the duplicates. These multi-campus records often arise when students change schools mid-period, and it is not possible to determine which campus's record most accurately reflects the true number of days enrolled or absent. Averaging across duplicates prevents movers from being assigned spuriously extreme values—such as a full set of absences at one campus when the student had already transferred—which could otherwise mechanically inflate absence measures.

Our final set of contemporaneous outcomes comes from student-level disciplinary records. These data are reported at the student-by-incident level and include the type of action (in-school suspension, out-of-school suspension, or expulsion) and, for suspensions, the number of days assigned. We aggregate these records to the student-by-year level, constructing measures of the total number of disciplinary actions (overall and by type), the total days assigned, and indicators for ever receiving each type of action in a given year.

Finally, we use TEA administrative data to measure high school graduation as a long-run outcome. The graduation files list all students who graduate in a given year and the campus from which they graduate. Our primary measure is an indicator equal to 1 if a student ever appears in the graduation records. We observe graduation outcomes through 2022 and code students who never appear in these files as non-graduates. Our findings are robust to alternative definitions, including graduating by age 18, by age 20, or at any point between ages 16 and 20. It is extremely rare for a student to have multiple graduation records (<1%), but when this occurs, we assign the first observed graduation year when constructing age-based measures.

We link TEA records to college enrollment data from THECB using ERC-provided student identifiers. The THECB data cover enrollment at Texas higher education institutions from 1992 to 2019. Enrollment is observed separately in the fall, spring, and summer terms, and institutions are categorized into four groups: 2-year colleges, 4-year colleges, health institutions, and independent institutions. Our analysis focuses on 2-year and 4-year colleges, which represent the public higher education sector in Texas. We define college attendance as ever enrolling in a given sector (2-year

or 4-year) between ages 18 and 22. Out-of-state enrollment is not observed for most of our sample; National Student Clearinghouse data are only available in the most recent years. As a result, students who attend college exclusively out of state cannot be distinguished from non-attenders and are coded as not attending. However, Texas has an unusually high rate of in-state enrollment: in 2020, 84% of first-time college enrollees from Texas remained in-state, roughly 10 pp above the national average (U.S. Department of Education, 2022).

Finally, we use earnings data from the Texas Workforce Commission (TWC) to validate the earnings results from the national sample shown in Section V.C. The TWC data consist of quarterly unemployment-insurance (UI) wage records for all workers in Texas covered by UI. These records include worker–employer–quarter observations of earnings; coverage excludes the self-employed, independent contractors, military personnel, and informal-sector workers (Stevens, 2007). We aggregate earnings to the worker-year level by summing across all employers and inflate all dollar values to 2025 dollars using the CPI-U. Because the UI data cannot distinguish between unemployment, employment in non-covered jobs, and out-migration from Texas, we follow standard practice and classify individuals as "active workers" if they ever have positive UI-covered earnings between ages 18 and 25. We impute zero earnings only in years with missing earnings data for this active-worker sample. Using these annual earnings, we construct within-cohort earnings ranks at each age following Chetty et al. (2026).

After merging all data sources, our final dataset contains a set of invariant student characteristics (sex, birth cohort, race/ethnicity); year-varying characteristics (campus attended, free- or reduced-price lunch status, and dropout-risk indicators); year-specific contemporaneous outcomes (test scores, attendance, and disciplinary actions); and student-invariant long-run outcomes (high school graduation, college attendance, and earnings).

We use data provided directly by CIS to identify which schools operated CIS programs and the year each program began. The original file lists 1,216 CIS schools as of 2019–20. We first exclude 38 records with unusable NCES school IDs (missing, duplicated, or not corresponding to Texas schools). We then drop 38 schools classified as "alternative schools" (e.g., alternative education centers or career centers). An additional 101 schools lack information on the program start date, and 27 schools have NCES IDs that do not match to a TEA campus code, preventing linkage to the microdata. After applying these restrictions, we are left with 1,012 usable CIS program records.

To construct the analysis sample, as described in Section IV, we focus on middle schools and match CIS-treated schools to never-treated controls. For each CIS implementation cohort, we take the set of middle schools receiving a CIS program in a given year (e.g., schools first treated in 2004) and compile school- and student-level data from six years prior to program adoption (1998 in this example). We then append data from all never-treated middle schools in that same baseline year. This forms the matching dataset for the 2004 implementation cohort. We repeat this procedure for

all implementation cohorts from 2001 through 2019.

Using the stacked data, we estimate a regression predicting whether a school adopts a CIS program in year t based on its characteristics in year t-6. The predictors include the fraction of students classified as "at risk of dropping out," the fraction White, the schoolwide absence rate, the fraction of high-risk and low-risk students (see Appendix B for details), and the predicted high school graduation rate based on contemporaneous characteristics (Section VI). We also include implementation-cohort fixed effects. For each school, this model yields a predicted probability of CIS adoption.

For each treated school, we compare its predicted probability to those of all potential control schools in the same implementation cohort, excluding never-treated schools in the same district. We then select the control school with the closest predicted probability of CIS.

B Additional Details on Case Management Prediction

As discussed in Section IV, our analysis focuses on students who would be most and least likely to interact with a CIS site coordinator. This requires constructing predicted probabilities of case management for students in both treated and control schools.

We estimate these predictions using data on actual case-management status between 2014 and 2019. We assemble a dataset containing 103,000 students who were ever case managed in grades 6–8 during this period and 612,000 other students in the same schools and years who were never case managed in middle school.[2] For each student, we construct a set of predictors based on their elementary school characteristics in the five years prior to their middle-school observation.[3]

We construct a series of in-school outcomes from test score, attendance, and disciplinary records. From the testing data, we include standardized math scores and an indicator for missing test scores in a given year.[4] From the attendance records, we include days enrolled and days absent in each six-week period, the annual absence rate, and an indicator for being chronically absent (absent more than 10% of enrolled days). From the disciplinary data, we include the number of in-school and out-of-school suspensions, an indicator for ever receiving either type of suspension, and the total number of days assigned.

Finally, in addition to school-based outcomes, we include a set of demographic predictors. For each student, we use invariant race/ethnicity and sex, as well as year-specific indicators for free- or reduced-price lunch eligibility and TEA "at-risk" status. We do not include birth cohort

[2] We exclude students who attended a CIS elementary school so that their elementary characteristics are not a function of prior case management.

[3] This avoids issues arising from possible grade retention. For example, if a student repeated third grade, we include both years of third-grade data and do not include the year the student was in first grade.

[4] Across these years, not all grades administered standardized tests in every year. For example, in 2015 no STAAR exam was given in first grade.

or school identifiers, as our goal is to generate predictions that can be applied out of sample to different schools and years.

Using these data, we estimate the probability of case management with a gradient-boosted decision-tree algorithm (Ke et al., 2017). We reserve 30% of the sample as a validation set, which is not used in training. The validation data are used solely to assess model performance (e.g., Figure III*C*).

Using the remaining 70% of the data, we train the decision-tree model. Because CIS case management is a relatively rare outcome—14% of students are case managed in the prediction sample—we combine a standard GBM with a second model that up-weights case-managed observations.[5] We implement this using five-fold cross-validation. For each fold, we proceed as follows:

1. Hold out the fold (20% of the training data);
2. Estimate the basic GBM on the remaining data;
3. Generate predictions for the held-out fold;
4. Estimate the up-weighted GBM on the remaining data;
5. Generate predictions for the held-out fold.

This procedure yields out-of-sample predictions from both models for every observation in the training data. We then regress case-management status on these two sets of predictions using a logit model and save the resulting coefficients.

Finally, we re-estimate each GBM model on the full training sample. This leaves us with two fitted GBM models—one standard and one up-weighted—and the coefficients from the logit model used to blend their predictions. Importantly, because the logit was trained only on out-of-sample predictions generated during cross-validation, its coefficients are not fitted on the same predictions produced by the final GBM models. This prevents overfitting and ensures proper separation between training and blending.

To apply these predictions to the full sample (beyond the training and validation data), we construct an analogous version of the prediction microdata for all years in our dataset. We then apply both GBM models to these back-filled records to generate predicted probabilities from each model. Using the previously estimated logit coefficients, we blend the two sets of predictions to obtain the final predicted likelihood of case management for every student. We classify students in the top quartile of this predicted distribution as high-risk and those in the bottom quartile as low-risk (see Figure III*C*).

[5]This approach follows standard methods for learning from imbalanced data (e.g., He and Garcia, 2009). We up-weight positive cases by a factor of 10.

TABLE A.1: Characteristics of Case-Managed Students

A. Student Characteristics (Share %)	
Male	46.6
White	12.8
Hispanic	66.7
Black	17.9
Other race/ethnicity	2.6
Eligible for FRPL	71.0
At risk of dropping out	76.2
B. Student Outcomes	
Standardized Math Score in 6th Grade (SDs)	-0.54
Absence Rate in 6th Grade	0.05
Share ever received an in-school suspension in 6th grade	26.5
Share ever received an out-of-school suspension in 6th grade	14.6
C. Reason for Referral (Share %)	
Referral includes (not mutually exclusive)	
Academics	83.4
Attendance	22.3
Behavior	65.5
Social services	34.7

	Mean	p25	p50	p75
	(1)	(2)	(3)	(4)
D. Information on Interactions				
Days after August 1 parent consent was received	77	35	64	126
Number of interactions in year	38	18	29	48
Number of unique services received	14	9	13	18
Share of services that were individual (%)	31.9	1.9	9.1	64.0
Share of services from an external organization (%)	14.8	0.0	0.0	0.0

Notes: This table presents summary statistics for students who were case managed by CIS in a middle school between 2014 and 2019, using the CIS microdata described in Section III B. Panels A–C report shares in percentage points. Panel D reports distributional statistics for students' interactions with CIS navigators. Referral categories are not mutually exclusive. A unique service refers to a distinct service code in the CIS microdata. Services are categorized as individual or group-based and are classified by whether they were delivered directly by TEA/CIS staff or by an external provider. The sample contains 170,000 student-years and 132,000 students.

TABLE A.2: Needs Addressed by CIS

	Number of Student-Years (1,000s) (1)	Fraction of Case-Managed Students-Years (2)
Targeted Need Includes		
Grades	70.3	0.409
Academic Readiness	69.6	0.405
Social Skills	61.9	0.360
Self Esteem	55.5	0.323
Basic Needs	53.1	0.309
College Readiness	36.9	0.215
Homework Completion	35.4	0.206
Absences	30.5	0.177
Classroom Conduct	26.3	0.153
Language Development	24.5	0.142
Tardies	18.7	0.109
Classroom Participation	17.9	0.104
Family Conflict	13.2	0.077
Mental Health and Wellness	12.3	0.072
Life Skills	10.4	0.061
Career/Employment	9.2	0.054
Delinquent Conduct	5.8	0.034
Health	3.3	0.019
Grief/Loss	1.9	0.011
Housing	1.4	<0.01
Suspected Substance Use	1.0	<0.01
Violence	0.9	<0.01
Gang Involvement	0.6	<0.01

Notes: This table provides information on the detailed needs of students who are case managed by CIS in middle schools from 2014-2019, using the CIS microdata described in Section III B. For each need category, we provide the fraction of student-year observations who are flagged as having this need (not mutually exclusive).

TABLE A.3: Effect of CIS on Test Scores under Alternative Risk Definitions

	High-Risk		Low-Risk	
	Risk Model	Math + Abs.	Risk Model	Math + Abs.
	(1)	(2)	(3)	(4)
$\beta_{post,3-5}$	0.106	0.074	0.011	0.007
	(0.023)	(0.020)	(0.024)	(0.020)
Pre-Period Mean	-0.7	-0.8	0.7	0.7
Frac. Case Managed (2014–2016)	25.2	20.0	2.4	6.7
Observations (1,000s)	417	406	286	290

Notes: This table reports the difference-in-differences effect of CIS on standardized math scores, estimated using Equation 2, under alternative definitions of high- and low-risk students. Columns (1) and (3) use our baseline gradient-boosted risk model. Columns (2) and (5) instead classify students by the bottom (top) of the elementary-school distribution of a combined index of elementary-school math scores and attendance, where the cutoffs match the size of each group using the risk model. The $\beta_{post,3-5}$ row reports the DID coefficient with its standard error in parentheses; *Pre-Period Mean* reports the mean standardized math score in the pre-period; *Frac. Case Managed (2014–2016)* reports the share of each group that is actually case managed between 2014 and 2016; and the final row reports the number of observations. Standard errors are clustered at the school-by-birth-cohort level.

TABLE A.4: Effect of CIS for Low- vs. High-Risk Students
School Clustering

	High-Risk	Low-Risk	Difference
	(1)	(2)	(3)
Years of CIS Exposure			
$\beta_{post,3-5}$	1.699	1.745	-0.046
	(0.030)	(0.047)	(0.040)
Pre-Period Mean	0.0	0.0	
Observations (1,000s)	455	329	784
Standardized Math Scores			
$\beta_{post,3-5}$	0.106	0.011	0.094
	(0.043)	(0.043)	(0.049)
Pre-Period Mean	-0.7	0.7	
Observations (1,000s)	417	286	703
True HS Graduation			
$\beta_{post,3-5}$	3.088	-1.082	4.170
	(1.247)	(0.829)	(1.331)
Pre-Period Mean	64.5	89.3	
Observations (1,000s)	429	316	745
2-Year College Attendance			
$\beta_{post,3-5}$	2.654	-3.162	5.815
	(2.479)	(2.392)	(2.810)
Pre-Period Mean	29.2	52.7	
Observations (1,000s)	161	111	272
Predicted HS Graduation			
$\beta_{post,3-5}$	2.363	0.301	2.062
	(0.813)	(0.675)	(0.845)
Pre-Period Mean	53.9	83.2	
Observations (1,000s)	455	329	784

Notes: This table presents an analogous table to that in Table III, but with standard errors clustered at the school level instead of the school-by-birth-cohort level. See the notes to Table III for details.

TABLE A.5: Effect of CIS for High-Risk Students by Demographic Group

	Sex			Ethnicity		
	Male	Female	Difference	Hispanic	Non-Hispanic	Difference
	(1)	(2)	(3)	(4)	(5)	(6)
First Stage						
$\beta_{post,3-5}$	1.702	1.697	0.005	1.734	1.624	0.110
	(0.026)	(0.026)	(0.012)	(0.030)	(0.027)	(0.030)
Pre-Period Mean	0.0	0.0	0.0	0.0	0.0	0.0
Observations (1,000s)	176	278	454	289	166	454
Std. Math Scores						
$\beta_{post,3-5}$	0.127	0.091	0.036	0.106	0.117	-0.011
	(0.027)	(0.025)	(0.025)	(0.024)	(0.031)	(0.033)
Pre-Period Mean	-0.8	-0.7	-0.7	-0.7	-0.7	-0.7
Observations (1,000s)	160	258	417	268	149	417
True HS Grad.						
$\beta_{post,3-5}$	3.225	3.024	0.201	3.264	2.358	0.906
	(1.405)	(1.201)	(1.554)	(1.276)	(1.466)	(1.750)
Pre-Period Mean	58.4	68.2	64.5	65.5	62.9	64.5
Observations (1,000s)	166	263	429	272	157	429
2-Yr Coll.						
$\beta_{post,3-5}$	1.295	3.212	-1.917	4.675	1.091	3.580
	(2.481)	(2.147)	(2.922)	(2.229)	(2.262)	(2.779)
Pre-Period Mean	22.1	33.1	29.2	27.6	31.7	29.2
Observations (1,000s)	59	102	161	97	64	161

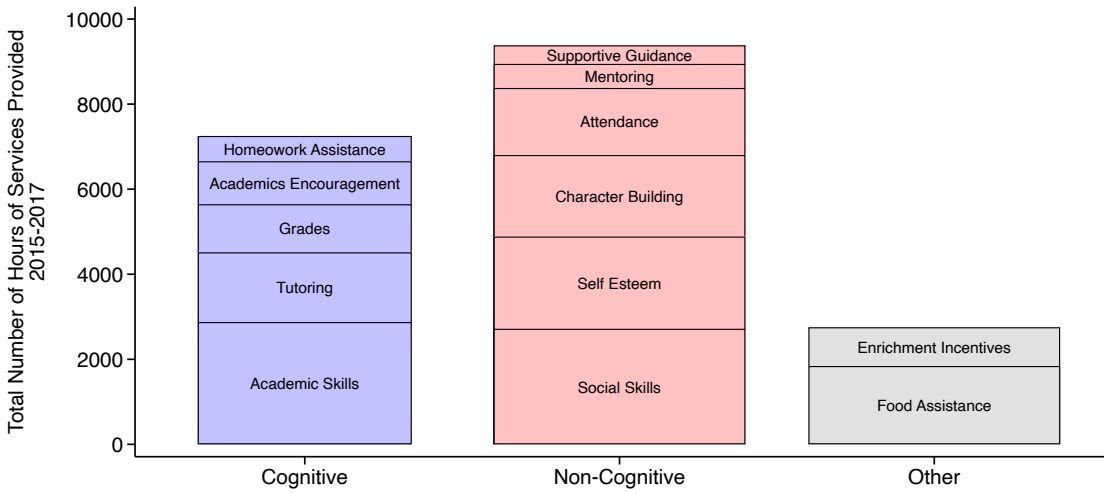
Notes: This table presents difference-in-differences coefficients estimated using Equation 2 and its long-run analog (see Section IV.C for details). In all columns we restrict to high-risk students. Columns (1) and (2) report coefficients for male and female students, respectively; columns (4) and (5) report coefficients for Hispanic and non-Hispanic students. In columns (3) and (6) we report the difference estimated from a single regression where we fully interact the DID regression with an indicator for being male and Hispanic, respectively. Each horizontal panel corresponds to a different outcome. Standard errors are clustered at the school-by-birth-cohort level.

TABLE A.6: Effect of CIS on Behavioral Outcomes

	High-Risk		Low-Risk	
	All	Male	All	Male
	(1)	(2)	(3)	(4)
Days Out of Class				
$\beta_{post,3-5}$	-0.491	-0.313	-0.387	-0.404
	(0.164)	(0.222)	(0.110)	(0.143)
Pre-Period Mean	9.4	10.3	5.1	5.2
Days Absent				
$\beta_{post,3-5}$	-0.446	-0.133	-0.315	-0.327
	(0.145)	(0.186)	(0.106)	(0.137)
Pre-Period Mean	7.9	8.0	4.8	4.9
Days Discipline				
$\beta_{post,3-5}$	-0.045	-0.180	-0.072	-0.078
	(0.059)	(0.101)	(0.022)	(0.033)
Pre-Period Mean	1.5	2.3	0.3	0.4
Days Out $\geq$ *18*				
$\beta_{post,3-5}$	-1.774	-1.285	-0.799	-1.199
	(0.602)	(0.941)	(0.278)	(0.412)
Pre-Period Mean	15.4	18.5	3.1	3.5
Days Out $\leq$ *1*				
$\beta_{post,3-5}$	0.949	0.524	2.120	2.468
	(0.551)	(0.736)	(0.910)	(1.075)
Pre-Period Mean	14.2	11.5	27.3	26.9
Has Disc. Record				
$\beta_{post,3-5}$	0.392	-1.834	-1.414	-1.555
	(0.931)	(1.429)	(0.545)	(0.763)
Pre-Period Mean	32.6	45.7	8.2	11.5
Fighting				
$\beta_{post,3-5}$	-0.010	-0.027	-0.005	-0.009
	(0.005)	(0.009)	(0.002)	(0.003)
Pre-Period Mean	0.0	0.1	0.0	0.0
Observations (1,000s)	422	156	326	179

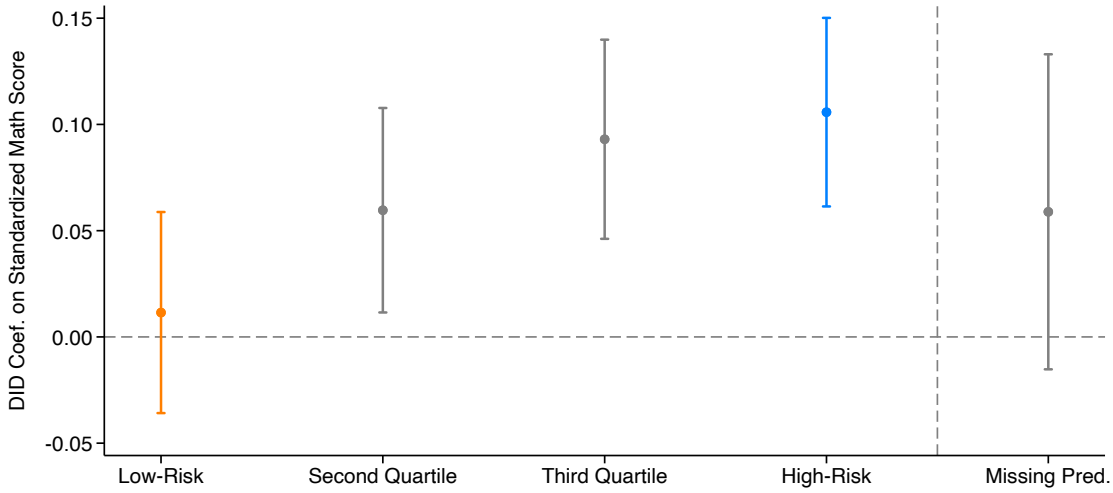
Notes: This table presents difference-in-differences coefficients on behavioral outcomes for high- and low-risk students. We estimate Equation 2 using data on middle school students between 1998 and 2016. The first two columns report results for high-risk students and the last two for low-risk students; within each pair, the first column includes all students and the second restricts to male students. Each row corresponds to a different outcome. *Days Out of Class* is the total number of days a student is out of class due either to absence or discipline (see Section V.B for details); *Days Absent* and *Days Discipline* report these two components separately. *Days Out* $\geq$ *18* and *Days Out* $\leq$ *1* are indicators for a high and low number of days out of class, respectively; event studies for *Days Out* $\geq$ *18* appear in Figure VIII. *Has Disc. Record* is an indicator for having any disciplinary record during the year, and *Fighting* is the number of times during the year a student is logged in the disciplinary records for fighting. The number of observations is constant across outcomes within a column and is reported once at the bottom. Standard errors, clustered at the school-by-birth-cohort level, are in parentheses.

FIGURE A.1: Top 10 Services Provided by CIS 2015-2017



Notes: This figure summarizes the most common services provided by CIS navigators, measured in total time. Using the CIS microdata described in Section III/B, we identify the top 10 most frequently provided services in each year from 2015 to 2017. We classify these services into three broad categories based on our own determination: cognitive, non-cognitive, and other. The height of each bar reflects the total time spent delivering that service across all three years.

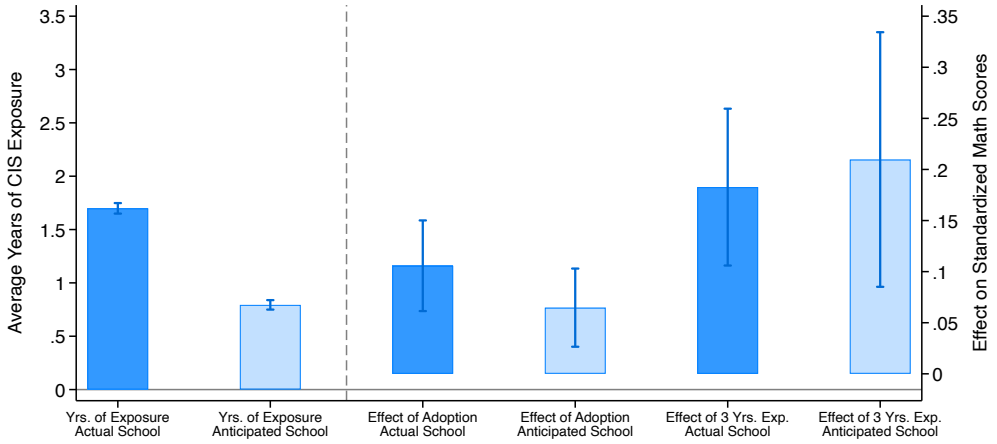
FIGURE A.2: Heterogeneity: Effect of CIS on Test Scores By Risk Group



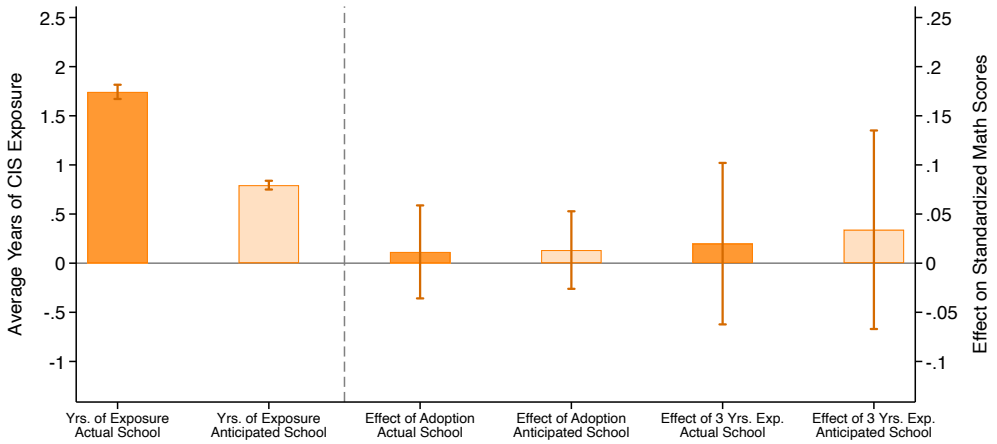
Notes: This figure reports difference-in-differences coefficients estimating the effect of CIS on standardized test scores, separately by student risk status. For each quartile of the risk distribution—as well as for students without a predicted risk score—we report the coefficient $\beta_{post,3-5}$ from Equation 2, comparing treated schools to matched control schools. The estimates for high- and low-risk students correspond to those shown in the event studies in Figure V. Students are missing a risk score if they lack elementary school data. Standard errors are clustered at the school-by-birth-cohort level.

FIGURE A.3: Effects of CIS on Standardized Test Scores: Actual School Assignment vs. First School Assignment

A: High-Risk Students



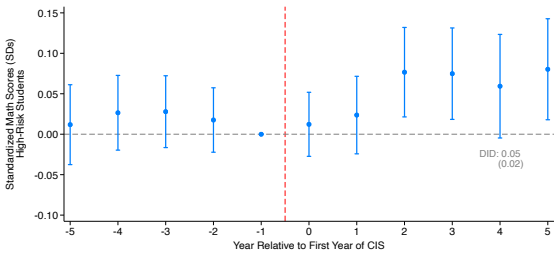
B: Low-Risk Students



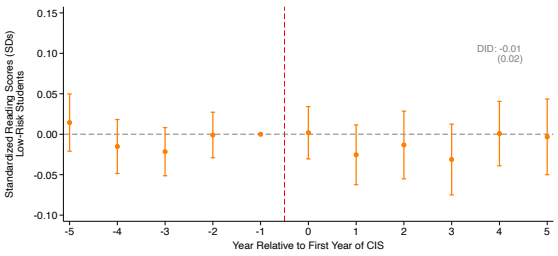
Notes: This figure compares the effects of CIS on exposure and on standardized test scores, along with the implied IV estimates, using students' actual versus anticipated middle school assignments. In Panel A, we restrict to high-risk students. The exposure and test-score estimates using actual school assignment replicate those reported in Figures IV and V, respectively. Anticipated school assignment follows the approach used in Figure VI: rather than using students' actual enrolled middle school, we assign them to an anticipated school based on the first elementary school they attend (see notes to Figure VI for details). The test-score estimate under anticipated assignment is the same as that shown in Figure VI Panel G. The IV estimates rescale the test-score effect by the corresponding effect on exposure to yield an effect per three years of CIS exposure. Panel B presents the analogous comparisons for low-risk students. Standard errors are clustered at the school-by-birth-cohort level.

FIGURE A.4: Additional Test Score Results

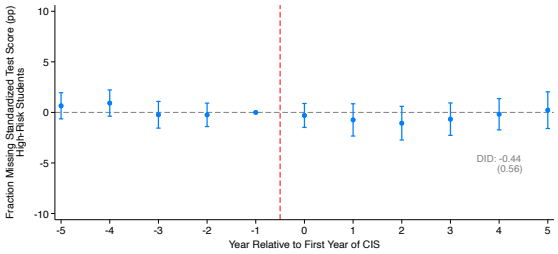
A: Effect of CIS on Std. Reading Scores
High-Risk Students



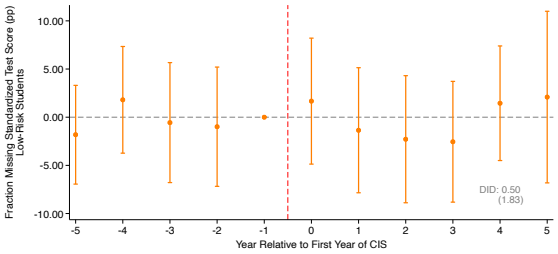
B: Effect of CIS on Std. Reading Scores
Low-Risk Students



C: Effect of CIS on Missing Std. Math
High-Risk Students



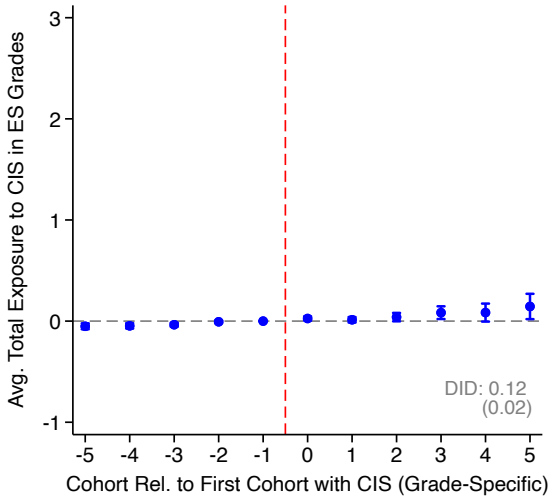
D: Effect of CIS on Missing Std. Math
Low-Risk Students



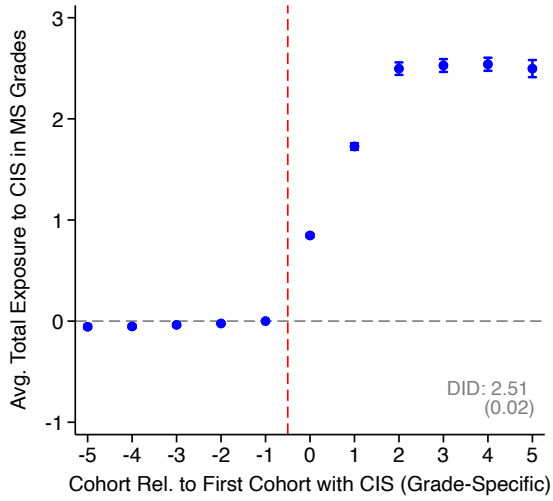
Notes: This figure presents additional results on the effect of CIS on student test score outcomes. Panels A and B show event studies analogous to those in Figure V, but for standardized reading scores instead of standardized math scores. In Panels C and D, the specification is the same, but the outcome is an indicator variable equal to 1 if a student is missing a standardized math score in year t. See notes for Figure V for additional details.

FIGURE A.5: Exposure to CIS by School Level

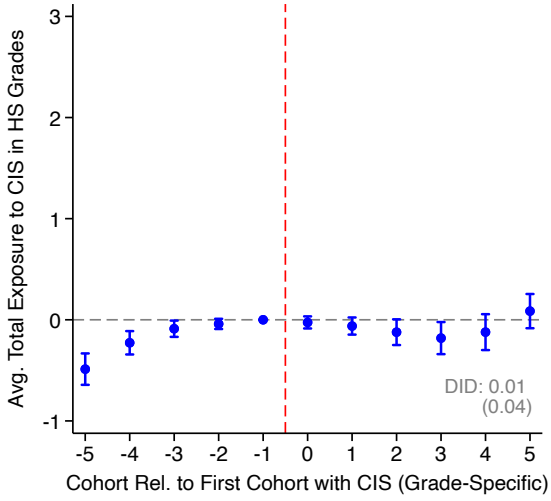
A: Total Exposure to CIS in Elementary School



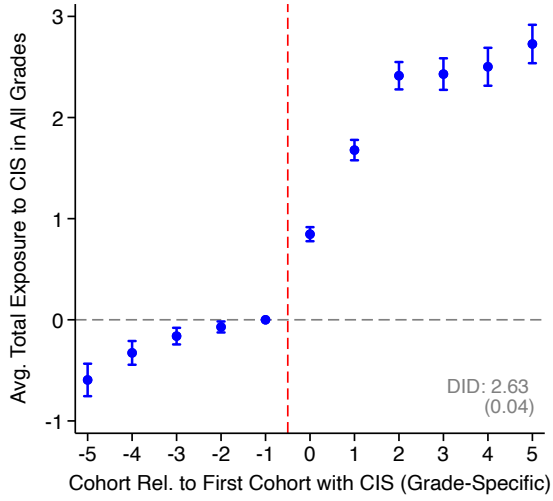
B: Total Exposure to CIS in Middle School



C: Total Exposure to CIS in High School

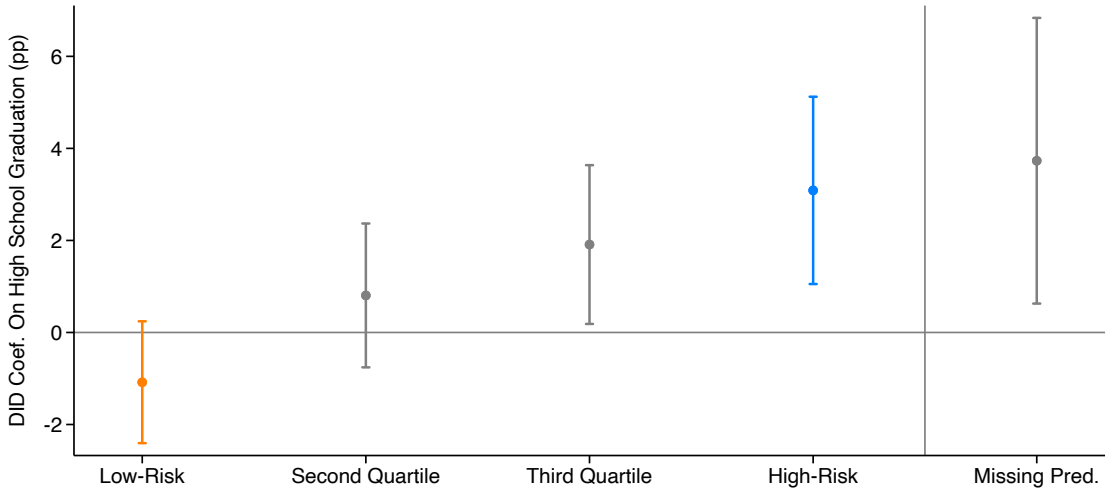


D: Total Exposure to CIS in All Grades



Notes: This figure presents event study estimates of CIS exposure by school level. In each panel, we estimate Equation 3, where the dependent variable is cumulative years of exposure to CIS during elementary school (Panel A), middle school (Panel B), high school (Panel C), and across all grades (Panel D). All specifications use the matched control design and leverage variation from the rollout of CIS to middle schools, as described in Section IV. These estimates assess whether variation in middle school exposure is correlated with exposure at other grade levels. Standard errors are clustered at the school-by-birth-cohort level.

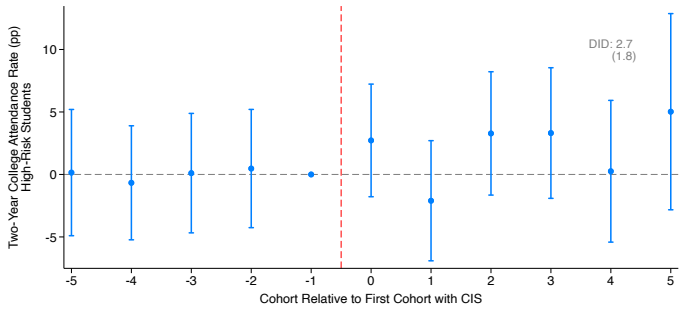
FIGURE A.6: Heterogeneity: Effect of CIS on High School Graduation By Risk Group



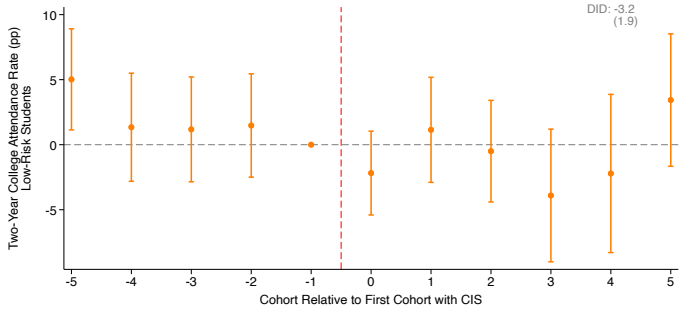
Notes: This figure reports difference-in-differences coefficients estimating the effect of CIS on high school graduation rates, separately by student risk status. For each quartile of the risk distribution—as well as for students without a predicted risk score—we report the coefficient $\beta_{post,3-5}$ from the difference-in-differences analog of Equation 3, comparing treated schools to matched control schools. The estimates for high- and low-risk students correspond to those shown in the event studies in Figure IX. Students are missing a risk score if they lack elementary school data. Standard errors are clustered at the school-by-birth-cohort level.

FIGURE A.7: Effect of CIS on College Enrollment

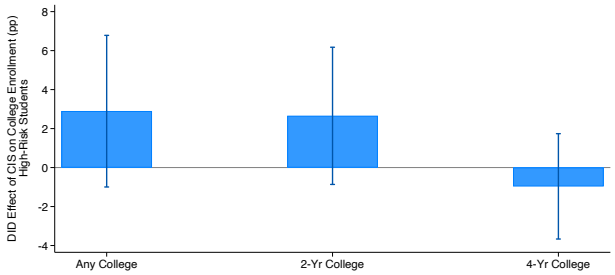
A: Two-Year College Attendance
High-Risk



B: Two-Year College Attendance
Low-Risk

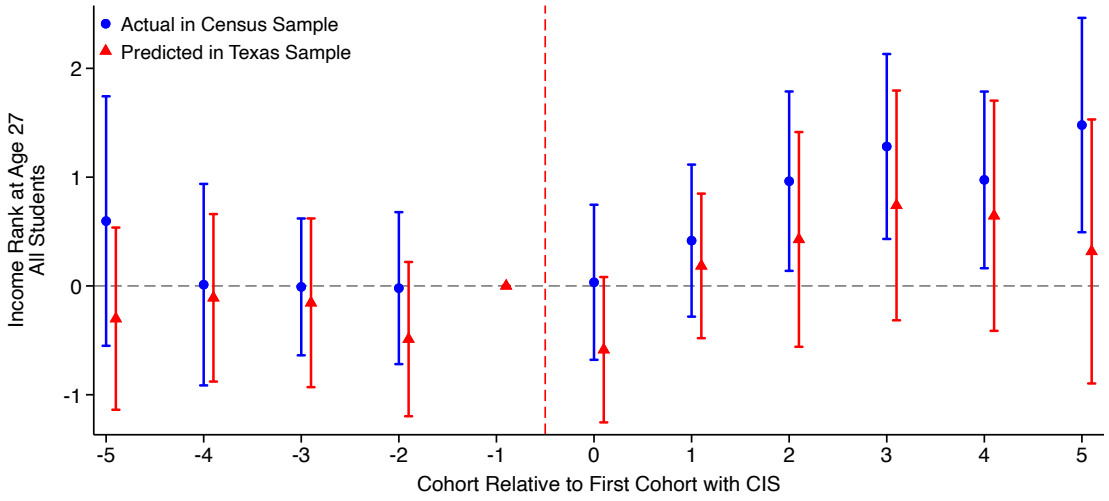


C: Two-Year College Attendance by College Type
High-Risk



Notes: This figure presents results on the effect of CIS on college enrollment. Panels A and B estimate effects on 2-year college attendance, separately by risk group using data on students in middle school between 1998 and 2009. We define college attendance as an indicator for ever enrolling in a 2-year college (in Texas) between ages 18 and 22 by 2019. Panel C presents DID coefficients from three separate regressions showing the effect of CIS adoption on college attendance for high-risk students. The middle bar is the same as that shown in Panel A and restricts to two-year college attendance. The bar on the left captures college attendance at all levels and the bar on the right captures enrollment only at 4-year colleges. In each case, we define enrollment as ever enrolling between the ages of 18-22. Standard errors are clustered at the school-by-birth-cohort level.

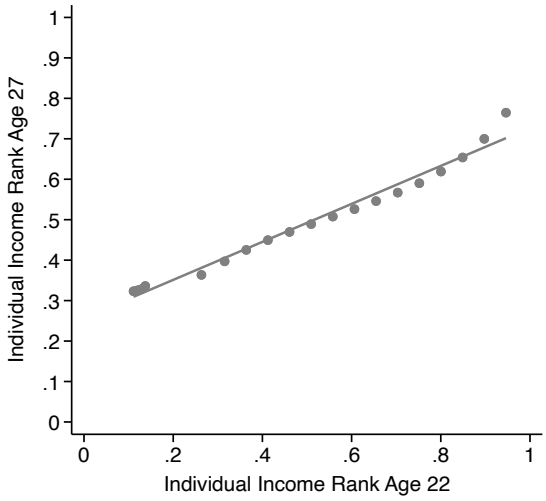
FIGURE A.8: Sample Comparison: Effect of CIS Income Rank at Age 27
All Students



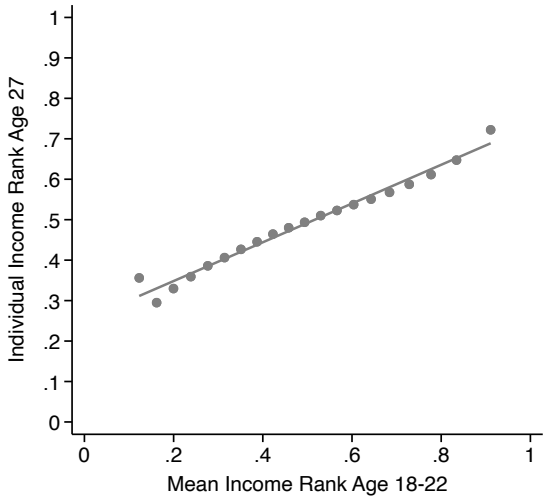
Notes: This figure compares the effect of CIS on adult earnings rank as measured in the national Census-IRS data to the effect on predicted earnings in the Texas sample. The Census series (circles) replicates the estimates from Figure X; see accompanying notes for additional details. The Texas series (triangles) presents coefficients from Equation 3, where the dependent variable is predicted earnings rank at age 27. These predictions are based on observed earnings and college attendance between ages 18 and 22. See Appendix Figure A.9 for further details on the earnings prediction model. Standard errors in the Census sample are clustered at the high school catchment level. Standard errors in the Texas sample are clustered at the school level.

FIGURE A.9: Predicting Individual Income Rank at Age 27

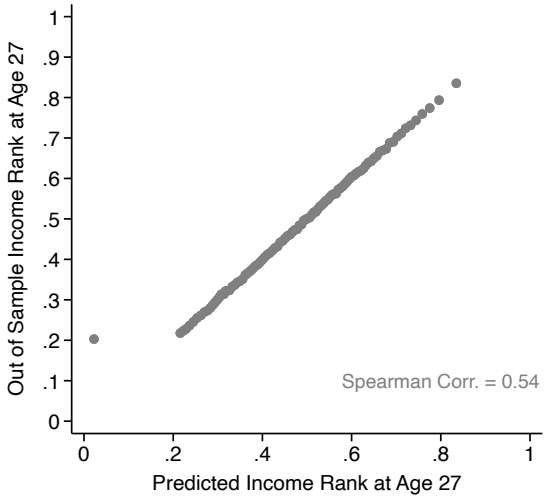
A: Relationship with Age 22 Rank



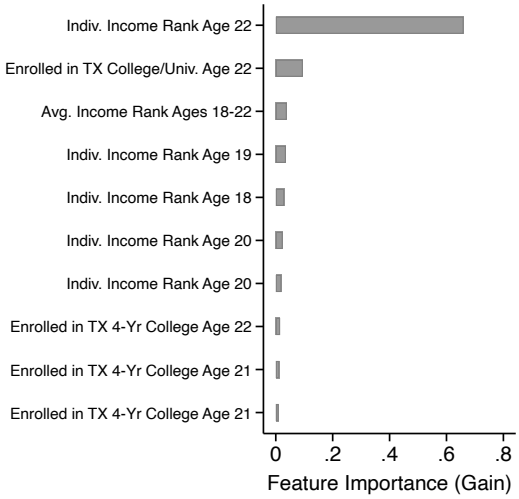
B: Relationship with Mean Rank Ages 18-22



C: Out-of-Sample Accuracy of Age 27 Income Rank

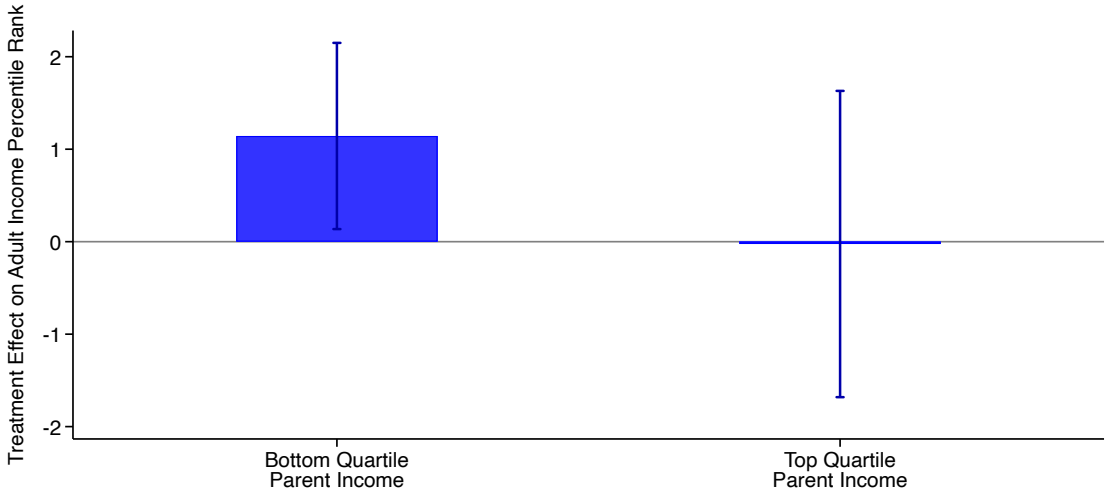


D: Feature Importance – Model Gain



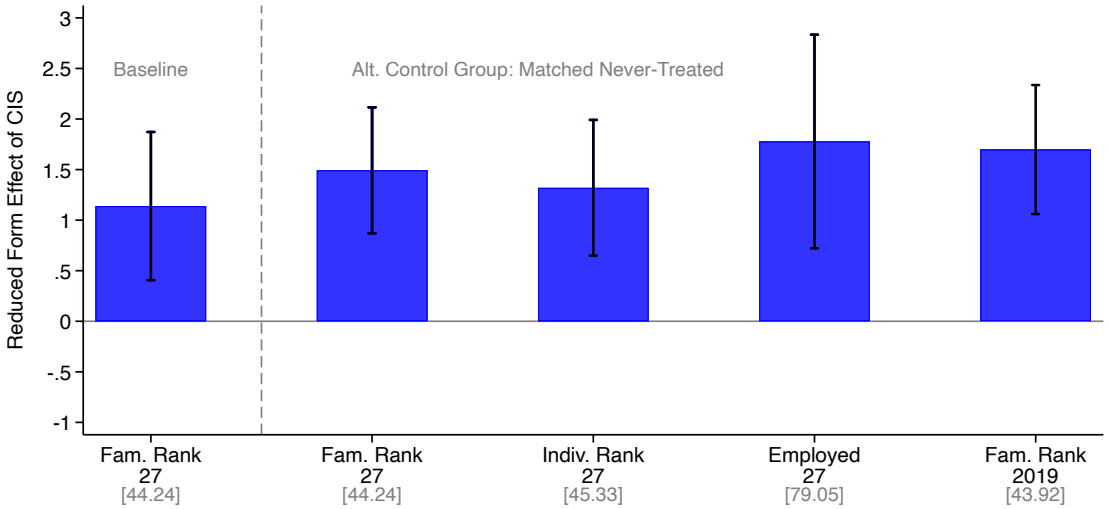
Notes: This figure summarizes the empirical prediction of individual income rank at age 27 using earnings and college attendance data from ages 18 to 22. All plots are based on a 20% hold-out test set not used in model training, using statewide data students born between 1978-1992. Panel A shows a binned scatterplot of the relationship between income rank at age 22 and income rank at age 27. Panel B presents the analogous relationship using mean income rank across ages 18 to 22. Panel C shows a binned scatterplot of observed income ranks against predicted values from a gradient boosting model trained on earlier outcomes. Predictors include income rank at each age (18–22), average rank, change in rank over time, and indicators for 2-year and 4-year college attendance at each age. Panel D reports the top 10 predictors by average model gain.

FIGURE A.10: Effect of CIS Household Income Rank at Age 27
Heterogeneity by Parent Income Quartile



Notes: This figure displays the DID effect of CIS on adult earnings rank as measured in the national Census-IRS data, separately for children with parents in the bottom and top quartiles of the parent income distribution. Standard errors are clustered at the high school catchment zone boundary.

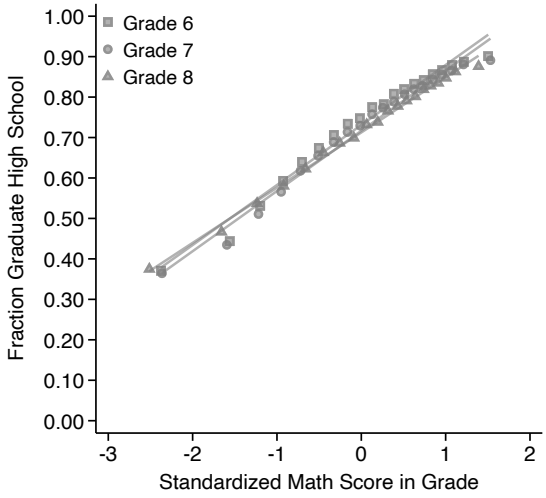
FIGURE A.11: Reduced Form Effect of CIS on Additional Outcomes
Census-IRS Sample: All Children



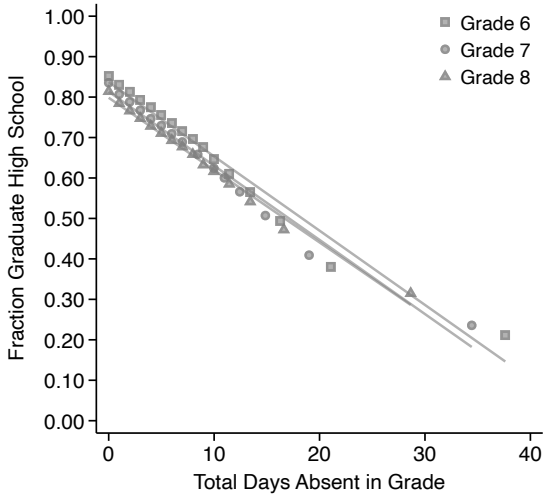
Notes: This figure presents effects of CIS adoption using the Census-IRS data described in Section III C. The first bar reproduces the DID coefficient shown in Figure X Panel *B*, which is estimated using a control group of later-treated CIS tracts and Equation 4, as described in Section V D. The remaining bars show robustness to an alternative control group and present results for additional outcomes. Coefficients to the right of the dashed line are estimated using a control group of never-treated neighborhoods, matched to treated neighborhoods following a procedure similar to the one described in Section IV. The match is restricted to neighborhoods in a different commuting zone than the treated school and is based on the share of children from White families, the population density in the 2000 Census, the level of parent income and adult household income rank in pre-exposure cohorts, and the trend in adult household income rank in pre-exposure cohorts. In addition to the effect of CIS on household income rank at age 27, we also report effects on individual income rank at age 27, employment at age 27 (measured by the presence of W-2 income), and household income rank in 2019, when these cohorts are in their mid- to late-thirties. In each regression, standard errors are clustered at the high school catchment zone boundary level, and the pre-treatment means of each outcome are reported in brackets below the axis labels.

FIGURE A.12: Predicting High School Graduation

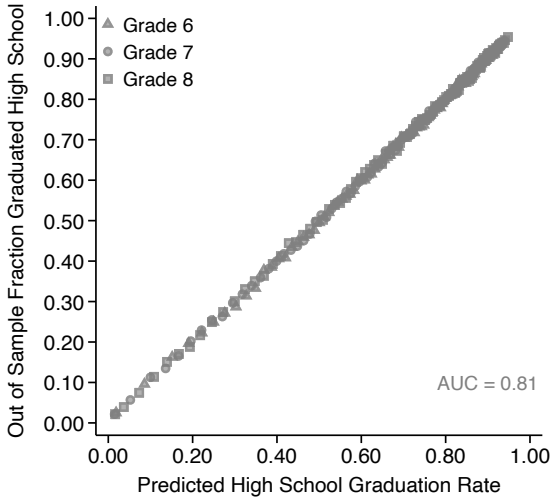
A: Relationship with Test Scores



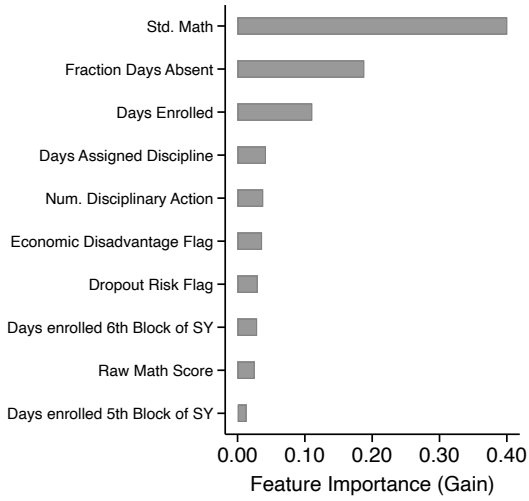
B: Relationship with Attendance



C: Out-of-Sample Accuracy of HS Graduation Prediction



D: Feature Importance – Model Gain



Notes: This figure summarizes the empirical prediction of high school graduation using academic and behavioral outcomes observed in middle school. All plots are based on a 20% hold-out test set not used in model training, using statewide data on middle school students from 1998–2004. Panel A shows a binned scatterplot of the relationship between standardized math scores and high school graduation, separately for grades 6–8. Panel B presents the analogous relationship using days absent in each grade. Panel C shows a binned scatterplot of observed graduation rates against predicted graduation probabilities from grade-specific gradient boosting models. These models are trained on student characteristics and outcomes measured in grade g, including demographics (race, sex, FRPL, dropout-risk flag, economic disadvantage flag), standardized math scores, attendance (at six-week intervals and overall), and disciplinary records (in-/out-of-school suspensions, days assigned, number of infractions). Panel D reports the top 10 predictors by average model gain across the grade-specific models.