By Nicholas Gage,
Tracy Ulrich, and
Lauren Wong

Effects of the California Integrated Supports Project on Student Discipline and Attendance:

A Staggered Difference-in-Differences Study

May 2026

Contents

Abstract

This study examined the impact of the California Integrated Supports Project (CA-ISP), a statewide professional learning initiative designed to strengthen multi-tiered systems of support, on student disciplinary outcomes and attendance. Using administrative data from 27 school districts, the analytic sample included 646,010 student-year observations representing 291,317 students across 381 schools over four academic years (2021–2022 through 2024–2025). The intervention was implemented using a staggered cohort rollout, allowing program effects to be estimated using weighted multilevel difference-in-differences models with random intercepts at the school and district levels. Outcomes included the probability of experiencing any disciplinary incident, in-school suspension (ISS), out-of-school suspension (OSS), total incident counts, suspension days, and student attendance rates. Results indicated modest but meaningful reductions in disciplinary incidents during initial implementation periods, with particularly large reductions observed in the first year of implementation for both cohorts. The intervention also produced reductions in ISS days, suggesting changes in the severity of disciplinary responses. Effects on OSS outcomes and attendance were small and inconsistent across cohorts. Moderator analyses indicated that impacts were substantially larger for students with elevated baseline disciplinary risk, suggesting that universal schoolwide supports may produce their greatest benefits for students already at risk of disciplinary involvement. Exploratory subgroup analyses suggested some variation in program impacts across racial and ethnic groups. Findings contribute new evidence regarding the potential for large-scale implementation initiatives to influence student behavioral outcomes across diverse school systems.

Keywords: school discipline, positive behavioral interventions and supports, multi-tiered systems of support, difference-in-differences, school climate.

Overview

Student disciplinary exclusion remains a persistent challenge in United States schools. Disciplinary practices such as office referrals, suspensions, and classroom removals are widely used to address behavioral challenges but have been associated with a range of negative academic and developmental outcomes for students. Research has linked exclusionary discipline to reduced instructional time, lower academic achievement, and increased risk of school disengagement and dropout (Gregory et al., 2010; Skiba et al., 2011). As a result, educators and policymakers have increasingly sought preventative and systemic approaches to improving student behavior and school climate while reducing reliance on exclusionary discipline.

One widely adopted approach to improving school climate and student behavior is the use of multi-tiered systems of support (MTSS). MTSS frameworks provide coordinated academic, behavioral, and social–emotional supports through a tiered prevention model in which universal supports are delivered to all students and more intensive interventions are provided to students with greater needs. The goal of MTSS is to create an integrated system that addresses academic and behavioral challenges while promoting positive school experiences for all students. By aligning resources, data systems, and instructional practices, MTSS frameworks aim to improve both academic and social–emotional outcomes across the entire student population.

Within MTSS, schoolwide Positive Behavioral Interventions and Supports (PBIS) represents one of the most widely implemented frameworks for improving school climate and student behavior. PBIS emphasizes proactive behavioral expectations, consistent reinforcement of positive behavior, and data-driven decision-making to guide intervention and support (Sugai & Horner, 2002). A substantial body of research has demonstrated that PBIS implementation can reduce office discipline referrals and suspensions while improving social behavior, classroom climate, and teacher perceptions of student behavior. For example, randomized controlled trials have found that schools implementing PBIS experience improvements in student behavioral outcomes and reductions in disciplinary incidents relative to comparison schools (Bradshaw et al., 2010; Bradshaw et al., 2012). Additional research suggests that PBIS can contribute to improvements in school climate, student engagement, and social–emotional functioning (McIntosh et al., 2016).

Despite the growing evidence base supporting PBIS and MTSS frameworks, implementing these systems with fidelity across large and diverse school systems remains challenging. Successful implementation often requires sustained professional learning, leadership engagement, and ongoing technical assistance to help schools integrate behavioral, social–emotional, and academic supports. In response to these challenges, statewide initiatives have increasingly focused on providing structured professional development and implementation support to help schools adopt and sustain integrated systems of support.

The California Integrated Supports Project (CA-ISP) represents one such initiative. CA-ISP is a statewide professional learning initiative designed to strengthen school systems that support students' academic, social–emotional, and behavioral development. The project provides professional learning and implementation support for educators in areas including social–emotional learning (SEL), trauma-informed practices, and culturally relevant and sustaining instructional practices aligned with

California's MTSS framework. These supports are delivered through partnerships among regional technical assistance providers and educational agencies, with the goal of building school capacity to create supportive learning environments and improve student outcomes.

Large-scale professional learning initiatives that focus on improving school climate and behavioral supports may influence student outcomes through several pathways. Universal behavioral expectations and consistent reinforcement systems may reduce disruptive behavior and disciplinary incidents, while improved teacher capacity to respond to behavioral challenges may reduce reliance on exclusionary discipline. In addition, SEL and trauma-informed practices can strengthen students' self-regulation and engagement with school, which may contribute to improvements in attendance and classroom participation (Durlak et al., 2011).

Despite substantial investments in initiatives designed to strengthen schoolwide support systems, there remains limited empirical evidence examining whether large-scale implementation initiatives produce measurable improvements in student outcomes when implemented across diverse school contexts. Much of the existing research on PBIS and MTSS has focused on individual schools or smaller district-level implementations. Evaluations using large administrative datasets can provide important evidence about whether statewide professional learning initiatives translate into measurable improvements in student behavioral and attendance outcomes.

The present study addresses this gap by examining the impact of CA-ISP on student disciplinary outcomes and attendance using a longitudinal administrative dataset spanning four academic years. The intervention was implemented using a staggered cohort rollout across schools, allowing program impacts to be evaluated using a difference-in-differences design that compared outcomes in treatment schools with outcomes in comparison schools over time. This design provides a rigorous approach to estimating program effects while accounting for baseline differences across schools and broader temporal trends in student outcomes.

Study Purpose and Research Questions

This study examined the impact of CA-ISP, a professional learning and implementation support initiative designed to strengthen multi-tiered systems of support and positive school climates, on student behavioral and attendance outcomes. Using a staggered difference-in-differences design across two implementation cohorts, this study evaluated whether participation in CA-ISP was associated with changes in students' disciplinary experiences and attendance patterns relative to comparison schools that did not receive the intervention during the study period. Specifically, the study examined whether the intervention reduced the likelihood of disciplinary incidents and suspensions, reduced the frequency and severity of disciplinary consequences, and improved student attendance rates. In addition, the study examined whether program impacts varied across implementation cohorts and years of exposure, and whether effects differed for students with elevated baseline disciplinary risk and across racial and ethnic groups.

The study addressed the following research questions (RQs):

- RQ1. Does participation in CA-ISP reduce the likelihood that students experience disciplinary incidents, in-school suspensions (ISS), or out-of-school suspensions (OSS)?
- RQ2. Does participation in CA-ISP reduce the frequency and severity of disciplinary consequences, including the total number of disciplinary incidents and the number of days students spend in ISS or OSS?
- RQ3. Does participation in CA-ISP improve student attendance rates?
- RQ4. Do program effects vary across implementation cohorts and years of exposure to the intervention?
- RQ5. Are program impacts larger for students with elevated baseline disciplinary risk?
- RQ6. Do program effects differ across student racial and ethnic groups?

Method

Participants and Data Sources

Student-level administrative data were obtained through data-sharing agreements with participating school districts. Data were requested from all schools participating in CA-ISP as well as comparison schools within the same districts that did not receive the intervention during the study period. The dataset included student enrollment records, demographic characteristics, attendance records, and disciplinary incident data across four academic years (2021–2022 through 2024–2025).

The full analytic dataset included 646,010 student-year observations representing 291,317 unique students across 381 schools in 27 school districts. Schools participating in the intervention were organized into two implementation cohorts based on the timing of program adoption. Cohort 1 included 42 schools serving 41,250 students and began implementation during the 2023–2024 academic year. Cohort 2 included 40 schools serving 36,954 students and began implementation during the 2024–2025 academic year. The comparison group consisted of 299 schools serving 233,425 students that did not participate in the intervention during the study period.

Across the four-year observation period, Cohort 1 schools contributed 79,251 student-year observations, Cohort 2 schools contributed 65,209 student-year observations, and comparison schools contributed 501,550 student-year observations. These data were used to construct the longitudinal analytical file used in the staggered difference-in-differences analyses.

Student demographic characteristics were obtained from district administrative records and included grade band, gender, race and ethnicity, disability status, English Learner status, and student mobility indicators. These variables were used as covariates in the outcome models and were also included in the estimation of propensity score weights to improve baseline comparability across treatment and comparison schools.

As shown in Table A1 in the Appendix, the sample included students across all grade bands, with the largest proportion enrolled in grades 1–5 (35.1%), followed by grades 9–12 (25.6%) and grades 6–8 (21.0%). The gender distribution was balanced, with 48.3 percent female and 51.5 percent male students. Slightly more than half of students identified as Hispanic (52.9%), and 49.1 percent identified as White, with smaller proportions identifying as Black (7.7%), Asian (15.9%), or other racial groups. Approximately 13.7 percent of students identified as having a disability, and 37.5 percent had ever been classified as English Learners. Student mobility during the study period was relatively low (2.4%). Baseline characteristics varied somewhat across cohorts, particularly in grade distribution and racial and ethnic composition, supporting the use of propensity score weighting to improve baseline comparability across treatment and comparison groups.

Intervention and Implementation Exposure

CA-ISP is a statewide initiative designed to strengthen implementation of MTSS through professional learning focused on SEL, trauma-informed practices, and culturally relevant and sustaining instructional practices. The initiative is implemented through partnerships among county offices of education and research organizations to provide professional development, technical assistance, and implementation supports aligned with California's MTSS framework. CA-ISP is designed to build educator capacity to create supportive school environments that integrate behavioral, social–emotional, and academic supports. The project includes structured professional learning modules and training activities that emphasize identity, supportive learning environments, student voice, situational appropriateness, and data-informed decision-making to promote equitable outcomes for students.

Implementation occurred using a staggered cohort rollout. Schools in Cohort 1 began implementation during the 2023–2024 academic year, while schools in Cohort 2 began implementation during the 2024–2025 academic year. Control schools did not participate in the intervention during the study period. Because implementation began at various times across cohorts, exposure to the intervention varied by school and year. To capture this staggered rollout, analytic models included indicators for Cohort 1 in its first implementation year (2023–2024), Cohort 1 in its second implementation year (2024–2025), and Cohort 2 in its first implementation year (2024–2025). These exposure indicators allowed treatment effects to be estimated separately for each implementation period.

Measures

Disciplinary Incidents

Student disciplinary incidents were measured using administrative discipline records collected by the participating districts. Disciplinary involvement was captured using two indicators. First, a binary indicator of any disciplinary incident was created to identify whether a student experienced at least one recorded behavioral incident during the academic year. Second, a count measure of total disciplinary incidents captured the total number of incidents recorded for each student during the academic year. Because the distribution of incident counts was highly skewed with substantial zero inflation, the count variable was log-transformed using a $\log(1 + x)$ transformation prior to modeling.

In-School Suspension

ISSs were captured using two measures. A binary indicator of any ISS identified whether a student received at least one ISS during the academic year. In addition, a count measure of ISS days captured the number of school days the student spent in ISS. As with incident counts, the distribution of ISS days was highly skewed; therefore, ISS days were log-transformed using log(1 + x) for analysis.

Out-of-School Suspension

OSS outcomes were measured using binary and count indicators. The binary measure of any OSS indicated whether a student received at least one OSS during the academic year. The count measure of OSS days captured the total number of school days the student was suspended. Similar to other disciplinary count variables, OSS days were log-transformed using log(1 + x) to address skewness and improve model stability.

Student Attendance

Student attendance was measured using an annual attendance rate, calculated as the proportion of instructional days attended by the student during the academic year. The attendance rate ranged from 0 to 1, with higher values indicating better attendance.

Student Demographic Characteristic

Student demographic characteristics were drawn from administrative enrollment records and included grade band, gender, race, Hispanic ethnicity, disability status, English Learner (EL) status, and student mobility. Grade band captured the student's grade grouping during the academic year. Gender was coded using district administrative records. Race and Hispanic ethnicity were recorded as reported in administrative data systems. Disability status indicated whether a student was identified as receiving special education services. English Learner status identified students designated as English Learners. Student mobility captured whether the student changed schools during the academic year. These demographic variables were included as covariates in all outcome models and also used in the estimation of propensity score weights to improve baseline comparability across treatment Cohorts 1 and 2 and Control schools.

Baseline Disciplinary Risk

To examine heterogeneous program effects, a measure of baseline disciplinary risk was constructed using each student's earliest observed record in the dataset. Students were classified as having elevated baseline disciplinary risk if they experienced at least one disciplinary incident during their baseline year. This indicator was used in moderator analyses examining whether program effects differed for students with prior disciplinary involvement.

Race and Ethnicity Subgroups

Exploratory subgroup analyses examined variation in program impacts across race and ethnicity groups. Using students' baseline demographic records, a mutually exclusive four-category race-and-ethnicity variable was created: White, Black, Hispanic, and Other. The "Other" category included students identifying with racial groups not included in the first three categories, as well as multiracial students.

Data Analysis

Program impacts were estimated under a staggered implementation design in which Cohort 1 schools began implementing CA-ISP in the 2023–2024 academic year and Cohort 2 schools began implementing in the 2024–2025 academic year. Control schools did not implement the intervention during the study period. The analytic sample was restricted to students enrolled in Control, Cohort 1, or Cohort 2 schools, and students who attended multiple schools spanning different treatment conditions were excluded to avoid mixed treatment exposure.

The academic year was coded as a four-level factor (2021–2022, 2022–2023, 2023–2024, 2024–2025). A three-level cohort variable distinguished Control schools from treatment schools in Cohorts 1 and 2. Using these variables, several treatment indicators were constructed to represent the staggered rollout of the intervention.

First, a pooled treatment indicator identified whether a student attended a treatment school at any point during the study period. Second, cohort-specific post indicators were created to reflect the timing of program adoption for each cohort. For Cohort 1 schools, post-implementation years included 2023–2024 and 2024–2025, whereas for Cohort 2 schools, the post-implementation year was 2024–2025. These indicators were used to construct pooled and cohort-specific difference-in-differences variables.

To allow treatment effects to vary across implementation years, additional exposure-year indicators were constructed representing Cohort 1 during its first implementation year (2023–2024), Cohort 1 during its second implementation year (2024–2025), and Cohort 2 during its first implementation year (2024–2025). These exposure-year indicators formed the primary treatment variables in the staggered difference-in-differences models.

Propensity Score Weighting

To improve baseline comparability across treatment Cohorts 1 and 2 and Control schools, stabilized inverse probability of treatment weights were estimated using multinomial propensity score models. A baseline file was created by selecting each student's earliest observed record in the analytic dataset. Multinomial logistic regression was then used to estimate the probability of cohort membership (Control, Cohort 1, or Cohort 2) as a function of baseline student characteristics:

$$Pr(Cohort_i = c \mid X_i)$$

where X_i included

- grade band,
- gender,
- Hispanic ethnicity,
- race and ethnicity,
- disability status.
- English Learner status, and
- student mobility.

Stabilized weights were constructed as

$$w_i = \frac{Pr(Cohort = c)}{Pr(Cohort_i = c \mid X_i)}$$

where $\Pr(Cohort = c)$ represents the marginal probability of cohort membership. Extreme weights were trimmed at the 1st and 99th percentiles to reduce the influence of extreme observations. These weights were then merged back to all student-year observations and applied in subsequent outcome models.

Outcome Models

Program impacts were estimated using weighted linear mixed-effects difference-in-differences models. Random intercepts were included for schools and districts to account for clustering of students within institutions.

All models included student-level covariates

X_{it}=(grade band, gender,race, Hispanic, disability, EL, mobility

and fixed effects for the academic year.

The general model structure was

$$Y_{ist} = \beta_0 + \mathbf{T}_{it}\boldsymbol{\beta} + \mathbf{Year}_t + \boldsymbol{\gamma}'X_{it} + u_s + v_d + \varepsilon_{ist}$$

where

- Y_{ist} is the outcome for student i in school s, district d, year t;
- u_s is a school random intercept;
- v_d is a district random intercept; and
- $\in_{ist}$ is the residual error.

The vector T_{it} represents treatment indicators that differ across model specifications.

Pooled First-Year Intent-to-Treat Impact

The primary analysis estimated the average first-year impact of CA-ISP using a stacked two-period difference-in-differences design. For Cohort 1 schools, the comparison contrasted the common pre-implementation year (2022–2023) with the first implementation year (2023–2024). For Cohort 2 schools, the comparison contrasted the same baseline year (2022–2023) with the first implementation year (2024–2025). Control schools contributed observations to both comparisons. The resulting datasets were stacked to estimate a pooled first-year treatment effect.

The estimating equation was

$$Y_{ist} = \beta_0 + \beta_1 Treat_i + \beta_2 Post_t + \beta_3 (Treat_i \times Post_t)$$
$$+ \beta_4 Cohort_i + \boldsymbol{\gamma}'X_{it} + u_s + v_d + \varepsilon_{ist}$$

where

- $Treat_i$ indicates treatment school membership,
- $Post_t$ indicates the first implementation year for each cohort, and
- $Cohort_i$ identifies whether observations belong to the Cohort 1 or Cohort 2 stacked sample.
- The coefficient $ß_3$ represents the pooled first-year intent-to-treat effect across both cohorts.

Fully Pooled Implementation-Period Effects

A secondary analysis estimated the average program impact across all implementation years using the full longitudinal dataset. For Cohort 1 schools, the post-implementation period included 2023–2024 and 2024–2025. For Cohort 2 schools, the post-implementation period included 2024–2025. Control schools contributed observations during the same calendar years.

The estimating equation was

$$Y_{ist} = \beta_0 + \beta_1 Treat_i + \beta_2 Year_t + \beta_3 (Treat_i \times Post_{it})$$
$$+ \boldsymbol{\gamma}'X_{it} + u_s + v_d + \varepsilon_{ist}$$

Where $Post_{it}$ indicates whether the student was observed during an implementation year for their cohort. The coefficient ß3 represents the average treatment effect across all observations in the implementation period.

Exposure-Year Effects

To examine whether impacts varied across implementation years, a staggered difference-in-differences model with cohort-specific exposure indicators was estimated:

$$Y_{ist} = \beta_0 + \beta_1 C1_i + \beta_2 C2_i + \beta_3 Year_t + \beta_4 D_{C1Y1,it}$$
$$+\beta_5 D_{C1Y2,it} + \beta_6 D_{C2Y1,it} + \boldsymbol{\gamma}' X_{it} + u_s + v_d + \varepsilon_{ist}$$

where

- $D_{C1Y1,it}$ indicates exposure during Cohort 1 Year 1,
- $D_{C1Y2,it}$ indicates exposure during Cohort 1 Year 2, and
- $D_{C2Y1,it}$ indicates exposure during Cohort 2 Year 1.

The coefficients β_4, β_5, β_6 represent the staggered difference-in-differences estimates corresponding to each cohort-year exposure period.

Two supplementary specifications were also estimated:

- pooled treatment models with a single treatment-by-post indicator, and
- cohort-specific pooled models estimating separate treatment effects for Cohort 1 and Cohort 2.

These models were used to confirm the robustness of results across alternative parameterizations of the staggered difference-in-differences design.

Models for Disciplinary Counts

For disciplinary count outcomes—including total incidents, ISS days, and OSS days—the distributions were highly skewed with substantial zero inflation. Accordingly, outcomes were transformed using the log-plus-one transformation:

$$\log(1 + Y_{ist})$$

The same model described above was applied to these transformed outcomes:

$$\log(1 + Y_{ist}) = \beta_0 + \beta_1 C1_i + \beta_2 C2_i + \beta_3 Year_t + \beta_4 D_{C1Y1,it}$$
$$+\beta_5 D_{C1Y2,it} + \beta_6 D_{C2Y1,it} + \boldsymbol{\gamma}' X_{it} + u_s + v_d + \varepsilon_{ist}$$

To facilitate interpretation, coefficients were converted to approximate percentage changes using

$$100 \times (e^{\beta} - 1)$$

Descriptive Analyses and Visualization

Descriptive statistics were computed by cohort and academic year for all primary outcomes, including

- mean and median disciplinary incident counts,
- the proportion of students experiencing any disciplinary incident,
- the proportion experiencing any ISS,
- the proportion experiencing any OSS, and
- average attendance rates.

Line plots were created to illustrate cohort trends over time in disciplinary incidents and suspension rates for Control, Cohort 1, and Cohort 2 schools.

Moderator Analyses

Two sets of moderator analyses examined heterogeneity in program effects.

First, moderation by baseline disciplinary risk was examined. A baseline indicator identified whether a student had experienced any disciplinary incident in their first observed year. Interaction terms between this baseline risk indicator and treatment exposure indicators were added to the models. Stratified models were also estimated separately for students with and without baseline disciplinary incidents.

Second, exploratory subgroup analyses examined variation in program impacts across racial and ethnic groups. A mutually exclusive race-and-ethnicity classification was constructed using students' earliest demographic records. Separate weighted mixed-effects models were estimated within each subgroup for key outcomes, including the probability of any disciplinary incident, log-transformed incident counts, and log-transformed ISS days.

Justification for Linear Probability Models

Binary outcomes—including the probability of any disciplinary incident, any ISS, and any OSS—were estimated using linear probability mixed models rather than logistic mixed models. This approach was selected because the analytic models incorporated stabilized inverse probability weights, multiple treatment indicators reflecting the staggered rollout, and nested random effects for schools and districts, which produced convergence instability when estimated using generalized linear mixed-effects models.

Linear probability models provide directly interpretable difference-in-differences estimates in percentage-point units and are commonly used in policy evaluation settings with large samples and complex fixed-effects structures. In the present data, predicted probabilities remained well within the unit interval, suggesting that the linear approximation was appropriate.

Sensitivity analyses across alternative difference-in-differences specifications—including pooled treatment models, cohort-specific models, and exposure-year models—produced substantively similar conclusions, supporting the robustness of the findings.

Results

Descriptive Trends

Table A2 in the Appendix presents descriptive statistics for all study outcomes by cohort and academic year, including the proportion of students experiencing any disciplinary incident, any ISS, and any OSS, as well as mean incident counts, ISS days, OSS days, and attendance rates.

Across the study period, disciplinary incidents increased in Control schools from 5.6 percent in 2021–2022 to 12.8 percent in 2022–2023, remained elevated at 14.1 percent in 2023–2024, and then declined slightly to 13.1 percent in 2024–2025. Mean incident counts followed a similar pattern, increasing steadily from 1.42 incidents per student in 2021–2022 to 7.46 incidents in 2024–2025.

In Cohort 1 schools, incident rates remained below those observed in Control schools throughout the study period. Rates increased from 4.5 percent in the pre-implementation year (2022–2023) to 4.9 percent in the first implementation year (2023–2024) and 7.6 percent in the second implementation year (2024–2025). Mean incident counts in Cohort 1 schools similarly increased over time, rising from 2.44 to 6.24 incidents per student across the same period.

In Cohort 2 schools, incident rates rose sharply to 19 percent in 2023–2024, the year immediately preceding implementation, and then declined to 8 percent during the first implementation year (2024–2025). In contrast to the percentage-based measure, mean incident counts in Cohort 2 schools showed a more moderate pattern, increasing from 1.45 in 2021–2022 to 3.27 in 2024–2025.

Rates of ISS and OSS were low across cohorts and years. The proportion of students receiving any ISS generally remained below 1 percent, and mean ISS days were near zero across all groups. OSS rates were also modest, ranging from approximately 2 to 5 percent across cohorts and years, with mean OSS days per student remaining below 0.15 days per year.

Attendance rates were relatively stable over time across all cohorts. Control schools showed gradual increases in attendance from 0.782 in 2021–2022 to 0.881 in 2024–2025. Cohort 1 schools exhibited similar improvements, increasing from 0.809 to 0.890, while Cohort 2 schools showed more modest fluctuations, ranging from 0.826 to 0.866 across the study period. These descriptive patterns provide context for the adjusted impact models reported below.

Main Impact Analyses

Pooled First-Year Intent-to-Treat Effects

Table A3 in the Appendix presents results from the pooled first-year intent-to-treat (ITT) model, which estimated the average effect of CA-ISP during the first year of implementation across the two cohorts. In this specification, participation in CA-ISP was associated with a statistically significant improvement in student attendance. Students in treatment schools demonstrated attendance rates that were 0.013 points higher than those in comparison schools (SE = 0.002, $p < .001$), corresponding to anapproximate 1.3 percentage-point increase in attendance.

No statistically significant first-year effects were observed for disciplinary outcomes. The estimated effect on the probability of experiencing any disciplinary incident was small and not statistically significant (b = 0.003, p = .277). Similarly, the estimated effects on the probability of any ISS, the probability of any OSS, and the number of ISS and OSS days were near zero. For total disciplinary incidents, the estimated effect was negative and approached statistical significance (b = -0.047, p = .066), suggesting a modest reduction in incident counts during the first year of implementation. However, this estimate did not meet conventional thresholds for statistical significance. Taken together, the pooled first-year ITT results suggest that the earliest detectable program effect was concentrated in student attendance rather than disciplinary outcomes.

Fully Pooled Implementation-Period Effects

Table A4 in the Appendix presents results from the fully pooled implementation-period model, which estimated the average treatment effect across all years in which schools were implementing CA-ISP. In contrast to the first-year ITT model, this specification revealed a statistically significant reduction in the likelihood of experiencing any disciplinary incident. Students in treatment schools were 0.024 points less likely to experience a disciplinary incident than students in comparison schools (SE = 0.002, $p < .001$), corresponding to an approximately 2.4 percentage-point reduction in the probability of a disciplinary incident.

Effects on suspension outcomes remained small and not statistically significant. No meaningful pooled effects were observed for the probability of any ISS, the probability of any OSS, ISS days, or OSS days. The attendance effect observed in the first-year ITT model attenuated in the fully pooled model and was only marginally significant (b = 0.002, $p = .068$).

Taken together, the two pooled specifications suggest a temporal pattern in program effects. The first-year ITT model indicates an early improvement in attendance. In contrast, the fully pooled implementation-period model indicates that reductions in disciplinary incidents became more apparent when impacts were averaged across the broader implementation period.

Cohort-Specific and Exposure-Year Results

To better understand the timing of these pooled effects, we next examined cohort-specific pooled models and exposure-year models for each outcome. Detailed exposure-year estimates for all outcomes are presented in Table A5 in the Appendix.

Any disciplinary incident: For the probability of experiencing any disciplinary incident, the pooled treatment model indicated a statistically significant reduction following implementation ($b = -0.024$, $SE = 0.002$, $p < .001$), corresponding to an approximate 2.4 percentage-point decrease relative to comparison schools. Cohort-specific models suggested that this reduction was not uniform across cohorts. The pooled effect for Cohort 1 was not statistically significant, whereas Cohort 2 showed a substantial reduction in the likelihood of incident cases ($b = -0.050$, $SE = 0.002$, $p < .001$).

The exposure-year models clarified this pattern. In Cohort 1 schools, the probability of experiencing any disciplinary incident declined during the first implementation year ($b = -0.016$, $SE = 0.003$, $p < .001$), indicating an approximately 1.6 percentage-point reduction relative to the comparison condition. However, the second implementation year yielded a positive coefficient (b = 0.014, $SE = 0.003$, $p < .001$), indicating that the first-year reduction was not sustained in this specification. In Cohort 2 schools, the first implementation year was associated with a larger decline in incident probability (b = -0.047, SE = 0.002, $p < .001$), corresponding to roughly a 4.7 percentage-point decline. Overall, these findings indicate that CA-ISP was associated with meaningful reductions in disciplinary incidents during initial implementation, particularly in Cohort 2.

In-school suspension: Results for the probability of receiving any ISS were generally small and statistically nonsignificant. The pooled treatment model showed no overall effect ($b \approx 0.000$, $p = .931$), and cohort-specific models indicated only limited variation across cohorts. Cohort 1 showed a marginally negative pooled coefficient, whereas Cohort 2 showed a small positive estimate.

Exposure-year models similarly provided little evidence of a consistent reduction in the probability of receiving ISS. The only suggestive pattern was a marginal reduction for Cohort 1 during the second implementation year ($b = -0.0013$, $SE = 0.0007$, $p = .053$). Given the very low base rates of ISS events across the sample, these null or weak findings are not unexpected.

Out-of-school suspension: Results for OSS were uniformly null across model specifications. The pooled treatment model indicated no statistically significant difference in the probability of experiencing any OSS event ($b = -0.00038$, $p = .706$), and cohort-specific and exposure-year models showed no consistent pattern of reduction across cohorts or implementation periods.

These findings suggest that CA-ISP did not meaningfully alter the likelihood of OSS during the study period.

Attendance: Attendance effects varied by model specification. As noted above, the stacked first-year ITT model showed a statistically significant improvement in attendance during the first implementation year. However, in the fully pooled implementation model, the effect attenuated and did not reach conventional levels of significance. Cohort-specific and exposure-year models likewise suggested small positive but generally nonsignificant coefficients. Accordingly, the attendance findings are most consistent with a short-term improvement during initial implementation rather than a sustained effect across all implementation periods.

Disciplinary incident frequency: To examine whether the intervention influenced the number of disciplinary incidents experienced by students, log-transformed count models were estimated . The pooled implementation model did not show a statistically significant overall effect on incident counts ($b = 0.016$, $p = .312$). Cohort-specific pooled models suggested that Cohort 1 had a positive coefficient relative to comparison schools ($b = 0.066$, $p = .005$), though this average obscures variation across implementation years.

The exposure-year models showed a more differentiated pattern. In Cohort 1 schools, the first implementation year was associated with a reduction in incident counts ($b = -0.059$, $SE = 0.027$, $p = .031$), corresponding to an approximate 5.7 percent decrease. However, the second implementation year showed a positive coefficient ($b = 0.177$, $SE = 0.026$, $p < .001$), indicating higher incident counts than in comparison schools. Cohort 2 did not show a statistically significant first-year effect on incident counts. These results suggest that reductions in incident frequency were not consistent across cohorts or implementation periods and should therefore be interpreted cautiously.

In-school suspension days: Additional models examined the number of days students spent in ISS using log-transformed outcomes. The pooled implementation model did not detect a statistically significant overall effect ($b = -0.00026$, $p = .490$). However, cohort-specific pooled models suggested a significant reduction for Cohort 1 ($b = -0.00138$, $SE = 0.00051$, $p = .007$).

Exposure-year models were broadly consistent with this pattern. In Cohort 1 schools, the number of ISS days declined during both the first implementation year ($b = -0.00124$, $SE = 0.00060$, $p = .038$) and the second implementation year ($b = -0.00153$, $SE = 0.00060$, $p = .011$). Cohort 2 showed no statistically significant effect. Although these coefficients were small in magnitude, they suggest that the intervention may have modestly reduced the duration of ISS in Cohort 1 schools.

Out-of-school suspension days: Models examining OSS days were largely null. The pooled treatment model was not statistically significant ($b = 0.00022$, $p = .881$), and cohort-specific models similarly showed no consistent reductions. Exposure-year models indicated a small positive coefficient for Cohort 1 in the second implementation year. However, the overall pattern did not support a meaningful intervention effect on the duration of OSS.

Heterogeneity Analyses

Moderation by baseline disciplinary risk: To examine whether program impacts differed by students' pre-intervention disciplinary risk, moderator models were estimated using an indicator of whether a student experienced any disciplinary incident in their baseline year. Stratified models were also estimated separately for students with and without prior disciplinary incidents. Full moderator model estimates, stratified by baseline disciplinary risk, are presented in Table A6 in the Appendix.

Across outcomes, intervention effects were substantially stronger for students with elevated baseline disciplinary risk. For the probability of experiencing any disciplinary incident, interaction models indicated large negative effects for higher-risk students during key implementation periods. In Cohort 1's first implementation year, students with prior disciplinary incidents experienced a 16 percentage-point reduction in the probability of a subsequent incident (b = -0.160, SE = 0.015, p < .001). In Cohort 2's first implementation year, the corresponding reduction was even larger, at 35.3 percentage points (b = -0.353, SE = 0.009, p < .001).

By contrast, effects among students without prior disciplinary incidents were much smaller. For example, among lower-risk students, Cohort 1's first implementation year was associated with a 2.3 percentage-point reduction in incident probability (b = -0.023, SE = 0.002, p < .001). At the same time, the corresponding estimate for Cohort 2 was not statistically significant.

Moderator analyses of incident counts and ISS days showed a similar general pattern, with stronger favorable effects concentrated among students who entered the study with prior disciplinary involvement. Overall, these analyses indicate that average treatment effects masked substantial heterogeneity, and that the clearest benefits of CA-ISP were concentrated among students at elevated baseline behavioral risk.

Exploratory subgroup analyses by race and ethnicity: To examine whether exposure-year effects differed across baseline race and ethnicity groups defined as Hispanic, Black, White, and Other, we analyzed exploratory subgroups by race and ethnicity. These results suggested that intervention effects were not uniform across student groups. Full exposure-year estimates by race and ethnicity subgroup are reported in Table A1 in the Appendix.

The clearest favorable patterns emerged for Hispanic students. Hispanic students experienced significant reductions in the probability of any disciplinary incident during Cohort 1's first implementation year (b = -0.013, SE = 0.004, p < .001) and Cohort 2's first implementation year (b = -0.059, SE = 0.003, p < .001). Hispanic students also showed significant reductions in ISS days during both the first and second implementation years in Cohort 1 schools.

Patterns for White, Black, and Other students were more mixed and less consistent across outcomes and implementation periods. White students showed some evidence of reduced incident probability during Cohort 2's first implementation year, whereas Black students did not exhibit consistent reductions across disciplinary outcomes. Students classified in the Other category showed an initial reduction in incident probability during Cohort 1's first implementation year, but this pattern did not persist across later implementation periods.

Because these subgroup analyses were exploratory, they should be interpreted cautiously. Nonetheless, they suggest that intervention effects may have varied across student populations and may have been especially favorable for Hispanic students on selected disciplinary outcomes.

Summary of Findings

Table A7 in the Appendix summarizes the statistically significant effects expressed as percentage-point and percentage changes. Across model specifications, the clearest full-sample findings were an early improvement in attendance during the first year of implementation and a reduction in the probability of disciplinary incidents when impacts were averaged across the implementation period. Effects on suspension outcomes were generally small and not statistically significant. More detailed exposure-year models suggested that program impacts varied across cohorts and over time, and heterogeneity analyses indicated that benefits were concentrated most strongly among students with prior disciplinary involvement. Together, these results suggest that CA-ISP produced modest average impacts overall, with larger and more consistent benefits for students at elevated baseline risk.

Discussion

This study examined the impact of a school-based intervention on student disciplinary outcomes and attendance using a staggered difference-in-differences design across two implementation cohorts. Overall, the results suggest that the intervention produced modest but meaningful improvements in disciplinary outcomes during the early stages of implementation, with the strongest effects concentrated among students with elevated baseline disciplinary risk.

The clearest evidence of program effectiveness emerged for the likelihood of experiencing any disciplinary incident. Both cohorts showed reductions during their first year of implementation, with particularly large reductions observed for Cohort 2. These findings suggest that the intervention may have been effective at reducing the probability that students experienced disciplinary incidents during the initial period of program adoption. The stronger effects observed for Cohort 2 may reflect improvements in implementation supports, organizational readiness, or staff familiarity with the program following the first cohort rollout.

Evidence for sustained reductions in disciplinary incidents over time was more mixed. For Cohort 1, the reduction in incident probability observed during the first implementation year was not maintained during the second year. Descriptive trends indicated that incident rates increased during the second implementation year for Cohort 1 schools, although these rates remained substantially lower than those observed in comparison schools. This pattern suggests that the initial improvements associated with program implementation may have attenuated over time, potentially due to implementation fatigue, changes in student composition, or broader contextual factors affecting disciplinary trends.

The intervention demonstrated more consistent effects on the severity of disciplinary responses, particularly regarding ISS. Cohort 1 schools experienced statistically significant reductions in ISS days during both the first and second implementation years. Although these reductions were small in magnitude at the aggregate level, they suggest that the intervention may have influenced how schools responded to behavioral incidents, potentially encouraging less exclusionary or shorter-duration disciplinary responses even when incidents occurred.

By contrast, the intervention had limited effects on OSS outcomes. Both the probability of experiencing an OSS event and the number of OSS days showed little consistent change across cohorts and implementation years. This finding is consistent with prior research suggesting that OSS decisions may be influenced by district policies, legal requirements, or safety considerations that are less responsive to classroom- or school-level behavioral interventions.

Attendance outcomes also showed little evidence of improvement. Although the estimated effects were generally positive, they were small in magnitude and not statistically significant. This suggests that while the intervention may have influenced student disciplinary experiences, it did not substantially alter patterns of school attendance during the study period.

Importantly, moderator analyses indicated that program effects were substantially stronger for students with elevated baseline disciplinary risk. Students who entered the study with prior disciplinary incidents experienced large reductions in subsequent incidents during key implementation periods, including reductions of approximately 16 percentage points during Cohort 1's first implementation year and more than 35 percentage points during Cohort 2's first implementation year. These results suggest that the intervention functioned as a particularly strong preventive support for students already demonstrating behavioral risk. In contrast, effects among students without prior disciplinary incidents were considerably smaller in magnitude.

Exploratory subgroup analyses by race and ethnicity suggested additional heterogeneity in program impacts. The most consistent favorable patterns were observed for Hispanic students, who experienced reductions in both incident probability and ISS days during key implementation periods. Effects for other racial and ethnic groups were more mixed across outcomes and implementation years. Because these subgroup analyses were exploratory, they should be interpreted cautiously. However, they suggest that future research should examine whether implementation processes or contextual factors contribute to differential program impacts across student populations.

Taken together, these findings suggest that while the intervention produced modest average effects, its most meaningful impacts may lie in reducing disciplinary risk among students who already exhibit behavioral challenges. Universal interventions are often expected to produce relatively small overall effects when implemented across entire student populations. However, the larger reductions observed among higher-risk students suggest that universal behavioral supports may yield their greatest benefits by preventing the escalation of behavioral problems among students already at elevated risk for disciplinary involvement.

Limitations

Several limitations should be considered when interpreting the results of this study.

First, although the staggered difference-in-differences design strengthens causal inference relative to simple observational comparisons, the intervention was not randomly assigned. Schools entered implementation in different cohorts based on program rollout decisions rather than random allocation. While the analytic models included extensive covariate adjustment, cohort indicators, and school-level random effects, unobserved differences between treatment and comparison schools may still have influenced the estimated effects. As with any quasi-experimental design, causal interpretations should therefore be made cautiously.

Second, the study relied on administrative disciplinary records, which capture formally recorded incidents but may not fully reflect students' underlying behavioral experiences. Changes in reporting practices, staff interpretation of behavioral policies, or shifts in school-level disciplinary norms can influence disciplinary data. For example, increases or decreases in incident counts may reflect changes in documentation practices rather than true changes in student behavior. Similarly, reductions in suspension days may reflect changes in disciplinary responses rather than reductions in behavioral incidents themselves.

Third, several disciplinary outcomes examined in this study had relatively low base rates, particularly ISS and OSS events. When outcomes occur infrequently, statistical power to detect meaningful differences can be limited. This constraint may partially explain the lack of statistically significant findings for some outcomes, particularly OSS events and OSS days. Future studies with longer follow-up periods or larger samples may be better positioned to detect changes in rare disciplinary events.

Fourth, the study period captured the early years of program implementation, and the results may therefore reflect initial adoption effects rather than longer-term program impacts. The mixed findings observed between the first and second implementation years for Cohort 1 highlight the possibility that intervention effects may evolve over time as implementation practices stabilize, staff turnover occurs, or school contexts change. Additional follow-up would be valuable to determine whether the observed patterns persist, strengthen, or attenuate as the program matures.

Fifth, the analyses primarily focused on student-level outcomes, without direct measures of implementation fidelity or program dosage. As a result, the study cannot directly assess whether variation in program implementation quality contributed to differences in observed impacts across cohorts or years. Future research integrating fidelity measures or implementation data could provide a clearer understanding of the mechanisms by which the intervention influences disciplinary outcomes.

Finally, although the moderator analyses provided evidence that the intervention was particularly beneficial for students with elevated baseline disciplinary risk, the study was not designed specifically to test targeted intervention strategies. The observed moderation patterns, therefore, should be interpreted as exploratory and hypothesis-generating rather than definitive evidence regarding differential program effectiveness.

References

Bradshaw, C. P., Mitchell, M. M., & Leaf, P. J. (2010). Examining the effects of schoolwide positive behavioral interventions and supports on student outcomes: Results from a randomized controlled effectiveness trial in elementary schools. *Journal of Positive Behavior Interventions*, 12(3), 133–148.

Bradshaw, C. P., Waasdorp, T. E., & Leaf, P. J. (2012). Effects of schoolwide positive behavioral interventions and supports on child behavior problems. Pediatrics, *130*(5), e1136–e1145.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432.

Gregory, A., Skiba, R. J., & Noguera, P. A. (2010). The achievement gap and the discipline gap: Two sides of the same coin? *Educational Researcher*, *39*(1), 59–68.

McIntosh, K., Horner, R. H., Chard, D. J., Dickey, C. R., & Braun, D. H. (2016). Reading skills and function of problem behavior in typical school settings. *Journal of Special Education*, *50*(2), 81–92.

Skiba, R. J., Horner, R. H., Chung, C. G., Rausch, M. K., May, S. L., & Tobin, T. (2011). Race is not neutral: A national investigation of African American and Latino disproportionality in school discipline. *School Psychology Review*, *40*(1), 85–107.

Sugai, G., & Horner, R. (2002). The evolution of discipline practices: School-wide PBIS. *Child & Family Behavior Therapy*, *24*(1–2), 23–50.

Appendix

Table A1. Baseline Student Characteristics by Study Condition

Characteristic	Overall (n = 291,317)	Control (n = 221,442)	Cohort 1 (n = 37,233)	Cohort 2 (n = 32,642)
Grade Band				
PreK and TK	14,880 (5.1%)	10,329 (4.7%)	2,622 (7.0%)	1,929 (5.9%)
K	38,277 (13.1%)	26,190 (11.8%)	7,547 (20.3%)	4,540 (13.9%)
1–5	102,135 (35.1%)	69,392 (31.3%)	19,247 (51.7%)	13,496 (41.3%)
6–8	61,314 (21.0%)	47,319 (21.4%)	6,012 (16.1%)	7,983 (24.5%)
9–12	74,697 (25.6%)	68,198 (30.8%)	1,805 (4.8%)	4,694 (14.4%)
13	14 (<0.1%)	14 (<0.1%)	0 (0.0%)	0 (0.0%)
Gender				
Female	140,692 (48.3%)	106,982 (48.3%)	18,060 (48.5%)	15,650 (47.9%)
Male	150,127 (51.5%)	114,086 (51.5%)	19,148 (51.4%)	16,893 (51.8%)
X or Other	498 (0.2%)	374 (0.2%)	25 (<0.1%)	99 (0.3%)
Hispanic Ethnicity				
Not Hispanic	132,094 (45.3%)	97,357 (44.0%)	20,972 (56.3%)	13,765 (42.2%)
Hispanic	154,110 (52.9%)	121,356 (54.8%)	16,077 (43.2%)	16,677 (51.1%)
Missing	5,113 (1.8%)	2,729 (1.2%)	184 (0.5%)	2,200 (6.7%)
Race				
White	142,910 (49.1%)	111,793 (50.5%)	18,328 (49.2%)	12,789 (39.2%)
Black	22,414 (7.7%)	13,911 (6.3%)	4,288 (11.5%)	4,215 (12.9%)
Asian	46,410 (15.9%)	34,809 (15.7%)	6,874 (18.5%)	4,727 (14.5%)
American Indian/Alaska Native	10,448 (3.6%)	7,542 (3.4%)	1,148 (3.1%)	1,758 (5.4%)
Pacific Islander	2,960 (1.0%)	2,052 (0.9%)	440 (1.2%)	468 (1.4%)
Multiple	469 (0.2%)	321 (0.1%)	103 (0.3%)	45 (0.1%)
Other	91 (<0.1%)	57 (<0.1%)	17 (<0.1%)	17 (<0.1%)
Missing	65,615 (22.5%)	50,957 (23.0%)	6,035 (16.2%)	8,623 (26.4%)
Disability Status				
No/Not reported	251,264 (86.3%)	191,170 (86.3%)	31,908 (85.7%)	28,186 (86.3%)
Disability	40,053 (13.7%)	30,272 (13.7%)	5,325 (14.3%)	4,456 (13.7%)

Characteristic	Overall (n = 291,317)	Control (n = 221,442)	Cohort 1 (n = 37,233)	Cohort 2 (n = 32,642)
English Learner Status				
Never-EL	160,949 (55.3%)	119,204 (53.8%)	23,705 (63.7%)	18,040 (55.3%)
Ever-EL	109,262 (37.5%)	87,392 (39.5%)	11,425 (30.7%)	10,445 (32.0%)
Missing	21,106 (7.2%)	14,846 (6.7%)	2,103 (5.6%)	4,157 (12.7%)
Mobility				
Yes	7,046 (2.4%)	6,697 (3.0%)	122 (0.3%)	227 (0.7%)

Note. Values are n (%). Baseline characteristics were derived from each student's first observed record in the analytic sample.

Table A2. Descriptive Statistics for Disciplinary Outcomes and Attendance by Cohort and Academic Year

Cohort	Academic Year	Any Incident (%)	Any ISS (%)	Any OSS (%)	Incident Count (M)	ISS Days (M)	OSS Days (M)	Attendance Rate (M)
Control								
	2021–22	5.6	0.0	4.6	1.42	0.00	0.13	0.782
	2022–23	12.8	0.6	3.4	5.28	0.01	0.12	0.863
	2023–24	14.1	0.6	2.9	6.67	0.01	0.11	0.870
	2024–25	13.1	0.4	2.8	7.46	0.01	0.10	0.881
Cohort 1								
	2021–22	2.3	0.0	1.2	1.27	0.00	0.03	0.809
	2022–23	4.5	0.5	2.1	2.44	0.01	0.06	0.870
	2023–24	4.9	0.5	2.0	2.60	0.01	0.06	0.882
	2024–25	7.6	0.3	2.0	6.24	0.00	0.06	0.890
Cohort 2								
	2021–22	3.6	0.0	2.1	1.45	0.00	0.07	0.837
	2022–23	8.7	0.5	3.4	2.54	0.01	0.11	0.826
	2023–24	19.0	0.3	2.7	2.04	0.00	0.08	0.862
	2024–25	8.0	0.2	2.2	3.27	0.00	0.07	0.866

Note. Values for Any Incident, Any ISS, and Any OSS represent the percentage of students experiencing each outcome during the academic year. Values for Incident Count, ISS Days, OSS Days, and Attendance Rate represent annual means.

Table A3. Pooled First-Year Intent-to-Treat Effects

Outcome	Estimate	SE	p	95% CI
Any Incident	0.003	0.003	.277	[-0.002, 0.008]
Any OSS	-0.002	0.002	.294	[-0.005, 0.001]
Any ISS	0.000	0.001	.940	[-0.001, 0.001]
Attendance Rate	0.013	0.002	<.001	[0.010, 0.017]
Incident Count (log)	-0.047	0.026	.066	[-0.097, 0.003]
ISS Days (log)	0.000	0.001	.520	[-0.002, 0.001]
OSS Days (log)	-0.001	0.002	.511	[-0.006, 0.003]

Note. Estimates represent pooled first-year intent-to-treat effects across cohorts.

Table A4. Fully Pooled Implementation-Period Effects

Outcome	Estimate	SE	p	95% CI
Any Incident	-0.024	0.002	<.001	[-0.028, -0.021]
Any OSS	0.000	0.001	.706	[-0.002, 0.002]
Any ISS	0.000	0.000	.931	[-0.001, 0.001]
Attendance Rate	0.002	0.001	.068	[0.000, 0.004]
Incident Count (log)	0.016	0.016	.312	[-0.015, 0.047]
ISS Days (log)	0.000	0.000	.490	[-0.001, 0.000]
OSS Days (log)	0.000	0.001	.881	[-0.003, 0.003]

Note. Estimates represent average treatment effects across all implementation years.

Table A5. Exposure-Year Difference-in-Differences Estimates

Outcome	Cohort 1, Year 1	Cohort 1, Year 2	Cohort 2, Year 1
Any Incident	−0.016** (0.003)	0.014** (0.003)	−0.047** (0.002)
Any ISS	−0.001 (0.001)	−0.0013† (0.0007)	0.000 (0.001)
Any OSS	0.000 (0.001)	0.000 (0.001)	0.000 (0.001)
Attendance Rate	0.002 (0.001)	0.001 (0.001)	0.001 (0.001)
Incident Count (log)	−0.059* (0.027)	0.177** (0.026)	0.000 (0.029)
ISS Days (log)	−0.00124* (0.00060)	−0.00153* (0.00060)	0.000 (0.001)
OSS Days (log)	0.000 (0.002)	0.005* (0.002)	0.000 (0.002)

*Note. Estimates represent staggered difference-in-differences coefficients from weighted linear mixed-effects models with school random intercepts and student-level covariate adjustment. Standard errors in parentheses (†p < .10; *p < .05; **p < .001).*

Table A6. Moderator Effects by Baseline Disciplinary Risk

Outcome	Exposure Year	Low-Risk Students	High-Risk Students
Any Incident	C1 Year 1	−0.023** (0.002)	−0.160** (0.015)
Any Incident	C1 Year 2	0.012** (0.002)	−0.130** (0.007)
Any Incident	C2 Year 1	−0.003 (0.002)	−0.353** (0.009)
Incident Count (log)	C1 Year 1	−0.066* (0.029)	−0.086* (0.035)
Incident Count (log)	C1 Year 2	0.104** (0.027)	0.181** (0.036)
ISS Days (log)	C1 Year 1	−0.001 (0.001)	−0.053** (0.006)
ISS Days (log)	C1 Year 2	−0.001 (0.001)	−0.065** (0.007)

Note. Baseline risk is defined as any disciplinary incident in the pre-intervention period.

*Standard errors in parentheses (*p < .05; **p < .001).*

Table A7. Effect Sizes for Statistically Significant Program Effects

Outcome	Group or Model	Exposure Period	Coefficient (SE)	Effect Size
Any Incident				
	Main model	Cohort 1 Year 1	-0.016 (0.003)	-1.6 percentage points
	Main model	Cohort 1 Year 2	0.014 (0.003)	+1.4 percentage points
	Main model	Cohort 2 Year 1	-0.047 (0.002)	-4.7 percentage points
Incident Count (log)				
	Main model	Cohort 1 Year 1	-0.059 (0.027)	-5.7%
	Main model	Cohort 1 Year 2	0.177 (0.026)	+19.4%
ISS Days (log)				
	Main model	Cohort 1 Year 1	-0.0012 (0.0006)	-0.12%
	Main model	Cohort 1 Year 2	-0.0015 (0.0006)	-0.15%
ISS Days (log)				
	Main model	Cohort 1 Year 2	0.0051 (0.0023)	+0.5%
Any Incident				
	High-risk students	Cohort 1 Year 1	-0.160 (0.015)	-16.0 percentage points
	High-risk students	Cohort 2 Year 1	-0.353 (0.009)	-35.3 percentage points
	Low-risk students	Cohort 1 Year 1	-0.023 (0.002)	-2.3 percentage points
	Low-risk students	Cohort 1 Year 2	-0.020 (0.002)	-2.0 percentage points
Incident Count (log)				
	High-risk students	Cohort 1 Year 1	-0.086 (0.035)	-8.3%
	High-risk students	Cohort 1 Year 2	0.181 (0.036)	+19.8%
	Low-risk students	Cohort 1 Year 2	0.311 (0.136)	+36.5%
	Low-risk students	Cohort 2 Year 1	-0.066 (0.029)	-6.3%
ISS Days (log)				
	High-risk students	Cohort 1 Year 1	-0.053 (0.006)	-5.2%
	High-risk students	Cohort 1 Year 2	-0.065 (0.007)	-6.3%

Outcome	Group or Model	Exposure Period	Coefficient (SE)	Effect Size
Any Incident				
	Hispanic students	Cohort 1 Year 1	-0.013 (0.004)	-1.3 percentage points
	Hispanic students	Cohort 1 Year 2	0.009 (0.004)	+0.9 percentage points
	Hispanic students	Cohort 2 Year 1	-0.059 (0.003)	-5.9 percentage points
	White students	Cohort 2 Year 1	-0.014 (0.006)	-1.4 percentage points
	Other students	Cohort 1 Year 1	-0.021 (0.005)	-2.1 percentage points
	Other students	Cohort 1 Year 2	0.011 (0.005)	+1.1 percentage points
	Black students	Cohort 1 Year 2	0.050 (0.014)	+5.0 percentage points
Incident Count (log)				
	White students	Cohort 1 Year 2	0.225 (0.057)	+25.2%
	Hispanic students	Cohort 1 Year 2	0.128 (0.034)	+13.7%
	Other students	Cohort 1 Year 2	0.278 (0.074)	+32.0%
	Other students	Cohort 2 Year 1	-0.114 (0.048)	-10.8%
ISS Days (log)				
	Hispanic students	Cohort 1 Year 1	-0.0021 (0.0008)	-0.21%
	Hispanic students	Cohort 1 Year 2	-0.0046 (0.0008)	-0.46%
	White students	Cohort 1 Year 2	0.0032 (0.0013)	+0.32%

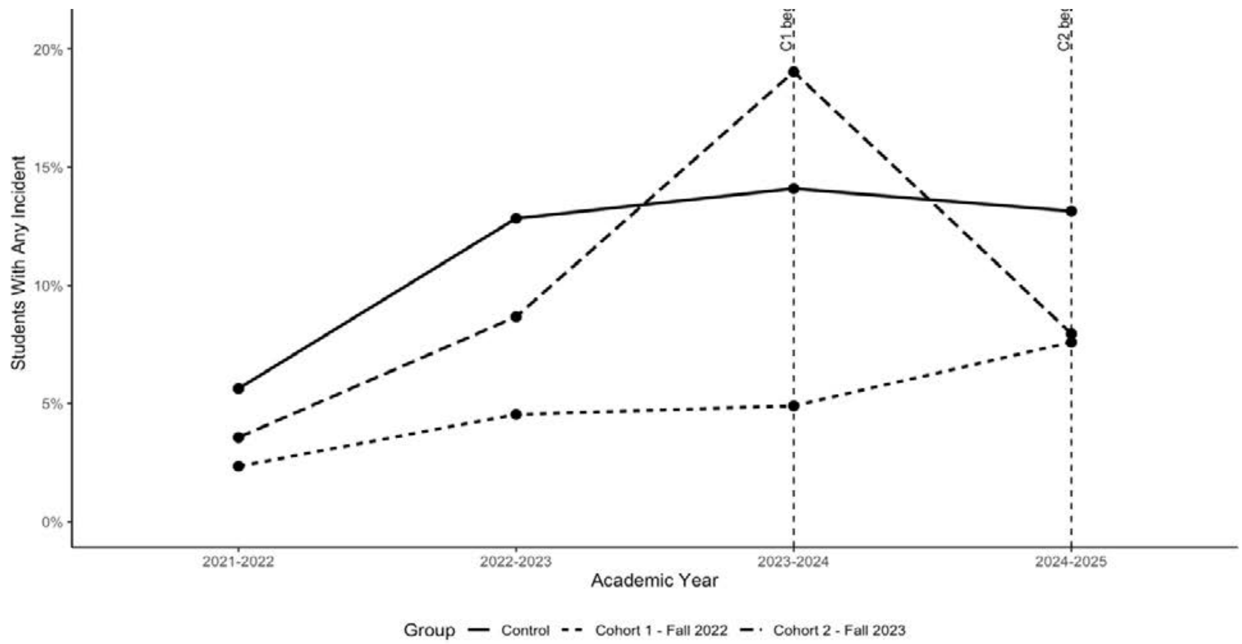
Note. For binary outcomes and attendance, effect sizes are reported as percentage-point changes (100 × ß). For log-transformed count outcomes, effect sizes are reported as approximate percent changes (100 × [exp(ß) − 1]). Only statistically significant coefficients are shown.

Table A8. Exploratory Exposure-Year Effects by Baseline Race and Ethnicity Group

Outcome	Group	Cohort 1, Year 1	Cohort 1, Year 2	Cohort 2, Year 1
Any Incident	White	−0.007 (0.005)	0.005 (0.005)	−0.014* (0.006)
Any Incident	Black	−0.017 (0.014)	0.050*** (0.014)	−0.007 (0.011)
Any Incident	Hispanic	−0.013*** (0.004)	0.009* (0.004)	−0.059*** (0.003)
Any Incident	Other	−0.021*** (0.005)	0.011* (0.005)	0.004 (0.005)
ISS Days (log)	Hispanic	−0.002** (0.001)	−0.005*** (0.001)	0.001 (0.001)

$^{*}p < .05$; $^{**}p < .01$; $^{***}p < .001$

Figure A1. Trends in Probability of any Disciplinary Incident by Cohort



Suggested citation: Gage, N., Ulich, T. & Wong, L. (2026). *Effects of the California integrated supports project on student discipline and attendance: A staggered difference-in-differences study.* WestEd.