Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 26-1435

Can We Save Failing Schools? Evidence From Los Angeles

Antonia Vazquez
Brown University

Can providing funding and support to failing schools help them improve? This paper studies this question using a natural experiment based on a 2017 lawsuit settlement that allocated substantial resources and support to the lowest-performing schools in the Los Angeles Unified School District (LAUSD). Using a difference-in-differences design, I compare 50 secondary schools that received an increase of 13.5% on average in their annual budgets for three years, to nearby public, noncharter schools that received no settlement funding. The intervention mandated hiring of additional staff members and allocating funds for professional development, but allowed discretionary spending on initiatives for high-need students. I find that, in line with the intent of the settlement, schools hired more personnel, including instructional staff such as teachers and counselors and support personnel such as paraprofessionals and school service staff. Settlement schools achieved a 0.75 percentage-points reduction in suspension rate relative to unfunded schools, reflecting a marked improvement in disciplinary outcomes. These reductions were particularly notable given that the settlement triggered demographic sorting, with the treated schools losing students overall and shifting toward higher concentrations of economically disadvantaged students and lower Black enrollment. A simple bounding exercise that accounts for demographic sorting indicates that the settlement had meaningful effects on suspension rates, suggesting real improvements on noncognitive dimensions of schooling. Survey evidence suggests that two key mechanisms for lower suspensions were improvements to school climate as reported by staff and students, and enhancements to educators' capacity and disciplinary approaches.

VERSION: August 2026

Suggested citation: Vazquez, Antonia. (2026). Can We Save Failing Schools? Evidence From Los Angeles. (EdWorkingPaper: 26-1435). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/acg9-ks68

Can We Save Failing Schools? Evidence From Los Angeles*

Antonia Vazquez

August 15, 2026

Abstract

Can providing funding and support to failing schools help them improve? This paper studies this question using a natural experiment based on a 2017 lawsuit settlement that allocated substantial resources and support to the lowest-performing schools in the Los Angeles Unified School District (LAUSD). Using a difference-in-differences design, I compare 50 secondary schools that received an increase of 13.5% on average in their annual budgets for three years, to nearby public, noncharter schools that received no settlement funding. The intervention mandated hiring of additional staff members and allocating funds for professional development, but allowed discretionary spending on initiatives for high-need students. I find that, in line with the intent of the settlement, schools hired more personnel, including instructional staff such as teachers and counselors and support personnel such as paraprofessionals and school service staff. Settlement schools achieved a 0.75 percentage-points reduction in suspension rate relative to unfunded schools, reflecting a marked improvement in disciplinary outcomes. These reductions were particularly notable given that the settlement triggered demographic sorting, with the treated schools losing students overall and shifting toward higher concentrations of economically disadvantaged students and lower Black enrollment. A simple bounding exercise that accounts for demographic sorting indicates that the settlement had meaningful effects on suspension rates, suggesting real improvements on noncognitive dimensions of schooling. Survey evidence suggests that two key mechanisms for lower suspensions were improvements to school climate as reported by staff and students, and enhancements to educators' capacity and disciplinary approaches.

*Annenberg Institute, Brown University, USA; antonia_vazquez@brown.edu. I thank Eric Chyn, Christopher Campos, Scott Carrell, Mary Evans, and Raissa Fabregas for their substantial comments and feedback on this paper. I also thank Manuela Angelucci, Jesse Bruhn, Emileigh Harrison, Yumin Hong, Paula López Villalba, Richard Murphy, Francisco Pardo, Ljubica Ristovska, Andrew Waxman, and seminar participants at the LBJ School Sack Lunch Seminar and Universidad de los Andes, Colombia, for helpful comments and suggestions.

1 Introduction

Should we invest in low-performing schools? This question is the subject of longstanding debate (Abdulkadiroglu et al., 2026, Atchison and Blair, 2026, Bifulco and Schwegman, 2020, Blair et al., 2026, Han et al., 2017, Sunderman et al., 2017). Supporters contend that low-performing schools fail due to systematic underresourcing and that substantial, sustained funding can turn them around. Detractors argue that closures are a more effective option because it redirects scarce resources toward more effective alternatives.

Evidence on this question is scarce. While there is now a large body of work on district finances, there are no studies that specifically ask whether funding for under-resourced schools can improve student outcomes (Abott et al., 2020, Baron, 2022, Biasi et al., 2025, Candelaria and Shores, 2019, Hyman, 2017, Jackson, 2020, Jackson and Mackevicius, 2024, Jackson et al., 2021, 2016, Johnson and Jackson, 2019, Lafortune et al., 2018). This scarcity of evidence is likely due to the fact that within-district variation exhibits substantial scope for selection. Low-performing schools are selected not only on poverty, which they typically share with neighboring schools in the same district, but on under-performance despite facing similar funding and demographics, a gap more likely to reflect leadership instability, staff turnover, or worse disciplinary climate than resource scarcity alone. Whether the logic of between-district reforms extends to this within-district margin therefore remains an open question.

I answer this question using a unique natural experiment. I study a 2017 lawsuit settlement in the Los Angeles Unified School District (LAUSD) that delivered substantial resources to the district schools struggling the most. Specifically, 50 secondary schools that officials identified as the lowest-performing in the LAUSD received $1,030 per student annually for three years, a figure representing approximately 13.5% of their existing budgets on average. Rather than imposing top-down structural mandates, the settlement combined targeted resources with local autonomy. It required schools to establish tiered student support teams and allocate funds to professional development, while granting local administrators discretion to invest the remaining funds in initiatives tailored to their high-need students. Using a difference-in-differences (DiD) design, I compare outcomes between the 50 funded secondary schools and nearby public, noncharter secondary schools that did not receive funding. To interpret the results causally, this design requires that funded schools would have followed similar trajectories to nonfunded schools absent the intervention; an assumption for which I find support in parallel pretreatment trends.

I examine multiple outcomes using administrative school level data from the California Department of Education and LAUSD, including staffing levels, suspension rates, ELA and math proficiency rates, enrollment, and demographics. To explore potential mechanisms, I take advantage of school experience survey data that capture detailed perceptions of school climate and environment from both students and staff. These data allow me to assess not only the direct effects of the intervention but also the mechanisms driving observed changes.

I show that schools successfully implemented the stipulated changes by increasing their employment of certified staff such as teachers and counselors by approximately 3 per 1,000 students and of support staff such as office staff by almost 2 per 1,000 students; these figures represent increases of approximately 5% over baseline levels and are statistically significant at standard levels. These resource increases translated into substantial improvements in the disciplinary climate, with reductions in suspension rates of 0.75 percentage points (pp), a substantial 44% decline from pretreatment levels. This effect size is more than double the magnitude of common school climate interventions such as restorative justice programs (Adukia et al., 2025, Augustine et al., 2018). While the disciplinary improvements were substantial, academic outcomes showed more limited progress. The intervention improved English language arts (ELA) achievement by 2.8 pp (12.8% from baseline), and for mathematics, the point estimate is positive but nonsignificant. However, because of data limitations that restrict the observation of academic outcomes to only two pretreatment periods, these test score results should be interpreted as suggestive rather than strictly causal. In contrast, the robust and causally identified reductions in suspension rates, paired with marked improvements in school climate measured by school experience surveys, suggest that the primary impact of the targeted funding was to improve the noncognitive and environmental dimensions of the schooling experience.

A key empirical question—one to which the answers predicted by theory are ex ante ambiguous—is whether the settlement also affected enrollment and student composition. Better-resourced schools might attract students seeking improved opportunities, while litigation stigma could signal poor performance and deter enrollment. In practice, the treated schools lost 48 students on average (a 5.8% decline from baseline), experienced a 0.9 pp decrease in Black enrollment (a 6.3% decline), and saw free lunch eligibility increase by 3.6 pp (a 4.2% increase). These patterns suggest that families with better school choice options left the treated schools, such that economic disadvantage became more concentrated in them.

These shifts in student composition raise questions about the interpretation of my main findings, particularly the reduction in suspension rates. If higher-risk students left treated schools, the drop in suspensions might reflect student turnover rather than better school practices. To address this concern, I implement a bounding exercise in the spirit of Horowitz and Manski (2000) that assesses how sensitive my estimates are to different assumptions about the characteristics of students who left. Even under highly conservative scenarios where I assumed that departing students would have suspension propensities twice as high as the school baseline, the estimated treatment effect remains negative and statistically significant. This exercise suggests that the observed improvements in the disciplinary climate are not a mechanical result of demographic sorting but instead reflect genuine, positive changes in school practices and educator approaches.

To understand whether the observed reduction in suspension rates reflects behavioral improvements, shifts in educator responses, changes in administrative practices, or better school climate, I examine school experience survey data from students and staff. My findings suggest that students in the treated schools perceived meaningful improvements in overall school climate (a 0.15 standard deviation (SD), a statistically significant effect at the 5% level) and better experiences with bullying (although this effect is not statistically significant). More strikingly, teachers and staff reported substantial improvements across multiple dimensions of their work environment. They reported receiving professional development on bullying and addressing behavioral issues when they arise (0.38 SD), and that bullying incidents had become a less significant problem (0.26 SD). Additionally, staff reported greater agreement that disciplinary practices at their schools were both fair and effective (0.31 SD), alongside notable improvements in their perceptions of the overall school climate (0.37 SD). These improvements were statistically significant at the 1% level. These patterns suggest that the intervention operated by improving school climate as reported by both staff and students and by enhancing educators' capacity and approaches to discipline.

Together, these results show that targeted, autonomy-preserving investment produced large improvements in school climate and discipline, even as academic outcomes were more mixed and, given data limitations, only suggestive rather than causally conclusive. At the same time, the intervention worsened the enrollment decline and compositional sorting typically associated with schools on a path toward closure, as families with more outside options left the treated schools. This pattern, however, may reflect the public, litigation-driven identification of these schools as the district's lowest-performing rather than the effect of additional resources per se: a quieter, less stigmatizing form of investment might not trigger the same sorting response. Money and resources paired with local autonomy can address some of the institutional constraints attached to failing schools, particularly around climate and discipline, but whether it can do so without also triggering the family-sorting dynamics likely depends on how visibly the intervention singles schools out.

This research contributes to several strands of literature. First, as discussed above, it speaks to an open question in the school finance literature: whether the logic of between-district reforms, where effects are concentrated in high-poverty, low-spending districts (Abott et al., 2020, Baron, 2022, Biasi et al., 2025, Candelaria and Shores, 2019, Hyman, 2017, Jackson, 2020, Jackson and Mackevicius, 2024, Jackson et al., 2021, 2016, Johnson and Jackson, 2019, Lafortune et al., 2018), extends to within-district resource allocation, where the selection into low-performing is different in kind. This paper provides direct evidence on that margin by studying a resources-and-support intervention targeted specifically at low-performing schools within a school district.

Second, existing work has primarily emphasized cognitive outcomes while overlooking other important dimensions of school quality, such as school climate, teacher quality, and student support services, that are strongly associated with overall school effectiveness (Bryk et al., 2010, Carrell et al., 2025, Jackson et al., 2020, Porter et al., 2023). Recent work also shows that families value school climate information when it is made salient and accessible (Crespin, 2026). This paper expands the analysis beyond the cognitive outcomes that dominate the school finance literature to include school climate and disciplinary outcomes. These noncognitive dimensions may be especially critical in low-performing schools, where restoring safety and a supportive learning environment represents a necessary first step toward long-term student success and a vital pathway to disrupting the school-to-prison pipeline.

The closest prior evidence on interventions to improve low-performing schools comes from research on School Improvement Grants (SIG), under which chronically underperforming schools were required to implement turnaround models designed by the federal government. These programs bundled money with dramatic structural changes such as replacement of principals and at least 50% of the schools' staff. The national evaluation of SIG found no significant impact on math or reading achievement, high school graduation, or college enrollment (Dragoset et al., 2017). The key distinction is that SIG represented a resource-plus-structural-change approach combined with strong accountability pressure, whereas the Los Angeles settlement provided resources and support only, functioning more as a "carrot" than a "stick." By preserving school autonomy and allowing local actors to direct funding toward their specific challenges, this approach contrasts with the punishment-based mechanisms that have dominated accountability policy (Booher-Jennings, 2005, Cullen and Reback, 2006, Figlio, 2006, Figlio and Rouse, 2006, Figlio and Getzler, 2006, Jacob, 2005, Jacob and Levitt, 2003, Reback, 2008), pressure that, while sometimes effective, tends to produce unintended consequences such as teaching to the test, strategic student exclusion, and gaming behaviors. Beyond effectiveness, resource-only approaches are more politically palatable than imposed turnarounds or school closures, which face significant resistance and are difficult to adopt at scale. My analysis thus tests whether a less disruptive, more supportive policy design can achieve meaningful improvements in the schools that need it most.

2 Background on the Los Angeles Unified School District (LAUSD)

The LAUSD is the second-largest public school district in the United States, serving over 500,000 students, and includes most of the city of Los Angeles, along with all or portions of 25 cities and unincorporated areas of Los Angeles County (Los Angeles Unified School District, 2025). The district operates approximately 1,300 schools, including elementary, middle, high, magnet, and special-education campuses. Because of its scale and heterogeneity, LAUSD faces significant challenges in resource allocation, equity, and accountability across schools with varying student needs.

2.1 School funding

In California, public schools are funded through a combination of state, local, and federal sources. The state provides the majority of funding (59%) through general revenues, while property taxes and other local sources contribute 32%, and the federal government provides less than 10%, primarily for specific programs.[1] California allocates these funds through the Local Control Funding Formula (LCFF), which determines how much each district receives based on enrollment and student needs.

For LAUSD, this means that the student funding takes into account the district's high concentration of disadvantaged populations. In the academic year 2023–2024, approximately 22% of students were designated as English language learners (ELLs), 81% were considered socioeconomically disadvantaged, and 2.7% were foster or homeless youth.[2]

California adopted the Local Control Funding Formula (LCFF) in 2013, which eliminated most categorical funding programs, replacing them with a simpler and more flexible funding system. Under the new formula, Local Education Agencies receive funding based on the demographic profile of the students they serve. School districts receive extra resources tied to high-need students (i.e., those classified as low-income, English language learners, and/or foster youth), with districts

[1] Public Policy Institute of California, "Targeted K-12 Funding and Student Outcomes," available online.

[2] Ed-Data, "Los Angeles Unified School District," available online.

receiving *supplemental* grants when these students are present, and *concentration* grants when they make up more than 55% of enrollment.[3] The goals of the LCFF were to distribute money equitably by providing more funding to high-need districts, to provide greater flexibility to local school boards, and to simplify the funding system.[4]

2.2 Emergence of the Community Coalition Lawsuit

Despite LCFF's goals, in July 2015, the Community Coalition of South Los Angeles (a local advocacy organization), along with parent plaintiff Reyna Frias, filed suit against LAUSD. They alleged that the district had misallocated approximately $450 million in funds designated for high-need students.[5] The plaintiffs argued that LAUSD counted prior special education expenditures, which the district was already required by law to provide, as if they had satisfied LCFF-mandated supports for low–income, English language learner, and foster youth populations. This accounting practice effectively reduced the additional resources that should have flowed to high-need students.

The Community Coalition litigation was the first to challenge a California district's implementation of LCFF under these terms. In a parallel development, a 2016 administrative decision by the California Department of Education sided with the plaintiffs, finding that LAUSD's accounting practices violated state requirements. Following protracted negotiations, LAUSD ultimately reached a settlement in July 2017, with formal ratification in September 2017 (see a detailed timeline in Figure 1).

2.3 The Settlement as Program "Treatment"

Under the settlement, LAUSD committed to reallocating $171.6 million over three years (2017–18 to 2019–20) to "school innovation funds" specifically earmarked for high-need student services (Los Angeles Unified School District, 2017c). The agreement identified 50 underperforming secondary schools (20 middle schools, 30 high schools) as priority recipients, selected via objective criteria including unduplicated student percentages, mathematics achievement, suspension rates, foster youth, and homelessness counts (King, 2017, Los Angeles Unified School District, 2017a,c).

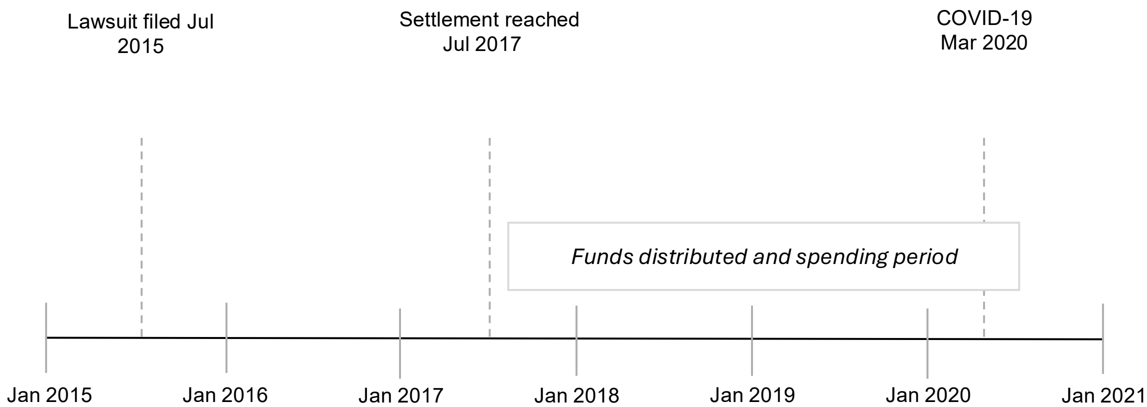
Funding was allocated proportionally to each school's duplicated count of high-need students, at approximately $1,030 per duplicated pupil per year (Los Angeles Unified School District, 2017b),

[3] Public Policy Institute of California, "Targeted K-12 Funding and Student Outcomes," available online.

[4] Brookings Institution, "Lessons Learned from 10 Years of California's Local Control Funding Formula," available online.

[5] ACLU of Southern California, "Community Coalition v. Los Angeles Unified School District," available online.

Figure 1: *Community Coalition v. LAUSD Timeline*



Notes: This figure illustrates the timeline of the LAUSD–ACLU settlement and its implementation. In July 2015, the American Civil Liberties Union (ACLU) and other advocacy organizations filed a lawsuit against the Los Angeles Unified School District (LAUSD) alleging improper use of hundreds of millions of dollars allocated under the Local Control Funding Formula (LCFF). The lawsuit reached settlement in July 2017, with LAUSD agreeing to allocate $171.6 million specifically for services for high-need students. The settlement funds were distributed over three academic years during the spending period of 2017/18–2019/20. Following the conclusion of the mandated spending period in 2020, the district retained discretionary authority over any remaining unspent funds. The figure shows the key phases of litigation, settlement negotiation, fund distribution, and postsettlement period that form the basis for the empirical analysis in this study.

corresponding to an average annual budget increase of 13.5%. The settlement's duplicated count methodology allows students who met multiple criteria to be counted more than once.

The settlement imposed two mandatory implementation constraints:

1. Each settlement school had to establish tiered student support teams (called "Achievement Through Support" Teams), scaling service intensity with school size. These teams were required to include restorative justice advisors, counselors, and psychiatric social workers (for larger schools).

2. Schools were required to devote at least 10% of the remaining funds toward foundational professional development in mathematics or English language arts.

Outside those requirements, schools retained autonomy to spend across six broad categories targeting high-need student services (see Appendix B for specifics).

From a program-evaluation standpoint, I treat the settlement's introduction of these funds and mandates as the "treatment." Settlement schools become the treated group; the comparison group are unfunded similar district schools. The timing, criteria-based selection, and funding structure afford a quasi-experimental opportunity to estimate the effect of injecting targeted resources plus constrained accountability into high-need schools.

3 Data

My analysis draws on publicly available administrative data at the school level from the California Department of Education (CDE) and the Los Angeles Unified School District (LAUSD), combined with information from the settlement agreement.

3.1 Sample construction

The analysis sample includes noncharter LAUSD secondary schools that were operational in academic year 2016/17, right before the settlement was reached (202 schools). I impose several restrictions to improve comparability between treatment and control schools. First, I exclude continuation and special education schools (7 schools), as these serve fundamentally different student populations with distinct educational missions. Second, I retain only secondary schools in the same local districts as treated schools (21 schools excluded) to ensure geographic comparability and similar neighborhood characteristics. LAUSD is subdivided into eight local districts, and I exclude schools in the North West local district because none of the treated schools are located there. This area includes some of the wealthiest neighborhoods in Los Angeles, such as Sherman Oaks, Encino, and Woodland Hills. Third, I restrict the sample to schools operating continuously throughout 2014/15 to 2019/20 (5 schools excluded). I set this end date to avoid potential confounding from the COVID-19 pandemic and this start date to ensure all schools have pretreatment data, as two of the beneficiary schools opened in August 2014.

The selection criteria yield an initial sample of 171 schools, consisting of 121 control schools and 50 treatment schools observed over 6 academic years from 2014/15 to 2019/20. However, because one treated school merged with another institution following the settlement, the final sample consists of 49 treated schools and 121 control schools, totaling 170 schools.[6] Figure A.1 shows the geographical distribution of the sample schools.

[6] Results are robust to dropping this school and the two beneficiary schools that opened in August 2014; estimates remain virtually unchanged in sign, magnitude, and statistical significance across all outcomes.

3.2 Outcome data

I examine multiple school-level outcomes to assess the implementation and effects of the settlement. All outcome data are publicly available and come aggregated at the school level. First, I analyze staffing outcomes to understand how schools allocated the additional funds. Second, I examine student disciplinary and academic outcomes. Third, I use student and staff responses to a school experience survey to measure perceptions of school climate and environment.

3.2.1 Staffing outcomes

To measure resource allocation, I examine two categories of school-level employment data.[7] Both measures are expressed as full-time equivalent (FTE) positions per 1,000 students enrolled.

Certified employees. These are school staff required to hold valid teaching credentials or permits and include (1) teachers who provide direct instruction, (2) administrative personnel in positions requiring certification but not providing direct instruction, and (3) pupil services staff such as counselors, librarians, and psychologists who hold specialized credentials. These data from the CDE are available through academic year 2018/19.

Classified staff. These are employees in positions not requiring certification and include (1) paraprofessionals who provide instructional support under teacher supervision (teaching assistants, aides, translators), (2) office/clerical staff supporting administration and business services, and (3) other support staff including health services, maintenance, food service, security, and transportation personnel. These data from the CDE are available through academic year 2019/20.

3.2.2 Student outcomes

Given that the settlement mandated significant investments in services for high-need students, including academic support and mental health services, and in school climate initiatives such as restorative justice programs, I focus on outcome measures that directly reflect these intervention areas. Both outcomes are measured at the school level using CDE administrative data.

Suspensions. I measure annual suspension rates as the unduplicated count of students suspended at least once during the academic year divided by cumulative enrollment. I use school-level data from 2014/15 to 2019/20.

[7] While actual school-level expenditure records for the fiscal years 2019 and 2020 are unavailable from public sources, school-level staffing data provide a direct and reliable measure of real resource allocation. Because the settlement specifically mandated the hiring of additional personnel, tracking changes in full-time equivalent positions per student captures the primary mechanism through which the funding was operationalized at the school site.

Test scores. The CDE publicly provides school-level performance data only in categorical levels: the percentage of students whose performance exceeded standards, met standards, nearly met standards, or did not meet standards. I focus on the percentage of students who met standards as my measure of academic achievement.[8] I use data from 2015/16 to 2018/19 due to data availability.

3.2.3 School experience survey outcomes

To measure experiences and perceptions of school climate, I use data from LAUSD's School Experience Survey (SES), an annual survey that captures comprehensive feedback from teachers, staff, students, and parents across all LAUSD schools. I examine aggregated school-level survey responses for the 2014/15–2019/20 academic years.

For students, I focus on measures of bullying experiences and overall perceptions of school climate. The survey has high response rates of 80% on average across years in my analysis sample. The bullying indicators assess four key forms of peer victimization occurring on school property: physical aggression (being pushed, shoved, slapped, hit, or kicked), relational aggression (being subject to the spreading of mean rumors or lies), sexual harassment (experiencing sexual jokes, comments, or gestures), and appearance-based teasing. The school climate measures capture students' perceptions across six dimensions of their school environment, including adult respect and help, feelings of safety and belonging, and overall satisfaction with their school experience.

The staff questionnaire surveys principals, administrators, teachers, counselors, and other school staff about their experiences, perceptions of school climate, and preparedness to address climate issues. The staff survey also has a high response rate of 82% on average across years in my analysis sample. The survey measures several key dimensions of school climate, including professional development preparedness (specifically regarding antibullying training), engagement in addressing bullying incidents, and personal safety perceptions during school hours. Staff members also evaluate the school's disciplinary climate through their perceptions of adult respect toward students and the fairness and effectiveness of the school's discipline. Additionally, staff assess the prevalence and severity of various climate challenges within their schools, including harassment and bullying among students, physical fighting, general disruptive behavior, racial or ethnic conflicts, and student respect toward staff members. See C for details on survey questions.

To analyze these multidimensional survey outcomes systematically, I construct summary in-

[8] This metric is particularly relevant for the treated schools in my sample, where approximately 45% of students fail to meet ELA standards and 70% fail to meet math standards, suggesting that interventions in these schools should primarily aim to move students from the lowest performance categories into at least approaching proficiency.

dices following the methodology outlined in Anderson (2008). This approach uses generalized least squares (GLS) weighting to aggregate standardized outcomes within each domain, where the weight assigned to each component outcome equals the sum of its corresponding row in the inverted covariance matrix. This weighting scheme optimally accounts for the correlation structure among outcomes, assigning lower weight to highly correlated measures and greater weight to outcomes that provide independent information. As an example, to create the student bullying index, I combine responses across the four victimization categories (physical aggression, relational aggression, sexual harassment, and appearance-based teasing), taking the proportions of students reporting each type of incident as "not a problem" and creating a summary measure using the Anderson (2008) procedure. Similarly, I construct indices for school climate, staff preparedness for handling bullying, perceptions of disciplinary effectiveness, and overall staff assessments of school climate. This indexing approach provides clear, interpretable measures of the treatment effects across conceptually related outcomes. Prior to standardization, I recoded all outcomes such that the positive direction uniformly indicates improved performance, ensuring that positive summary index values correspond to positive effects across all domains. Each index is then standardized to have mean zero and unit variance using the full sample mean and standard deviation across all schools and years. See Appendix C for a list of the questions I used to construct each index.

3.3 Summary statistics

Before comparing beneficiary and control schools in detail, Table A.1 situates the study sample within the broader California context in the academic year 2016/17 using comparable data from the California Department of Education. *CA* refers to public schools outside Los Angeles County, *LA* to schools in the county outside LAUSD, and *LAUSD* to district schools not in the study sample; the last two columns correspond to the comparison and beneficiary groups. LAUSD schools as a whole serve more disadvantaged students than the rest of the state and county: even the non-sample LAUSD schools show a free lunch rate of 80.6% and an English Learner share of 15.4%, compared to roughly 52–58% and 13–14% elsewhere. Within this already high-need district, beneficiary schools are the most disadvantaged group across nearly all dimensions. However, for most indicators, particularly free lunch (87.4% versus 83.6%) and homeless students (4.2% versus 3.0%), the gap between beneficiary and comparison schools is modest relative to the gap separating both groups from the rest of California. The main exception is English Learners, where beneficiary schools (26.5%) have a higher share than the comparison group (15.1%), a difference comparable in magnitude to the gap between LAUSD and the rest of the state.

Table 1 compares pretreatment characteristics (measured in 2016/17) between the 49 beneficiary schools and the 121 control schools. Beneficiary schools are smaller on average (834 versus 1,187 students) and serve student populations with higher concentrations of historically underserved groups. The share of students eligible for free lunch is 3.8 percentage points higher in beneficiary schools (87.4% versus 83.6%). Black students comprise 14.4% of enrollment in beneficiary schools compared to 7.5% in controls, nearly double the share. Conversely, beneficiary schools enroll fewer White students (1.2% versus 5.0%) and fewer Asian students (0.9% versus 3.2%). Beneficiary schools also serve substantially more English Learners (26.4% versus 15.1%, an 11.3 percentage point difference), more homeless students (4.2% versus 3.0%), and more than twice the share of foster youth (1.3% versus 0.6%).

In terms of academic outcomes, treated schools show substantially worse baseline achievement on state standardized tests. The share of students not meeting standards is 18.7 percentage points higher in ELA (44.4% versus 25.7%) and 22.7 percentage points higher in mathematics (69.6% versus 46.9%).

Beneficiary schools face greater disciplinary challenges and worse perceived climate. Suspension rates are nearly three times higher (1.70% versus 0.64%). Survey indices, standardized to have mean zero and unit variance across the full sample, reveal substantial gaps in perceived climate. Staff in beneficiary schools rate school climate 1.04 standard deviations lower than staff in control schools (−0.57 versus 0.47), and perceive discipline as less fair and effective (a gap of 0.88 standard deviations). Students in beneficiary schools also reported worse school climate (−0.73 versus −0.21, a difference of 0.52 standard deviations).

Figure A.2 shows the distribution of the lawsuit index separately for beneficiary and comparison schools in 2015/16. The index is the composite score stipulated by the settlement to identify eligible schools – a weighted sum that assigns positive weights to the share of homeless students (5%), foster students (5%), the overall suspension rate (20%), and the share of unduplicated pupils (30%), and a negative weight to the share of students meeting or exceeding the math standard (40%), so that lower academic performance raises the index. The settlement does not publicly report the school-level scores; the index used here is my replication and should be understood as an approximation of the selection rule. Figure A.2 shows that beneficiary schools occupy the worse end of the distribution: most score above 20, reflecting high concentrations of disadvantaged students and poor academic outcomes. Comparison schools are more spread out, with most falling below the threshold, though a subset score in the 20–25 range.

4 Empirical Framework

4.1 Estimation strategy

The central challenge in estimating the effect of the settlement funding is that beneficiary schools were not randomly selected, they were chosen because they were struggling. As Table 1 documented, these schools differ from other LAUSD secondary schools: they serve higher shares of low-income students, English Learners, and foster youth; they have substantially lower baseline test scores; and they exhibit worse school climate. Simply comparing outcomes between beneficiary and non-beneficiary schools after the settlement would conflate the effect of the funding with these preexisting differences.

To address this challenge, I use a difference-in-differences (DID) design where I compare *changes* in outcomes before and after the settlement between the two groups of schools. The baseline specification is:

$$Y_{st} = \alpha_s + \lambda_t + \delta \cdot (\text{Settlement}_s \times \text{Post}_t) + \varepsilon_{st} \tag{1}$$

where Y_{st} is the outcome variable, e.g., the suspension rate, for school s in year t. The α_s denotes school fixed effects, λ_t denotes year fixed effects, Settlement_s is an indicator equal to 1 for schools that received the settlement funds, Post_t is an indicator equal to 1 for years 2017/18 and after (postsettlement), $\text{Settlement}_s \times \text{Post}_t$ is the interaction term capturing the treatment effect, δ is the DID estimator, and ε_{st} is the error term. For statistical inference, I cluster the standard errors at the school level. This approach accounts for the serial correlation of outcomes within schools, since, for example, a school that performs well in one year tends to perform well in subsequent years, while maintaining the assumption that errors are independent across schools. School-level clustering is appropriate given that the "treatment assignment" is at the school level and schools represent the primary unit of analysis (Abadie et al., 2023). As a complement to the binary treatment specification, I also estimate models using a continuous measure of treatment intensity based on each school's initial settlement allocation as a share of its 2016/17 budget.

The school fixed effects (α_s) control for time-invariant unobserved differences between treated and control schools. Given the baseline differences documented in Table 1, these fixed effects

Table 1: Summary Statistics of Schools in Analysis Sample

	Lawsuit beneficiary				Comparison group			
	Mean	SD	Min	Max	Mean	SD	Min	Max
Demographics								
N of students enrolled	833.59	407.55	288	1961	1187.20	682.70	189	2953
% Hispanic	82.09	16.16	28	99	80.42	19.17	20	99
% Black	14.42	16.39	0	69	7.51	12.19	0	71
% White	1.21	0.89	0	5	4.98	7.32	0	35
% Asian	0.89	2.18	0	14	3.20	5.16	0	26
% less than 2% races	1.39	1.39	0	5	3.90	6.06	0	45
% free lunch	87.41	5.42	74	99	83.61	7.85	53	97
% english learner	26.35	6.93	9	42	15.12	8.49	0	41
% Homeless	4.21	1.88	1	10	3.03	1.98	0	12
% Foster	1.29	1.02	0	4	0.56	0.40	0	2
Employees								
Certified staff per 1,000 st	63.56	11.90	48	99	54.81	6.47	44	77
Classified staff per 1,000 st	37.80	14.54	16	83	27.58	9.65	6	61
Main outcomes								
Suspension rate	1.70	2.11	0	12	0.64	1.11	0	9
% met math standard	7.47	4.13	1	19	16.78	5.68	4	39
% met ELA standard	21.90	8.42	5	41	32.48	7.61	18	50
% not meet ELA standard	44.37	16.40	8	78	25.65	13.03	0	55
% not meet math standard	69.56	9.94	45	91	46.88	12.40	2	71
Survey indexes outcomes: students								
Bullying experiences	-0.51	1.11	-2.71	1.42	-0.42	1.04	-2.85	2.00
School climate	-0.73	0.68	-2.01	1.02	-0.21	0.94	-1.88	3.95
Survey indexes outcomes: staff								
Handling bullying	-0.35	1.06	-1.96	3.09	0.43	1.02	-1.86	3.34
Bullying among students	-0.49	1.12	-1.97	2.02	0.36	0.98	-1.42	2.91
Discipline	-0.35	0.82	-1.65	2.70	0.53	1.11	-1.41	3.79
School climate	-0.57	0.84	-1.91	1.47	0.47	1.03	-1.41	3.70
N Schools	49				121			

Notes: Summary statistics correspond to academic year 2016/17. The comparison group includes all public, noncharter secondary schools in Los Angeles Unified School District outside of the North West local district. I construct survey indices following Anderson (2008) and normalize them to have mean zero and standard deviation one using the full sample across all years. On all indices, a higher value means a better outcome. For students, the bullying index combines four survey questions about students' experiences with physical aggression (being pushed, shoved, slapped, hit, or kicked), relational aggression (being the subject of mean rumors or lies), sexual harassment (experiencing sexual jokes, comments, or gestures), and appearance-based teasing. The school climate index includes six survey questions on adult respect and help, feelings of safety and belonging at the school, and overall satisfaction with their school experience. For staff, the handling bullying index contains two survey questions related to receiving professional development on preventing bullying and whether they address bullying. The bullying index includes three questions on how much of a problem staff thinks bullying, physical fighting, and discrimination are among students. The discipline index includes two questions about the fairness and effectiveness of discipline at school. The school climate index has four questions on safety, respect, and disruptive behavior.

ensure that the estimated treatment effect is not confounded by the fact that beneficiary schools serve more disadvantaged populations, have fewer resources, or differ in other permanent ways from comparison schools. The year fixed effects λ_t, absorb district-wide shocks that affect all schools equally in a given year–such as changes in state testing standards, district-wide policy shifts, or macroeconomic conditions. The parameter of primary interest is δ, which measures the average treatment effect of the settlement funding on school outcomes in the postsettlement period. It captures the average difference in outcome trajectories between beneficiary and comparison schools after the settlement, relative to their difference before. The main specifications are unweighted, treating each school equally and estimating the average treatment effect across schools. As a robustness check, I also re-estimate them weighting observations by total enrollment, so that larger schools receive proportionally more weight and the estimand corresponds to the average treatment effect across students rather than schools.

4.2 Identification strategy

The validity of this DID approach relies on the parallel trends assumption. In this case, this means that in the absence of the settlement, outcomes at beneficiary schools would have evolved similarly to those at comparison schools. The plausibility of this assumption rests not only on observable pre-trends but also on the selection mechanism itself. The 50 beneficiary schools were chosen by the plaintiff organization (Community Coalition of South Los Angeles) and LAUSD, based on a ranking of schools using baseline measures of need: foster youth and homeless counts, targeted student population enrollment, suspension rates, and math achievement.[9]

Selection into the treatment group was determined by baseline levels of student need, academic achievement, and disciplinary incidents, rather than prior trajectories. While the exogenous nature of this selection process rules out strategic anticipation by school administrators, selecting schools based on extreme baseline levels introduces a potential threat of mean reversion. If the baseline performance of beneficiary schools reflected transitory shocks, these schools might naturally revert toward the district mean in subsequent periods, violating the parallel trends assumption. To address this concern, the empirical analysis relies on three complementary approaches. First, the

[9] Given that selection was based on a ranking of schools, a Regression Discontinuity Design (RDD) might appear to be the natural empirical approach. However, an RDD is infeasible in this setting for two primary reasons. First, the exact composite index scores used by the district to finalize the list of 50 schools are not publicly reported. Second, as demonstrated by the replicated index in Figure A.2, there is no strict cutoff; several comparison schools exhibit index scores that overlap with the lower bound of the beneficiary schools. Consequently, given the absence of a sharply observed running variable and a strict threshold a DID design is a better fit.

event study specification visually check for pre-existing differential trends. Second, I evaluate the main outcomes using log-transformations to ensure that the estimated proportional changes are not mechanically driven by baseline levels. Third, the analysis incorporates a synthetic difference-in-differences estimator that reweights the control group to match the pretreatment trajectories of the treated schools, thereby mitigating bias from baseline compositional differences.[10] [11]

To provide evidence supporting the parallel trends assumption, I use an event study specification that allows for dynamic treatment effects and a visual inspection of pre-treatment trends:

$$Y_{st} = \alpha_s + \lambda_t + \sum_{k=-3, k\neq -1}^{2} \delta_k \cdot (\text{Settlement}_s \times \mathbf{1}_{t=2018+k}) + \varepsilon_{st} \tag{2}$$

where Y_{st} is the outcome variable, e.g., the suspension rate, for school s in year t, α_s denotes school fixed effects, λ_t denotes year fixed effects, Settlement_s is an indicator equal to 1 for schools that received the settlement funds, $\mathbf{1}_{t=2018+k}$ represents indicators for years relative to the 2017/18 settlement (where $k = -1$ is academic year 2016/17 and omitted as the reference), δ_k is the treatment effect coefficients for each of the years relative to settlement, and ε_{st} is the error term, clustered at the school level.

The event study specification allows me to examine the parallel trends assumption by analyzing whether the treated and control schools exhibited differential pretreatment trends. Evidence of parallel pretreatment trends would support the counterfactual assumption that the treated schools' outcomes would have evolved similarly to control schools' in the absence of the intervention. The

[10] A potential concurrent shock is the introduction of school climate measures into California's accountability system through the California School Dashboard, launched in Fall 2017. The Dashboard incorporated suspension rates as a formal accountability indicator, creating pressure on schools to reduce disciplinary actions. However, several features of the setting mitigate this concern. First, the Dashboard applied uniformly to all LAUSD schools, meaning both treated and control schools faced identical accountability pressure; any common response would be differenced out in the DiD design. The threat to identification requires that this pressure differentially affect treated schools. Second, while treated schools had higher baseline suspension rates (mean 1.70%) than control schools (mean 0.64%), the treated group's mean falls within the Dashboard's "Low" status category for unified districts (1.1–2.5%), where schools typically receive Green or Yellow ratings rather than the Red designation that triggers formal accountability consequences. Only 1 treated school and only 1 control school have a rate above 8.1% and would have faced heightened scrutiny.

[11] A potential concern is whether California's Local Control Funding Formula (LCFF), adopted in 2013, differentially affected treated and control schools prior to the treatment period. LCFF directs supplemental and concentration grants to districts based on their share of high-need students, those who are low-income, English learners, or foster youth, at the district level rather than the school site level. Because both treated and control schools belong to LAUSD, all schools in our sample faced the same LCFF-induced funding environment, and any district-wide LCFF-driven changes would be absorbed by our year fixed effects. Within-district allocation of LCFF funds was determined by LAUSD's Local Control Accountability Plan (LCAP), and both treated and control schools have similarly high shares of free lunch-eligible students (87% vs. 84%), suggesting broadly comparable baseline need. Finally, the event study shows flat pre-trends for the main outcomes prior to the settlement, providing direct evidence against differential pre-treatment trajectories induced by LCFF or any other concurrent shock.

pretreatment coefficients (δ_{-3} and δ_{-2}) check whether the treated and control schools' outcomes followed parallel trends before the intervention. Ideally, these point estimates should be small in magnitude and statistically insignificant.

As a robustness check, I use the synthetic difference-in-differences (SDID) estimator developed by Arkhangelsky et al. (2021). This estimator has been increasingly adopted in education policy research, including recent work on religious school competition in Indonesia (Bazzi et al., 2025). SDID combines strengths from the synthetic control and traditional DID methods. Similarly to synthetic control, SDID reweights the control units and matches the pretreatment trends, thereby weakening reliance on the strict parallel trends assumption. Similarly to DID, SDID remains invariant to additive unit-level shifts and permits valid large-panel inference.

The SDID estimator assigns greater weight to control units whose pretreatment trajectories more closely resemble those of the treated units and emphasizes time periods that are similar to the treatment periods. This weighting scheme enhances robustness by ensuring that comparisons draw primarily on the most relevant controls and time periods, rather than relying equally on all the available data. By focusing on similar units and periods, SDID reduces the potential for bias from poor counterfactual matches, making it particularly well suited for settings where the treatment and control groups may differ in their baseline characteristics.

5 Results

5.1 Effects on school spending

Following the settlement, the beneficiary schools received substantial funding. While the settlement terms mandated expenditure in certain categories such as hiring of staff or professional development training, schools retained considerable discretion over the allocation of the remaining funds as long as they used them on services for high-need students.

Table 2 in Column (2) presents the DID estimates examining how the settlement funding affected school spending decisions. The results provide strong evidence that the settlement schools used their additional funding to expand staffing capacity, confirming that the financial intervention tangibly improved school resources. Relative to control schools, the treated schools increased their certified staff by approximately 3 per 1,000 students, an effect that is statistically significant at the 5% level. To put this magnitude in perspective, for a treated school of average size with 834 students, this estimate implies the addition of approximately 2.6 full-time equivalent certified staff members, such as classroom teachers, counselors, or psychiatric social workers. This substantial increase represents a 4.9% increase relative to the baseline level of 63.6 certified staff per 1,000 students in the treated schools. It encompassed teachers providing direct instruction, administrative personnel in certified positions, and specialized pupil services staff, all critical components of school capacity who require specialized credentials.[12] Additionally, the treated schools expanded their classified staff by almost 2 per 1,000 students, a significant effect at the 10% level corresponding to a 4.8% increase from the baseline of 37 classified staff per 1,000 students. For the average-sized treated school, this corresponds to adding approximately 1.5 full-time equivalent support personnel, including paraprofessionals, clerical staff, and other support staff who do not require certification but provide essential operational and instructional assistance. The event study results presented in Panels (a) and (b) in Figure 2 confirm that the pretreatment trends of treated and control schools followed similar paths for both staffing outcomes, supporting the parallel trends assumption underlying the DID design. In summary, the schools that received settlement funds meaningfully expanded their human capital resources rather than using the money for other purposes.

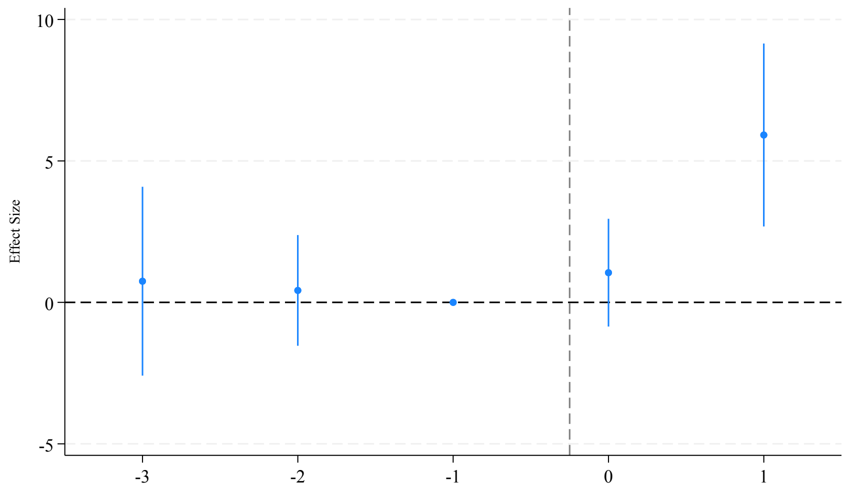
Table 2: Effects on School Spending Outcomes

	(1) Treated group mean (2017)	(2) DiD	(3) SDiD	(4) N
Certified staff per 1,000 st	63.564	3.093**	3.410***	850
		(1.199)	(1.119)	
Classified staff per 1,000 st	37.798	1.828*	1.709*	1020
		(0.942)	(0.904)	

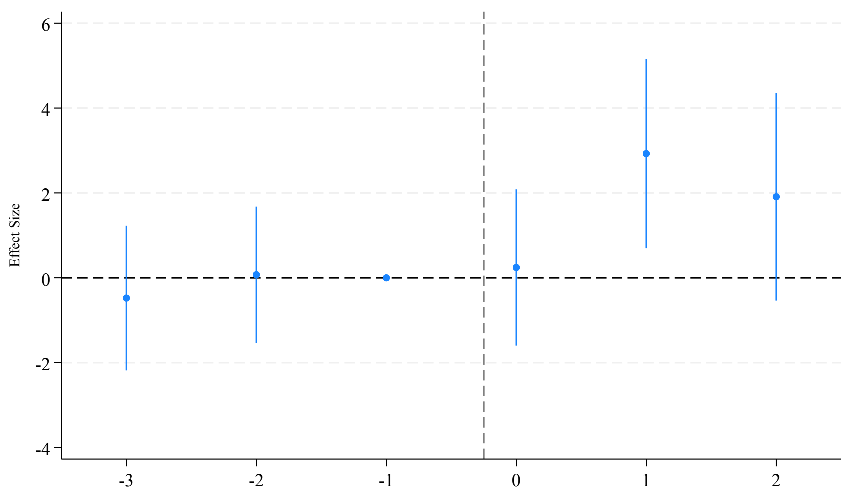
Notes: Certified staff includes teachers, administrators, and pupil services personnel requiring credentials. Certified staff data are available for years through academic year 2018/19. Classified staff includes paraprofessionals, clerical, and other support personnel not requiring certification. Column (1) shows the treated group mean in the academic year 2016/17 (N=49). Column (2) shows DiD estimates from regressions that include school and year fixed effects. Standard errors are clustered at the school level. Column (3) presents estimates from synthetic difference-in-differences estimation. Bootstrapped standard errors in parentheses (1,000 replications). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[12] One may worry that the increase in certified staff per 1,000 students is mechanical: if enrollment is declining in treated schools, the per-student ratio could rise even without any actual hiring. However, staffing decisions in LAUSD are driven by enrollment levels, so declining enrollment would if anything lead to *fewer* hires, not more. The fact that we observe an increase in certified staff per 1,000 students despite enrollment pressure thus reflects an increase in staffing resources, not a mechanical artifact of the denominator.

Figure 2: Effects on Spending Outcomes



(a) Certified Staff



(b) Classified Staff

Notes: This figure reports event study estimates of the settlement's effects on school staffing: Panel (a) for certified staff, which includes teachers, administrators, and pupil services personnel requiring credentials, and Panel (b) for classified staff, which includes paraprofessionals, clerical, and other support personnel. All regressions include school and year fixed effects and cluster standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed line indicates the start of the settlement period.

5.2 Effects on Student Outcomes

I begin by examining effects on school disciplinary climate, a natural starting point given two of the settlement's authorized expenditure categories: investments in high-need students, including academic support and mental health and social-emotional support; and school climate initiatives. The settlement also mandated that treated schools hire Restorative Justice Teacher Advisors, Pupil Services and Attendance Counselors, and/or Psychiatric Social Workers, positions directly linked to student behavioral outcomes through two pathways. First, restorative justice interventions have been shown to reduce disciplinary incidents by emphasizing reparative rather than punitive responses to student behavior (Adukia et al., 2025, Augustine et al., 2018). Second, research demonstrates that school counselors positively affect student outcomes, including reductions in disciplinary actions (Bell and Meyer, 2024, Carrell and Hoekstra, 2014, Carrell and Carrell, 2006, Hurwitz and Howell, 2014, Mulhern, 2023, Reback, 2010). Table 3, Column (2) presents the estimated effects on suspension rates, and Figure 3 displays the event study results. The event study shows no evidence of differential pre-trends between treated and control schools, supporting the validity of the parallel trends assumption. The treated schools experienced a significant reduction in suspension rates of 0.75 pp (significant at the 5% level), which represents a 44% decrease relative to their baseline suspension rate of 1.7%.

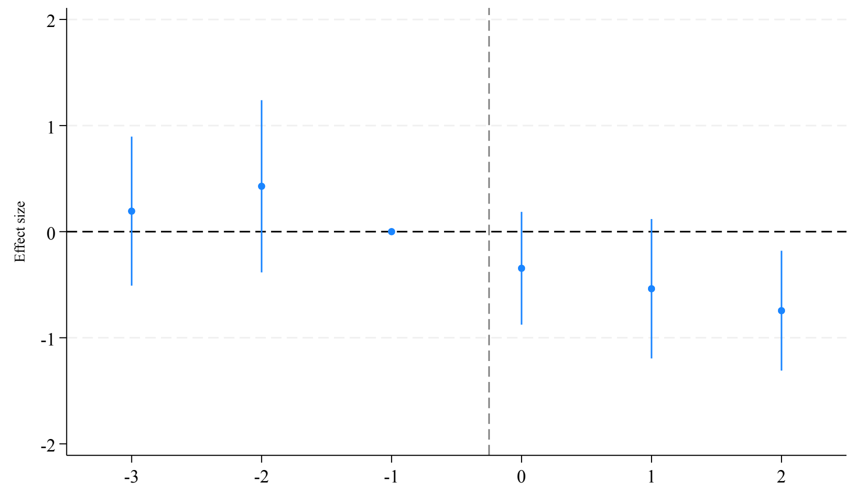
Table 3: Effects on students' outcomes

	(1) Treated group mean (2017)	(2) DiD	(3) SDiD	(4) N
Suspension rate	1.705	-0.750**	-0.575*	1020
		(0.313)	(0.321)	
% Math met	7.475	0.961	1.154*	679
		(0.605)	(0.590)	
% ELA met	21.897	2.804***	2.759***	679
		(0.867)	(0.820)	

Notes: Suspension rate is measured as the percentage of students suspended at least once during the academic year. The math outcome is the percentage of students who met the state standards for the math standardized test, and ELA outcomes is the percentage of students who met the state standards for the English Language Arts standardized test. Column (1) shows the treated group mean in the academic year 2016/17 (N=49). Column (2) shows DiD estimates from regressions that include school and year fixed effects. Standard errors are clustered at the school level. Column (3) presents estimates from synthetic difference-in-differences estimation. Bootstrapped standard errors in parentheses (1,000 replications). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To contextualize this magnitude, I compare these results to two randomized controlled trials of restorative practices interventions. Adukia et al. (2025) study Chicago Public Schools' rollout of restorative practices training to school staff across 73 high schools, which emphasized less punitive

Figure 3: Effect on Suspension Rate



Notes: This figure reports event study estimates of the settlement's effects on the student suspension rate. The suspension rate is measured as the percentage of students suspended at least once during the academic year. The regression includes school and year fixed effects and clusters standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed line indicates the start of the settlement period.

and more reparative strategies when engaging with students, including developing restorative protocols in response to disciplinary incidents and strengthening student-teacher relationships. They find a 14% reduction in suspension rates, approximately one-third the size of the effect I observe. Augustine et al. (2018) evaluate the International Institute for Restorative Practices' SaferSanerSchools program in Pittsburgh Public Schools, which provided intensive training and coaching to school staff on restorative practices over two years, and find a 2 pp reduction in the proportion of students suspended (a 13% decrease from a 15.8% baseline). The effect I observe in Los Angeles is larger than both of these estimates. This finding is robust to alternative specifications: using the SDID estimator (Arkhangelsky et al., 2021) to improve comparability between treated and control schools, I find a consistent reduction of 0.575 pp (significant at the 10% level; see Table 3, Column (3)). The consistency of estimates across specifications is reassuring and supports the interpretation that the settlement improved disciplinary climate in treated schools.

Next, I examine effects on academic outcomes, measured by the percentage of students who met the state standard in mathematics and the percentage who met the standard in English Language Arts. This analysis is motivated by the settlement's requirements: one authorized expenditure category specifically mandated significant investments in academic support for high-need students, and the settlement required expenditures on professional development for teachers. In a meta-analysis, Fryer (2017) shows that professional development interventions have positive effects on standardized reading and mathematics test scores. While these settlement provisions suggest potential for academic gains, the evidence I present is more limited than for the suspension results, as I observe only two pre-treatment periods for academic outcomes (see Figure A.3), which limits my ability to show parallel trends. With this caveat in mind, Table 3, Column (2) presents the main DiD estimates, which suggest improvements in both subjects. In English Language Arts, treated schools achieved a 2.8 pp increase in the share of students meeting the state standard (significant at the 1% level), representing a 12.8% improvement from the baseline of 21.9%. In mathematics, treated schools increased the share of students who met standards by 0.96 pp, though this effect is not statistically significant. [13] As an additional robustness check, I re-estimate the effects using the SDID estimator (See Table 3, Column (3)). The SDID estimates show a 2.8 pp increase in ELA proficiency (significant at the 1% level), the same as the main estimate, and a 1.15 pp increase in math proficiency (significant at the 10% level), an increase in both magnitude and precision compared to the main estimate. The consistency of estimates across specifications is reassuring, though given the limited pre-period data, these academic achievement results should be interpreted as suggestive rather than definitive evidence of causal effects.

A potential concern with the main outcomes I consider is mean reversion. Since suspension rates and math performance were part of the criteria used to select treated schools, if treated schools were at extreme values of any outcome at baseline, regression to the mean could bias estimates upward. To address this, I re-estimate the main specifications using log-transformed outcomes. Mean reversion operates mechanically on levels, that is, schools with high or low outcomes have more room to move in absolute terms, but proportional changes are harder to explain by mean reversion alone. Table A.2 presents the results. For suspension rates, I use $\ln(\text{suspension rate} + 1)$ to accommodate schools with zero suspensions; the log estimate is -0.269 (significant at the 1% level), implying a reduction of approximately 24% relative to the pre-treatment mean, in line with the 44% reduction in levels reported in the main specification. For ELA, the log estimate is 0.102 (significant at the 1% level), implying an approximately 10% improvement relative to baseline. The math estimate is positive but not significant (0.095, se: 0.064), again in line with the main result.

[13] As a robustness check, I re-estimate the main specifications weighting observations by total enrollment, so that larger schools receive proportionally more weight and the estimand corresponds to the average treatment effect across students rather than schools. The weighted estimates are -0.69 pp for suspension rates (significant at the 1% level), 0.93 pp for math (not significant), and 2.90 pp for ELA (significant at the 1% level), consistent with the main results.

The alignment of results across specifications suggests that the main findings are not driven by mechanical reversion toward the mean.

As a robustness check, I examined the dynamic effects of the settlement using the synthetic difference-in-differences event study estimator proposed by Ciccia (2024), which extends the SDID framework of Arkhangelsky et al. (2021) to allow for period-by-period treatment effect estimates. To assess the reliability of the SDID estimates, I first examine the control unit weights reported in Table A.3. The weights are remarkably flat: the highest-weighted school receives a weight of approximately 0.02 for the suspension rate outcome, while most schools receive weights in the range of 0.007–0.010, close to the uniform weight of $1/121 \approx 0.008$ that a standard difference-in-differences estimator would assign. No single school dominates the synthetic control for any of the main outcomes, and all comparison schools receive positive weight, suggesting that the SDID estimates are not sensitive to the inclusion or exclusion of any particular control school. Figure A.4 shows the SDID event study results for suspension rates. The placebo estimates are small in magnitude and statistically insignificant, supporting the parallel trends assumption. The post-treatment effects grow over time, with the largest and most precisely estimated reduction occurring in the third year of the settlement, consistent with schools requiring time to implement new practices and personnel. Figure A.5 shows corresponding results for math and ELA achievement. Pre-treatment placebo estimates are again close to zero and insignificant for both outcomes, while post-treatment effects are positive and significant, corroborating the main DiD findings.

5.3 Effects Using Continuous Treatment Intensity

The binary treatment specification in equation (1) treats the settlement as a uniform shock, but funding intensity in fact varied meaningfully across treated schools: the settlement allocated approximately $1,030 per duplicated pupil, generating budget increases ranging from 7.4 % to 17.8% across the 49 beneficiary schools. To assess whether outcomes responded to the magnitude of funding rather than simply its presence, I re-estimate the main specifications replacing the binary $Settlement_s$ indicator with each school's initial (2016/17) funding allocation as a percentage of its baseline budget, interacted with $Post_t$. I use each school's initial allocation, determined prior to the treatment period.

Table A.8 presents the results. The continuous specification confirms the pattern from the main analysis: a one percentage-point increase in funding intensity is associated with a 0.044 pp reduction in the suspension rate (significant at the 5% level) and a 0.206 pp increase in ELA proficiency (significant at the 1% level), while the effect on math proficiency remains small and statistically insignificant. Evaluated at the sample mean funding intensity of 13.5%, these estimates imply effects of similar magnitude to the binary DiD results in Tables 3, providing reassurance that the main findings are not an artifact of the binary treatment coding and that outcomes scale sensibly with the size of the resource shock.

5.4 Effects on enrollment

It is possible that the settlement and the money schools received affected enrollment and student body composition, although the expected direction of this effect is theoretically ambiguous. Prior research shows that parents respond to school quality information by enrolling their children in higher-quality schools (Andrabi et al., 2017, Campos, 2024, Corradini, 2024, Hastings and Weinstein, 2008, Hussain, 2023). However, the settlement in this context generated two competing information signals with potentially opposing effects on enrollment. First, the influx of additional resources could signal improved school quality and investment, attracting families seeking better educational opportunities for their children. Second, the public nature of the litigation and identifying schools as underfunded and underperforming could stigmatize these institutions and deter enrollment, particularly among families with greater school choice options.[14] In either case, it is unclear ex ante which types of students would be most likely to move, if any.

If the treated schools were perceived as improving because of the additional resources, they may have attracted students of varying academic abilities, such as high-achieving students seeking better opportunities, average students drawn to enhanced programs, or struggling students hoping for additional support. Each scenario would imply a different effect: The arrival of more high-achieving students would make outcomes such as test scores and suspension rates appear better than the true treatment effect, while an influx of struggling students could make these same outcomes appear worse. Conversely, if the litigation stigma caused the treated schools to be perceived as low-quality, high-achieving students might have exited to other schools, leaving behind a more disadvantaged student body.

I thus examine whether the settlement affected student enrollment patterns and demographic composition across the treated and control schools. Table 4, Column (2) presents the DiD estimated effects on enrollment and student demographics, while Figures A.6, A.7, and A.8 display the

[14] The settlement received substantial news coverage. See, for example, Article 1, Article 2, Article 3, Article 4, and Article 5.

event study results. The results reveal changes in the size and composition of the treated schools' student bodies. On average, the treated schools experienced an enrollment decline of 48 students (approximately 5.8% from their baseline levels), suggesting that the intervention was associated with net outflows rather than enrollment increases. In terms of student body composition, the treated schools experienced a 0.9 pp decrease in Black student enrollment and a 3.6 pp increase in their enrollments of students eligible for free lunch, indicating a concentration of economic disadvantage. Additionally, the share of students belonging to other racial categories increased by 0.5 pp, a small change given that each of these groups represents less than 2% of enrollment. The enrollment shares of Hispanic, White, and Asian students remained statistically unchanged. These patterns suggest that families with greater school choice options, likely those with more resources and flexibility, left the treated schools after the settlement, while more economically disadvantaged families remained or were drawn to these schools.

Table 4: Effects on Enrollment and Student Body Composition

	(1) Treated group mean (2017)	(2) DiD	(3) SDiD
N of students enrolled	833.592	-48.203**	-23.051
		(19.001)	(13.994)
% Hispanics	82.092	0.531	0.551*
		(0.344)	(0.332)
% Blacks	14.418	-0.908***	-0.767***
		(0.285)	(0.274)
% Whites	1.210	-0.067	-0.129
		(0.145)	(0.154)
% Asian	0.887	-0.045	-0.109
		(0.116)	(0.093)
% Other less 2% races	1.393	0.489***	0.378***
		(0.117)	(0.089)
% Free lunch	87.407	3.629***	3.143***
		(0.827)	(0.806)
Observations	49	1020	

Notes: Demographic variables are measured as percentages of total enrollment. Column (1) shows the treated group mean in the academic year 2016/17. Column (2) shows DiD estimates from regressions that include school and year fixed effects. Standard errors are clustered at the school level. Column (3) presents estimates from synthetic difference-in-differences estimation. Bootstrapped standard errors in parentheses (1,000 replications). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

To assess the robustness of these findings, I re-estimate the effects using the SDID estimator. Table 4, Column (3) presents these results. The demographic composition changes remain largely consistent: Black enrollment decreases by 0.77 pp (significant at the 1% level), free lunch eligibility increases by 3.1 pp (significant at the 1% level), and other racial categories increase by 0.38 pp (significant at the 1% level). The enrollment effect, however, is reduced to 23 students and is not statistically significant at conventional levels (p-value = 0.10). While this suggests the enrollment decline may be less robust than the compositional changes, I take a conservative approach in my subsequent analysis and treat the observed compositional shifts as potential treatment effects that could bias my main estimates.

A potential concern with the main estimates is that comparison schools located near beneficiary schools may have been indirectly affected by the settlement. The results in this section show that enrollment declined in treated schools following the settlement, and students who left may have transferred to nearby comparison schools, affecting their composition and biasing the estimates. To address this, I re-estimate the main specifications after dropping comparison schools within 2 miles of any beneficiary school at the same school level, leaving 74 of the original 121 comparison schools in the sample. Figure A.9 presents the event study results and Table A.4 the DiD estimates. The findings are consistent with the main results. The estimated reduction in suspension rates is 0.65 pp (significant at the 10% level), compared to 0.75 percentage points in the main specification. The ELA effect is 1.81 pp, positive but smaller in magnitude than the main estimate of 2.80 pp and remains statistically significant. The overall pattern suggests that proximity-based spillovers are unlikely to be driving the main findings.

6 Understanding Suspension and Enrollment: Bounding Evidence

The shifts in student composition could affect the interpretation of my main findings on disciplinary outcomes. If the families who left included students with higher (or lower) suspension risk, the observed decline in disciplinary incidents could be overestimated (or underestimated), reflecting changes in student body composition rather than changes in school practices alone. To assess this potential bias, I implement a bounding exercise in the spirit of Horowitz and Manski (2000) that assesses how sensitive my estimates are to different assumptions about the characteristics of students who left.

The bounding approach is informative because I cannot observe the counterfactual suspension outcomes for the students who exited the treated schools, which creates a missing data problem that could bias my suspension treatment effect estimates positively or negatively. To establish bounds around the true treatment effect, I consider four scenarios that span a range of plausible assumptions about the propensity of departing students to be suspended, from lowest (Scenario 1) to highest (Scenario 4).

In Scenario 1, I assume that the departures concentrated among well-behaved students, that is, students who would not have been suspended under any circumstances. Under this scenario, my estimated treatment effect understates the true policy impact, as the remaining student body would be more challenging to discipline. In Scenario 2, I assume that the departing students had the same suspension rate as the treated schools' baseline (1.7%), meaning approximately 0.81 of the 48 students who departed per school would have been suspended. In Scenario 3, I assume that the departing students had suspension rates matching their demographic group's baseline, with that of Hispanic students being 1.10% and that of Black students 4.67%; this results in approximately 1.11 expected suspensions among the 48 departing students per school. In Scenario 4, I assume that the departing students had elevated suspension rates reflecting twice the treated schools' baseline (3.4%), meaning approximately 1.63 of the 48 students who departed per school would have been suspended.

For each scenario, I impute counterfactual suspension outcomes for the students who would have remained under the control condition, reconstructing what the treated schools' suspension rates would have looked like with their original student composition. I then recompute the treatment effects using these adjusted denominators and suspension counts. This exercise provides bounds on the true treatment effect under different assumptions about selection into departure.

Across all scenarios, the estimated effects have the same sign and similar magnitude as the original estimate, a 0.75 pp reduction in suspension rates (significant at the 5% level) (see Table 5). Under Scenario 1, where well-behaved students departed, the treatment effect increases to -0.82 pp (significant at the 1% level), suggesting even stronger impacts than initially estimated. Under Scenario 2, where the departing students had the same suspension rate as the treated schools' baseline, the treatment effect remains substantial at -0.71 pp (significant at the 5% level). Under Scenario 3, where the departing students had suspension rates matching their demographic baseline, the estimated reduction is -0.67 pp (significant at the 5% level). Under Scenario 4, where the departing students had twice the treated schools' baseline suspension rate, the treatment effect is -0.60 pp (significant at the 10% level). Overall, the results of this bounding exercise suggest that the observed improvements in the suspension rate are not a mechanical result of the student composition but reflect genuine changes in school practices.

Table 5: Treatment Effects Under Different Assumptions about Characteristics of Departing Students

Variable	Original	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Treatment Effect	-0.750**	-0.815***	-0.708**	-0.669**	-0.600*
	(0.313)	(0.308)	(0.307)	(0.307)	(0.307)
Observations	1020	1020	1020	1020	1020

Notes: This table presents treatment effect estimates from difference-in-differences regressions under different assumptions about the characteristics of the students who departed the treated schools. The dependent variable is the suspension rate. Original shows the baseline estimates. Scenario 1 assumes that the departing students were well behaved (zero suspension probability). Scenario 2 assumes that the departing students had the same suspension rate as the treated schools' baseline (1.7%). Scenario 3 assumes that the departing students had suspension rates matching their demographic group's baseline (1.10% for Hispanic, 4.67% for Black students). Scenario 4 assumes that the departing students had suspension rates twice the treated schools' baseline (3.4%). All regressions include school and year fixed effects. Standard errors clustered at the school level are reported in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

7 Why did suspension rates improve?

The reduction in suspension rates following the funding intervention raises the question of what mechanisms drove this change. The observed decrease could reflect effects operating through several distinct pathways. First, schools under scrutiny or receiving targeted resources may strategically reduce recorded suspensions without meaningful changes to underlying practices–a form of administrative gaming analogous to concerns in the accountability literature about teaching to the test (Cullen and Reback, 2006, Figlio and Getzler, 2006, Jacob and Levitt, 2003, Neal and Schanzenbach, 2010). Second, student behavior itself might improve, though existing evidence suggests that variation in disciplinary outcomes is better explained by school-level factors than by student behavior alone (Welsh and Little, 2018). Third, educators may change how they respond to student behavior, either through enhanced capacity to manage incidents or through shifts in the threshold at which behaviors trigger formal discipline. Welsh and Little (2018) emphasize that disciplinary disparities emerge from "the policies, practices, and perspectives of teachers and principals," suggesting educator responses as a key lever for policy intervention. Fourth, broader improvements in school climate may reduce both the incidence of problematic behavior and the reliance on exclusionary responses. Evaluations of restorative justice programs support this interpretation: Adukia et al. (2025) find that Chicago's restorative practices initiative improved students' perceptions of school climate, with impacts driven by students' perceptions of their peers' classroom behavior, their sense of belonging, and feelings of school safety. Distinguishing among these mechanisms is crucial for interpreting the policy's effectiveness, as a reduction driven by administrative gaming would represent a very different policy success than one reflecting improved educator capacity or school climate.

I leverage comprehensive survey data from both students and school staff collected through LAUSD's SES. This annual survey captures feedback from teachers, staff, students, and parents across all LAUSD schools and provides detailed measures of school climate, bullying experiences, and disciplinary practices. The staff component of the survey provides insights, capturing responses from principals, administrators, teachers, counselors, and support staff. The staff questionnaire measures preparedness to address climate issues through professional development indicators, engagement in addressing bullying incidents, personal safety perceptions, and assessments of disciplinary fairness and effectiveness. Staff also evaluate the prevalence of various climate challenges including harassment, bullying, physical fighting, disruptive behavior, racial conflicts, and respectfulness of students toward staff. The student survey assesses experiences with four key forms of peer victimization on school property (physical aggression, relational aggression, sexual harassment, and appearance-based teasing), alongside six dimensions of school climate including perceptions of adult respect, safety, belonging, and overall satisfaction.[15]

To analyze these outcomes, I create all the indices so that higher positive values consistently indicate better outcomes: improved school climate, reduced bullying, enhanced safety, and more effective disciplinary practices. This uniform scaling allows straightforward interpretation of the treatment effects across all the survey measures, with positive coefficients representing improvements in the school environment and student experiences. One concern with survey-based outcomes is differential response rates. For example, if treated schools systematically increased survey participation following the settlement, the observed improvements could reflect compositional changes in who responds rather than genuine shifts in climate or practices. To assess this, Figure A.10 presents event study estimates using response rates as the outcome for staff and student surveys, respectively. Neither figure shows evidence of differential changes in response rates following the settlement: the point estimates remain close to zero and statistically insignificant in the post-treatment period,

[15]See Appendix C for details on all the questions I included in each index.

suggesting that the survey-based findings are not driven by selective participation.

Table 6 presents the estimated treatment effects on staff survey outcomes, with Figure 4 showing event study results that support the parallel trends assumption. The results point toward mechanism three–changes in educator capacity and approach. Staff reported substantial improvements in their ability to handle bullying incidents and received enhanced professional development on these issues, with the handling bullying index improving by 0.38 SD. Staff perceptions of disciplinary fairness and effectiveness improved by 0.31 SD. These findings suggest that the settlement successfully enhanced educators' tools for managing student behavior. Additionally, staff indicated that bullying incidents had become less prevalent (0.26 SD improvement) and reported improved overall school climate (0.37 SD), consistent with mechanism four. All effects are statistically significant at the 1% level.[16]

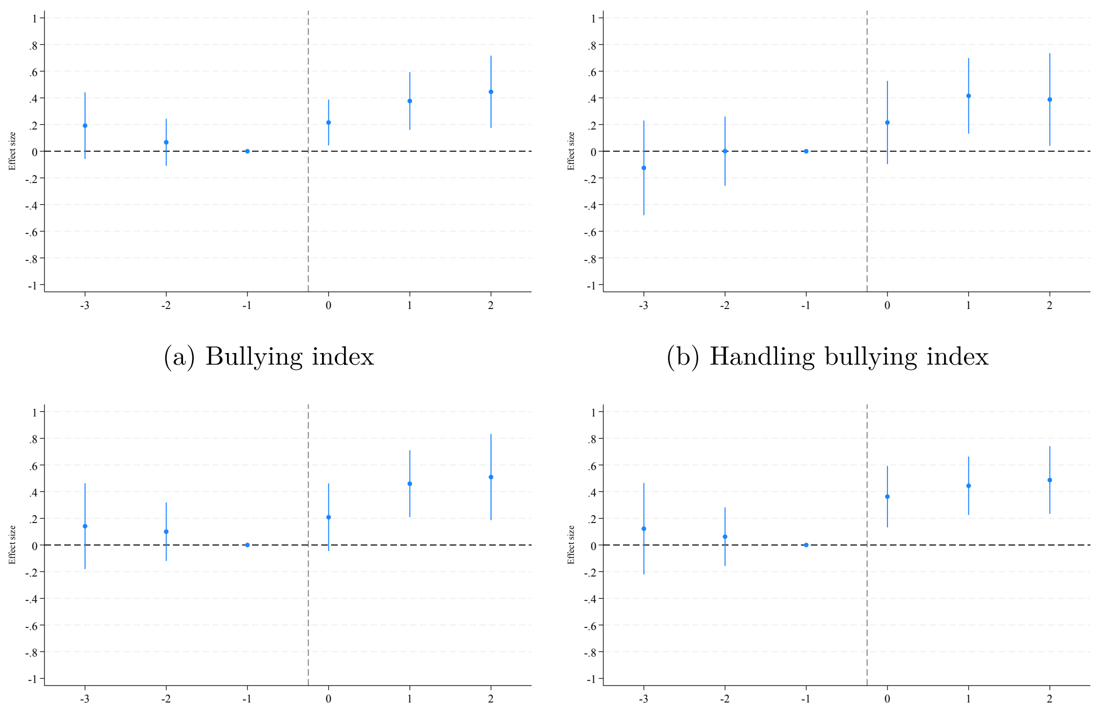
Table 6: Effects on Staff's School Experience Indices

	(1)	N
Bullying index	0.260***	912
	(0.082)	
Handling bullying index	0.381***	911
	(0.113)	
Discipline index	0.312***	912
	(0.106)	
School climate index	0.370***	912
	(0.092)	

Notes: All regressions include school and year fixed effects. Standard errors are clustered at the school level. I construct the survey indices following Anderson (2008) and normalize them to have mean zero and standard deviation one using the full sample across all years. On all indices, a higher value means a better outcome. The bullying index in Panel (a) includes three questions on how much of a problem staff thinks bullying, physical fighting, and discrimination are among students. The handling bullying index in Panel (b) contains two survey questions related to receiving professional development on preventing bullying and whether they address bullying. The discipline index in Panel (c) includes two questions about the fairness and effectiveness of discipline at school. The school climate index in Panel (d) has four questions on safety, respect, and disruptive behavior. Staff includes teachers, administrators, counselors, and other school personnel. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

[16]Table A.7 reports the effects on the individual components underlying each index.

Figure 4: Effects on Staff's School Experience Indices



(a) Bullying index

(b) Handling bullying index

(c) Discipline index

(d) School climate index

Notes: This figure reports event study estimates of the settlement's effects on survey outcomes on staff's school experience. The regressions include school and year fixed effects and cluster standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed line indicates the start of the settlement period. I construct the survey indices following Anderson (2008) and normalize them to have mean zero and standard deviation one using the full sample across all years. On all indices, a higher value means a better outcome. The bullying index in Panel (a) includes three questions on how much of a problem staff thinks bullying, physical fighting, and discrimination is among students. The handling bullying index in Panel (B) contains two survey questions related to receiving professional development on preventing bullying and whether they address bullying. The discipline index in Panel (c) includes two questions about the fairness and effectiveness of discipline at school. The school climate index in Panel (d) has four questions on safety, respect, and disruptive behavior. Staff includes teachers, administrators, counselors, and other school personnel.

Table A.5 and Figure A.11 present corresponding results for students. Students reported meaningful improvements in overall school climate (0.15 SD, significant at the 5% level), corroborating staff perceptions and providing additional support for mechanism four. Reported bullying experiences show a positive but statistically insignificant point estimate (0.12 SD), offering limited support for mechanism two–that student behavior itself fundamentally changed.[17]

Taken together, these patterns help distinguish among the competing mechanisms outlined above. The improvements in staff-reported preparedness, professional development, and percep-

[17] Table A.6 reports the effects on the individual components underlying each index.

tions of disciplinary effectiveness argue against mechanism one (pure administrative gaming) since strategic underreporting of suspensions would not generate genuine improvements in educators' self-assessed capacity to handle incidents. Instead, the findings point toward mechanisms three and four operating in tandem: the settlement enhanced educators' tools and approaches for managing student behavior while simultaneously improving the broader school climate. This interpretation aligns with the settlement's emphasis on restorative justice training and suggests that reduced suspension rates reflect more thoughtful, effective disciplinary practices rather than simply more lenient enforcement.

8 Conclusion

This paper provides new evidence on a fundamental question in education policy: Can investing in failing schools help them improve? Using a natural experiment from a 2017 settlement in LAUSD, I examine the effects of a substantial increase in the resources delivered to the district's 50 lowest-performing secondary schools. The settlement provided $1,030 per student annually for three years, a figure representing a 13.5% budget increase on average.

The findings reveal a nuanced picture of how investment affects low-performing schools. On one hand, the schools successfully implemented the required changes and achieved meaningful improvements in their disciplinary climate. They increased their employment of certified staff by approximately 3 per 1,000 students and support staff by almost 2 per 1,000 students, which represent 5% increases over baseline levels. More importantly, suspension rates fell by 0.75 pp, a remarkable 44% reduction from pretreatment levels. Survey evidence suggests this improvement reflects genuine changes in educator practices and school climate rather than administrative adjustments, with teachers reporting enhanced capacity to address behavioral issues and more effective disciplinary practices.

A back-of-the-envelope cost-effectiveness calculation further contextualizes the intervention. The settlement cost approximately $1,030 per high-need pupil per year, totaling $171.6 million over three years across 50 treated schools. With a baseline suspension rate of 1.70% and average enrollment of approximately 834 students per school, the 0.75 pp reduction implies roughly 6.3 fewer suspended students per school per year, or approximately 309 fewer suspensions annually across all treated schools. This translates to a cost of approximately $185,000 per avoided suspension per school per year. The $185,000 figure represents an upper bound, since the settlement funded a broad range of initiatives beyond discipline-related inputs. Approximately 50% of settlement funds were directed toward personnel — the category most directly linked to the observed suspension reductions through hiring of restorative justice advisors, counselors, and psychiatric social workers. Attributing only the personnel expenditure to the suspension outcomes implies a cost of approximately $92,500 per avoided suspension, which is high but reflects a broader investment in school capacity that generated improvements across multiple dimensions of school quality.

On the other hand, these improvements came with unintended consequences for school composition. The settlement's dual nature, providing additional resources while publicly identifying schools as underperforming, could have triggered enrollment responses, as prior research shows that families respond to school quality information by enrolling their children in higher-quality schools (Andrabi et al., 2017, Campos, 2024, Hastings and Weinstein, 2008). Consistent with this possibility, the treated schools experienced demographic sorting, losing 48 students per school on average, such that their concentration of economically disadvantaged students rose. Nonetheless, a simple bounding exercise that accounts for this demographic sorting indicates that the settlement still had meaningful effects on school climate.

These findings extend the robust evidence that money matters for schools to the specific context of low-performing institutions. I thereby address a critical gap between general school spending research and policy debates about how to improve the lowest-performing schools. While previous studies demonstrate positive effects of increased spending across diverse school populations, my research shows that investment can meaningfully improve practices even in the most challenging educational environments.

The results highlight how important it is to examine outcomes beyond test scores in evaluations of school interventions. They underscore that school quality encompasses multiple dimensions and that interventions may succeed along some margins while showing limited effects on others. Despite the limited changes in academic achievement, the substantial improvements in disciplinary climate represent meaningful progress in noncognitive dimensions of schooling. These reductions in exclusionary discipline matter for students' long-term development and have consequences beyond the classroom. Previous research has documented that suspensions increase contemporaneous arrests (Adukia et al., 2025) and that attending schools with stricter discipline policies raises the likelihood of suspension and arrest later in life (Bacher-Hicks et al., 2024). Applying the estimates from Adukia et al. (2025) to my findings suggests that the observed suspension reductions could reduce arrests by approximately 41%. While this applies elasticities from a different context, it illustrates that the economic returns to these school investments may extend well beyond measured academic outcomes.

References

Abadie, Alberto, Susan Athey, Guido W Imbens, and Jeffrey M Wooldridge, "When Should You Adjust Standard Errors for Clustering?," *The Quarterly Journal of Economics*, February 2023, *138* (1), 1–35.

Abdulkadiroglu, Atila, Parag A. Pathak, and Christopher R. Walters, "Who Gets What in Education: Can School Matching Improve Student Achievement?," NBER Working Paper 34936, National Bureau of Economic Research March 2026.

Abott, Carolyn, Vladimir Kogan, Stéphane Lavertu, and Zachary Peskowitz, "School District Operational Spending and Student Outcomes: Evidence from Tax Elections in Seven States," *Journal of Public Economics*, March 2020, *183*, 104142.

Adukia, Anjali, Benjamin Feigenberg, and Fatemeh Momeni, "From Retributive to Restorative: An Alternative Approach to Justice in Schools," Working Paper 31675, National Bureau of Economic Research, Cambridge, MA February 2025. Revised February 2025.

Anderson, Michael L., "Multiple Inference and Gender Differences in the Effects of Early Intervention: A Reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects," *Journal of the American Statistical Association*, 2008, *103* (484), 1481–1495.

Andrabi, Tahir, Jishnu Das, and Asim Ijaz Khwaja, "Report Cards: The Impact of Providing School and Child Test Scores on Educational Markets," *American Economic Review*, June 2017, *107* (6), 1535–1563.

Arkhangelsky, Dmitry, Susan Athey, David A. Hirshberg, Guido W. Imbens, and Stefan Wager, "Synthetic Difference-in-Differences," *American Economic Review*, 2021, *111* (12), 4088–4118.

Atchison, Drew and Damon Blair, "School Accountability, Underfunding, and Inefficiency," *AERA Open*, 2026, *12* (1), 1–23.

Augustine, Catherine H, John Engberg, Geoffrey E Grimm, Emma Lee, Elaine Lin Wang, Karen Christianson, and Andrea A Joseph, "Can Restorative Practices Improve School Climate and Curb Suspensions? An Evaluation of the Impact of Restorative Practices in a Mid-Sized Urban School District," Research Report RR-2840-DOJ, RAND Corporation, Santa Monica, CA 2018.

Bacher-Hicks, Andrew, Stephen B Billings, and David J Deming, "The School-to-Prison Pipeline: Long-Run Impacts of School Suspensions on Adult Crime," *American Economic Journal: Economic Policy*, 2024, *16* (4), 165–193.

Baron, E. Jason, "School Spending and Student Outcomes: Evidence from Revenue Limit Elections in Wisconsin," *American Economic Journal: Economic Policy*, February 2022, *14* (1), 1–39.

Bazzi, Samuel, Masyhur Hilmy, and Benjamin Marx, "Religion, Education, and the State," *Review of Economic Studies*, 2025. Advance access publication 26 September 2025.

Bell, Elizabeth and Katharine Meyer, "Does Reducing Street-Level Bureaucrats' Workload Enhance Equity in Program Access? Evidence from Burdensome College Financial Aid Programs," *Journal of Public Administration Research and Theory*, 2024, *34*, 16–38.

Biasi, Barbara, Julien Lafortune, and David Schonholzer, "What Works and for Whom? Effectiveness and Efficiency of School Capital Investments Across the U.S.," *The Quarterly Journal of Economics*, 08 2025, *140* (3), 2329–2379.

Bifulco, Robert and David J. Schwegman, "Who Benefits from Accountability-Driven School Closure? Evidence from New York City," *Journal of Policy Analysis and Management*, 2020, *39* (1), 96–130. Published online 2019.

Blair, Damon, Drew Atchison, and Kerstin Carlson Le Floch, "Investing in Improvement: A Critical Look at Funding for Low-Performing Schools," *Educational Policy*, 2026, *40* (1), 3–40. Published online 2025.

Booher-Jennings, Jennifer, "Below the Bubble: 'Educational Triage' and the Texas Accountability System," *American Educational Research Journal*, 2005, *42* (2), 231–268.

Bryk, Anthony S., Penny B. Sebring, Elaine Allensworth, Stuart Luppescu, and John Q. Easton, *Organizing Schools for Improvement: Lessons from Chicago*, Chicago: The University of Chicago Press, 2010.

Campos, Christopher, "Social Interactions, Information, and Preferences for Schools: Experimental Evidence from Los Angeles," December 2024. Working paper.

Candelaria, Christopher A. and Kenneth A. Shores, "Court-Ordered Finance Reforms in the Adequacy Era: Heterogeneous Causal Effects and Sensitivity," *Education Finance and Policy*, January 2019, *14* (1), 31–60.

Carrell, Scott E. and Mark Hoekstra, "Are school counselors an effective education input?," *Economics Letters*, 2014, *125*, 66–69.

_ **and Susan A. Carrell**, "Do Lower Student to Counselor Ratios Reduce School Disciplinary Problems?," *Contributions to Economic Analysis & Policy*, 2006, *5* (1), Article 11.

Carrell, Scott, Michal Kurlaender, Paco Martorell, Matthew Naven, and Christina Sun, "Do Schools Matter? Measuring the Impact of California High Schools on Test Scores and Postsecondary Enrollment," September 2025. Working paper.

Ciccia, Diego, "A Short Note on Event-Study Synthetic Difference-in-Differences Estimators," 2024. Sciences Po.

Corradini, Viola, "Information and Access in School Choice Systems: Evidence from New York City," October 2024. Working paper.

Crespin, Rene A., "The Value of School Social Climate Information: Evidence from Chicago Housing Transactions," *American Economic Journal: Economic Policy*, 2026, *18* (3), 163â–96.

Cullen, Julie Berry and Randall Reback, "Tinkering Toward Accolades: School Gaming under a Performance Accountability System," in Timothy J. Gronberg and Dennis W. Jansen, eds., *Advances in Applied Microeconomics*, Vol. 14, Bingley: Emerald Group Publishing Limited, 2006, pp. 1–34.

Dragoset, Lisa, Jaime Thomas, Mariesa Herrmann, John Deke, Susanne James-Burdumy, Cheryl Graczewski, Andrea Boyle, Rachel Upton, Courtney Tanenbaum, and Jessica Giffin, "School Improvement Grants: Implementation and Effectiveness," Technical Report NCEE 2017-4013, National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education, Washington, DC 2017.

Figlio, David N., "Testing, Crime, and Punishment," *Journal of Public Economics*, 2006, *90*, 837–851.

_ **and Cecilia Elena Rouse**, "Do Accountability and Voucher Threats Improve Low-Performing Schools?," *Journal of Public Economics*, 2006, *90* (1–2), 239–255.

_ **and Lawrence S. Getzler**, "Accountability, Ability, and Disability: Gaming the System?," in Timothy J. Gronberg and Dennis W. Jansen, eds., *Advances in Applied Microeconomics*, Vol. 14, Bingley: Emerald Group Publishing Limited, 2006, pp. 35–49.

Han, Chunping, Margaret E. Raymond, James L. Woodworth, Yohannes Negassi, W. Payton Richardson, and Will Snow, "Lights Off: Practice and Impact of Closing Low-Performing Schools, 2017, Volume I," Technical Report, Center for Research on Education Outcomes (CREDO), Stanford University, Stanford, CA 2017.

Hastings, Justine S. and Jeffrey M. Weinstein, "Information, School Choice, and Academic Achievement: Evidence from Two Experiments," *The Quarterly Journal of Economics*, November 2008, *123* (4), 1373–1414.

Horowitz, Joel L. and Charles F. Manski, "Nonparametric Analysis of Randomized Experiments with Missing Covariate and Outcome Data," *Journal of the American Statistical Association*, 2000, *95* (449), 77–84.

Hurwitz, Michael and Jessica Howell, "Estimating Causal Impacts of School Counselors With Regression Discontinuity Designs," *Journal of Counseling & Development*, July 2014, *92.*

Hussain, Iftikhar, "Housing market and school choice response to school quality information shocks," *Journal of Urban Economics*, November 2023, *138*, 103606.

Hyman, Joshua, "Does Money Matter in the Long Run? Effects of School Spending on Educational Attainment," *American Economic Journal: Economic Policy*, November 2017, *9* (4), 256–280.

Jackson, C. Kirabo, "Does School Spending Matter? The New Literature on an Old Question," in "Confronting Inequality: How Policies and Practices Shape Children's Opportunities," Cornell University Press, 2020.

_ **and Claire L. Mackevicius**, "What Impacts Can We Expect from School Spending Policy? Evidence from Evaluations in the United States," *American Economic Journal: Applied Economics*, January 2024, *16* (1), 412–446.

_ **, Cora Wigger, and Heyu Xiong**, "Do School Spending Cuts Matter? Evidence from the Great Recession," *American Economic Journal: Economic Policy*, 2021, *13* (2), 304–335.

_ **, Rucker C. Johnson, and Claudia Persico**, "The Effects of School Spending on Educational and Economic Outcomes: Evidence from School Finance Reforms *," *The Quarterly Journal of Economics*, February 2016, *131* (1), 157–218.

_ **, Shanette C. Porter, John Q. Easton, Alyssa Blanchard, and Sebastian Kiguel**, "School Effects on Socioemotional Development, School-Based Arrests, and Educational Attainment," *American Economic Review: Insights*, 2020, *2* (4), 491–508.

Jacob, Brian A., "Accountability, Incentives and Behavior: The Impact of High-Stakes Testing in the Chicago Public Schools," *Journal of Public Economics*, 2005, *89* (5–6), 761–796.

Jacob, Brian and Steven D. Levitt, "Rotten Apples: An Investigation of the Prevalence and Predictors of Teacher Cheating," *Quarterly Journal of Economics*, 2003, *118* (3), 843–877.

Johnson, Rucker C. and C. Kirabo Jackson, "Reducing Inequality through Dynamic Complementarity: Evidence from Head Start and Public School Spending," *American Economic Journal: Economic Policy*, 2019, *11* (4), 310–349.

Jr., Roland G. Fryer, "The Production of Human Capital in Developed Countries: Evidence From 196 Randomized Field Experiments," in "Handbook of Economic Field Experiments," Vol. 2, Elsevier, 2017, chapter 2.

King, Michelle, "New Allocation – Support for High Need Schools," July 2017. Interoffice correspondence, Los Angeles Unified School District. Unpublished.

Lafortune, Julien, Jesse Rothstein, and Diane Whitmore Schanzenbach, "School Finance Reform and the Distribution of Student Achievement," *American Economic Journal: Applied Economics*, April 2018, *10* (2), 1–26.

Los Angeles Unified School District, "School Innovation Fund Planning and Implementation Guide 2017-2020," Technical Report, Los Angeles Unified School District September 2017. Internal planning document.

_ , "School Innovation Funds 2017-2020: Frequently Asked Questions," Technical Report, Los Angeles Unified School District September 2017. Internal guidance document. Released 9-20-17.

_ , "Settlement Agreement and Mutual Release: Community Coalition of South Los Angeles and Ms. Reyna Frias v. Los Angeles Unified School District," September 2017. Case No. BS 156259, Superior Court of California, County of Los Angeles.

_ , "Fingertip Facts 2025–26," Technical Report, Los Angeles Unified School District, Communications Office 2025. Accessed: March 6, 2026.

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

Neal, Derek and Diane Whitmore Schanzenbach, "Left Behind by Design: Proficiency Counts and Test-Based Accountability," *The Review of Economics and Statistics*, May 2010, *92* (2), 263–283.

Porter, Sarah C., C. Kirabo Jackson, Shanette Kiguel, and John Q. Easton, "High School Climate and Organizational Context Shape Student Development and Educational Attainment," Technical Report 2023.

Reback, Randall, "Teaching to the Rating: School Accountability and the Distribution of Student Achievement," *Journal of Public Economics*, 2008, *92* (5–6), 1394–1415.

_ , "Schools' Mental Health Services and Young Children's Emotions, Behavior, and Learning," *Journal of Policy Analysis and Management*, 2010, *29* (4), 698–725.

Sunderman, Gail L., Erin Coghlan, and Rick Mintrop, "School Closure as a Strategy to Remedy Low Performance," Technical Report, National Education Policy Center, School of Education, University of Colorado Boulder, Boulder, CO May 2017.

Welsh, Richard O. and Shafiqua Little, "The School Discipline Dilemma: A Comprehensive Review of Disparities and Alternative Approaches," *Review of Educational Research*, 2018, *88* (5), 752–794.

Online Appendix

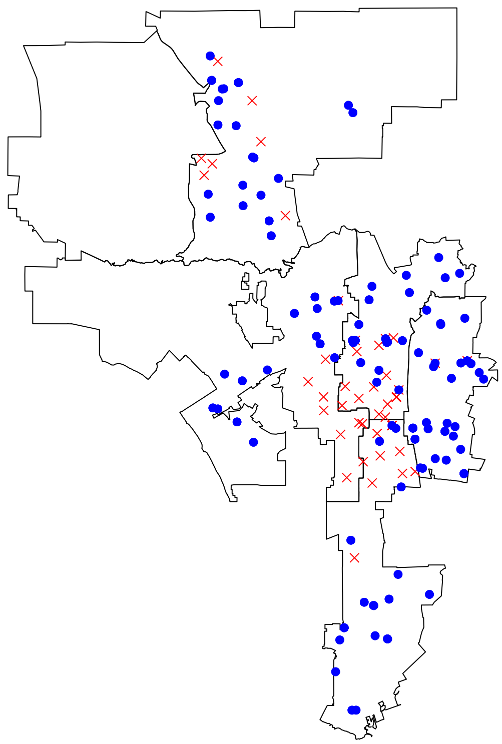
Can We Save Failing Schools? Evidence From Los Angeles

Antonia Vazquez

Table of Contents

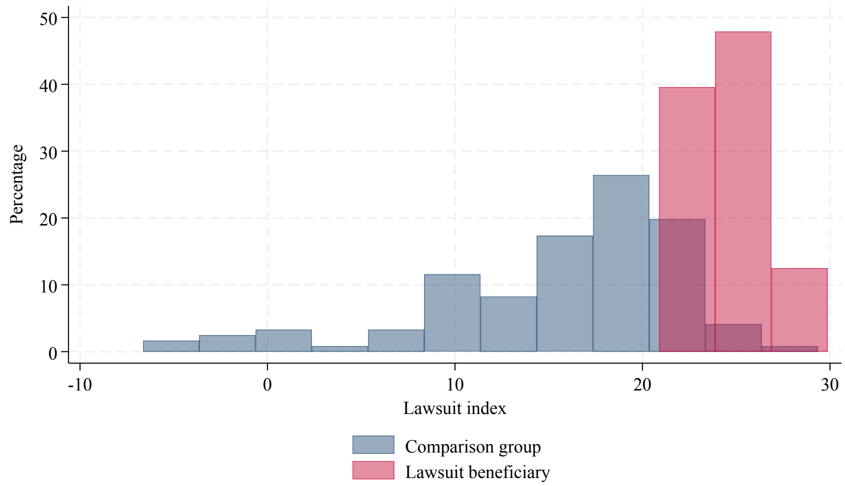
A Appendix Figures and Tables

Figure A.1: Geographic Distribution of the Schools Part of the Analysis Sample



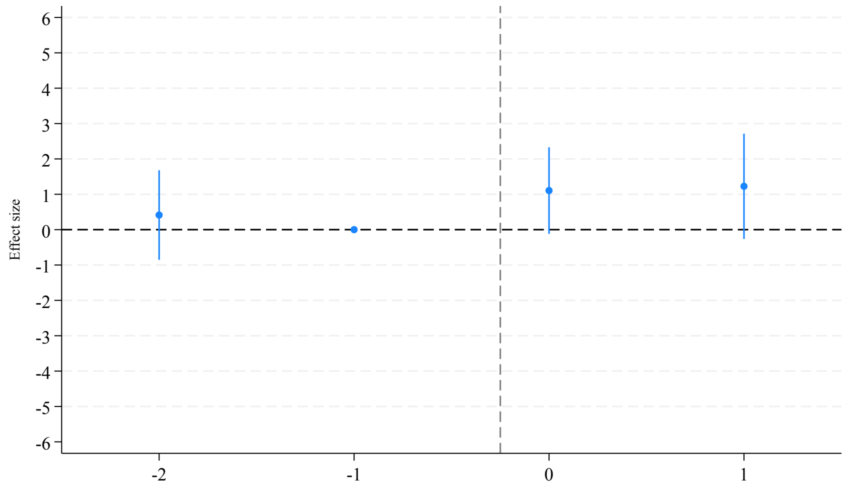
Notes: The map displays the geographic distribution of schools in the analysis sample across LAUSD local districts. Red crosses markers indicate treated schools (beneficiaries of the lawsuit settlement), and blue circles indicate control schools. District boundaries correspond to LAUSD's local district administrative divisions.

Figure A.2: Distribution of the Lawsuit Index by Lawsuit Beneficiary and Comparison Schools

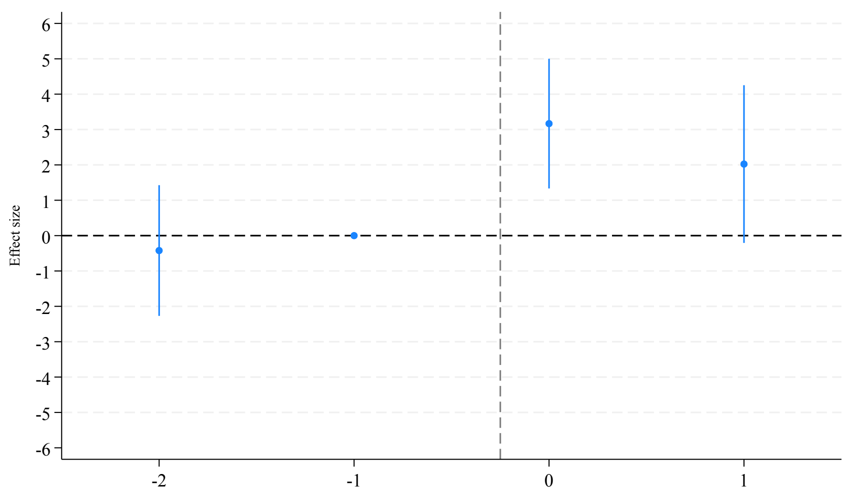


Notes: Each bar represents the percentage of schools within a group falling in a given lawsuit index bin. Navy bars correspond to the comparison group and red bars to lawsuit-beneficiary schools. The lawsuit index is a weighted composite score constructed from 2015/16 school-level data to capture the conditions targeted by the lawsuit settlement. It assigns positive weights to the share of homeless students (5%), the share of foster students (5%), the overall suspension rate (20%), and the share of unduplicated pupils (30%). Academic performance, measured as the share of students meeting or exceeding the math standard, enters with a negative weight (40%), so that lower achievement raises the index. Schools with higher values on the index were more likely to be named as beneficiaries of the lawsuit settlement.

Figure A.3: Effects on Students' Performance Outcomes



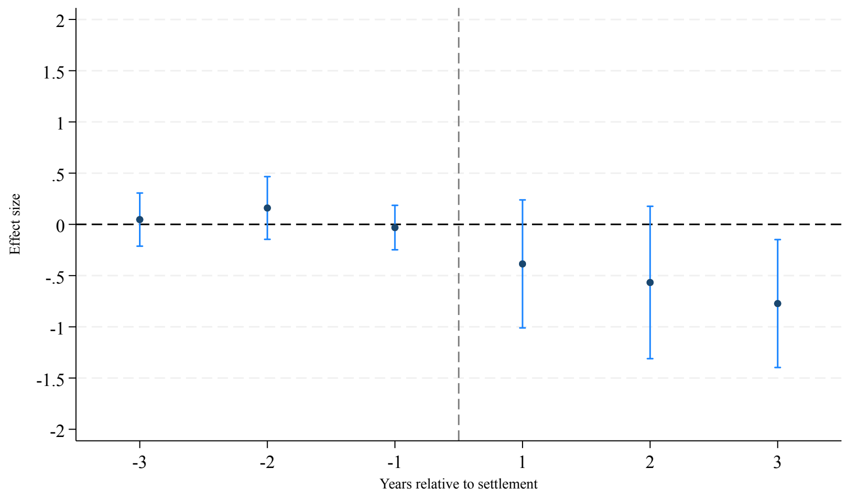
(a) Math



(b) ELA

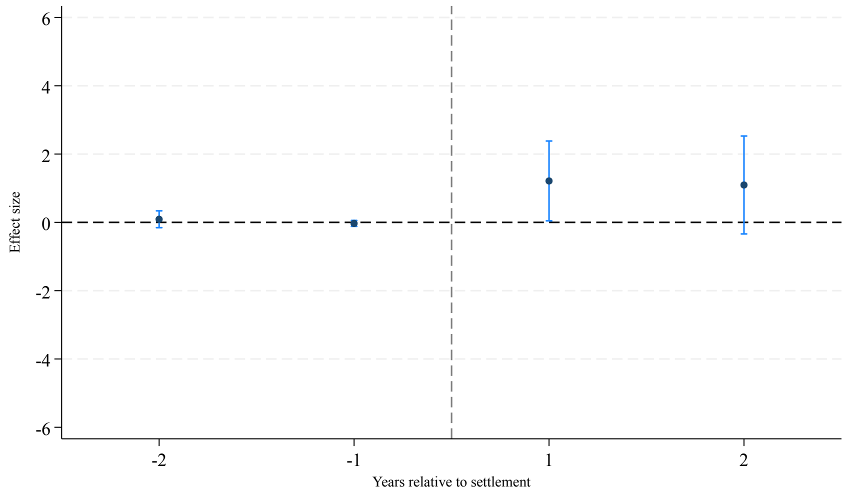
Notes: This figure reports event study estimates of the settlement's effects on students' performance outcomes. The math and English language arts (ELA) outcomes are the percentage of students who met the state standards for the math standardized test (Panel (a)) or the ELA standardized test (Panel (b)). The regressions include school and year fixed effects and cluster standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed line indicates the start of the settlement period.

Figure A.4: Synthetic DiD Event Study: Suspension Rate

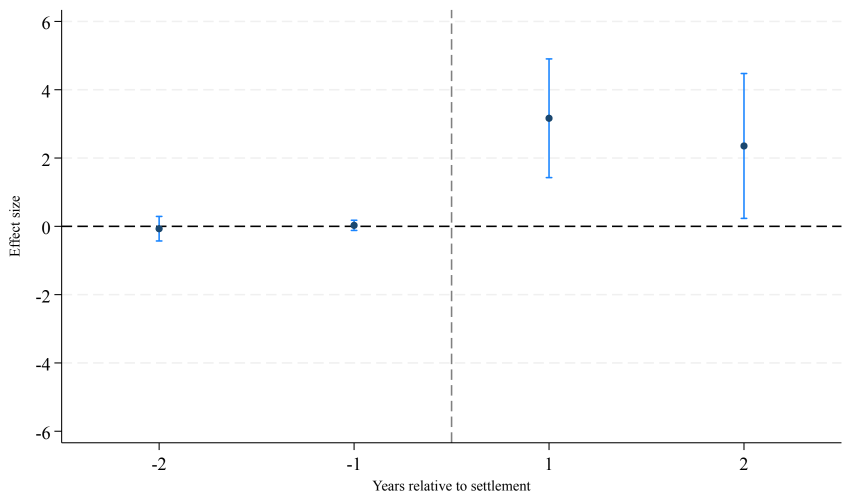


Notes: This figure reports synthetic difference-in-differences event study estimates of the settlement's effects on the student suspension rate, following Arkhangelsky et al. (2021). The suspension rate is measured as the percentage of students suspended at least once during the academic year. Placebo estimates use pre-treatment periods as pseudo-treatment dates. The vertical dashed line indicates the start of the settlement period. Year 0 (academic year 2017/18, the first year of the settlement) is omitted as the reference period. Standard errors are bootstrapped with 1,000 replications.

Figure A.5: Synthetic DiD Event Study: Math and ELA Achievement



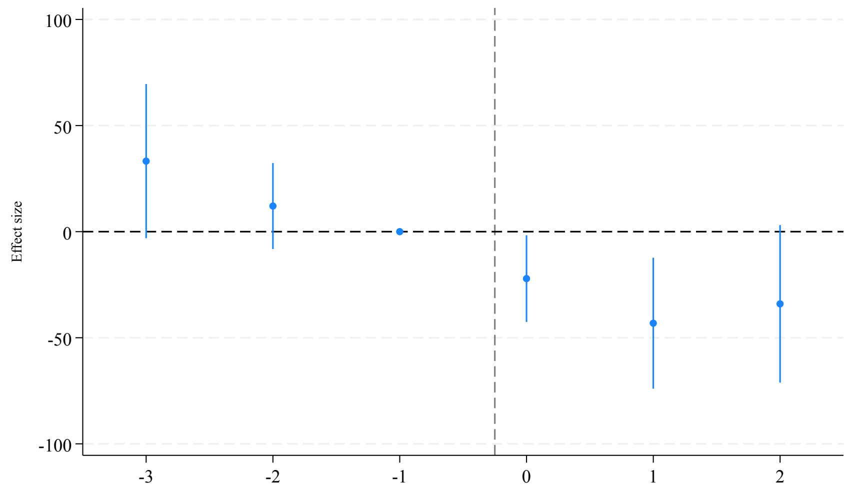
(a) Math



(b) ELA

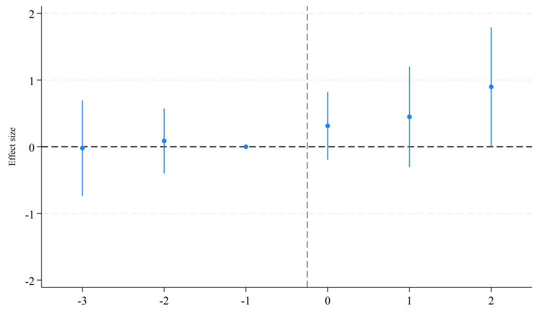
Notes: This figure reports synthetic difference-in-differences event study estimates of the settlement's effects on the percentage of students who met the state standards for the math standardized test (Panel (a)) and the English Language Arts standardized test (Panel (b)), following Arkhangelsky et al. (2021). The sample is restricted to 2016–2019 due to data availability. Placebo estimates use pre-treatment periods as pseudo-treatment dates. The vertical dashed line indicates the start of the settlement period. Year 0 (academic year 2017/18, the first year of the settlement) is omitted as the reference period. Standard errors are bootstrapped with 1,000 replications.

Figure A.6: Effect on Enrollment

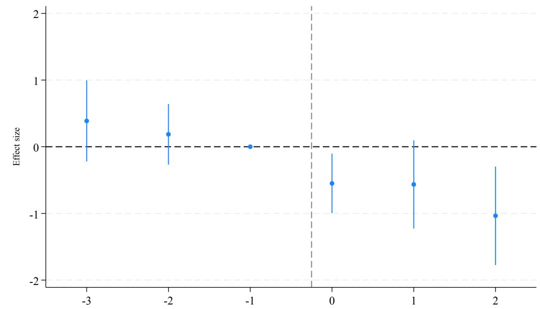


Notes: This figure reports event study estimates of the settlement's effects on school enrollment. The regression includes school and year fixed effects and clusters standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed line indicates the start of the settlement period.

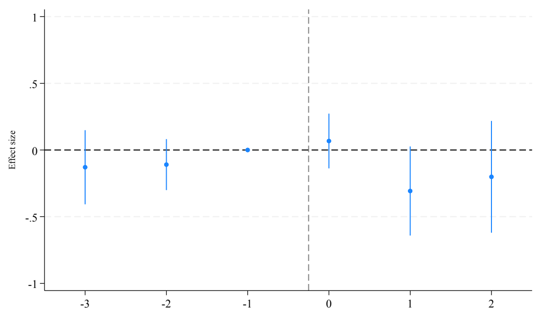
Figure A.7: Effects on Racial Composition



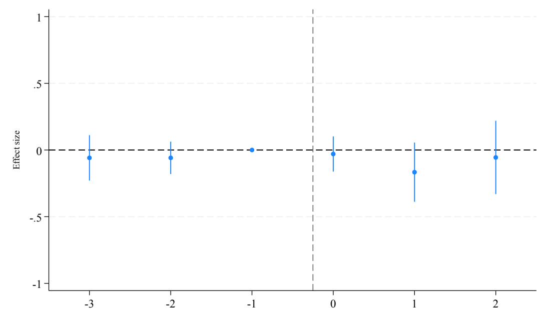
(a) % Hispanic



(b) % Black



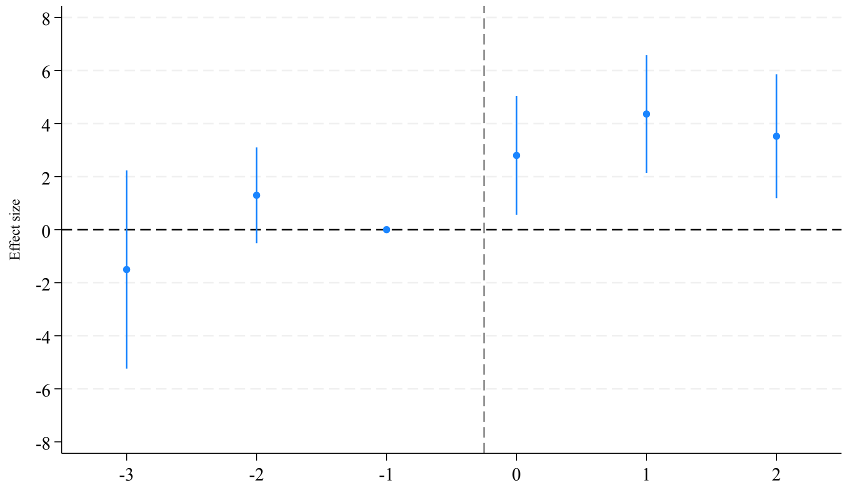
(c) % White



(d) % Asian

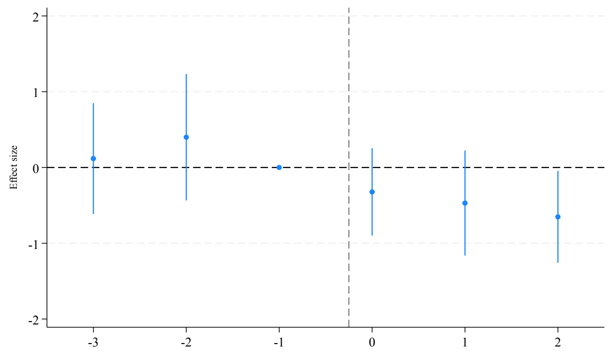
Notes: This figure reports event study estimates of the settlement's effects on schools' racial composition. The regressions include school and year fixed effects and cluster standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed line indicates the start of the settlement period.

Figure A.8: Effects on the Percentage of Students Eligible for Free Lunch

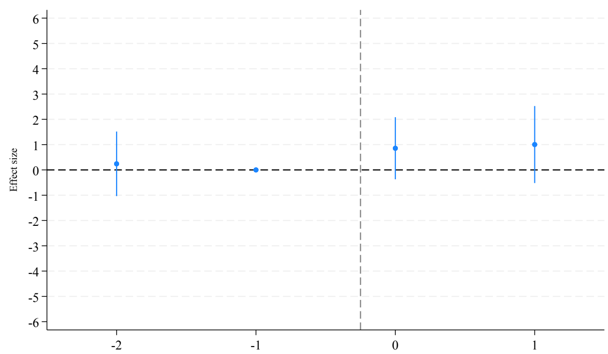


Notes: This figure reports event study estimates of the settlement's effects on the percentage of students eligible for free lunch. The regressions include school and year fixed effects and cluster standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed line indicates the start of the settlement period.

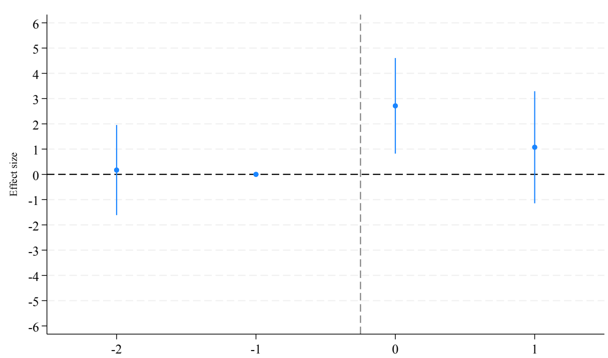
Figure A.9: Effects on Students' Outcomes Using a Further Away Comparison Group



(a) Suspension rate



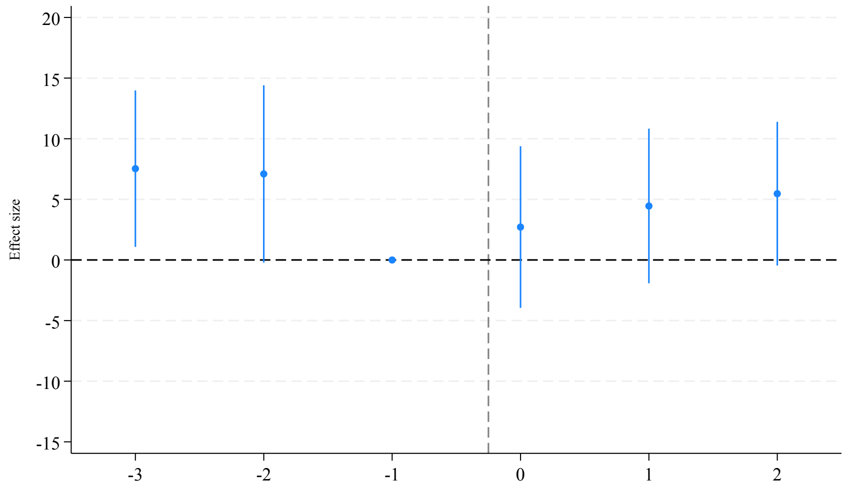
(b) Math



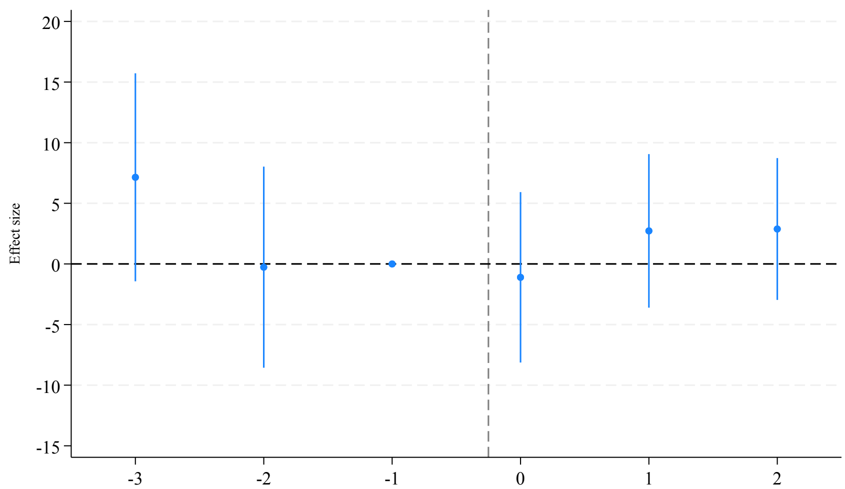
(c) ELA

Notes: This figure reports event study estimates of the settlement's effects on students' outcomes. Suspension rate is measured as the percentage of students suspended at least once during the academic year (Panel (a)). The math and English language arts (ELA) outcomes are the percentage of students who met the state standards for the math standardized test (Panel (b)) or for the ELA standardized test (Panel (c)). The regressions include school and year fixed effects and cluster standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed line indicates the start of the settlement period.

Figure A.10: Effects on Survey Response Rates



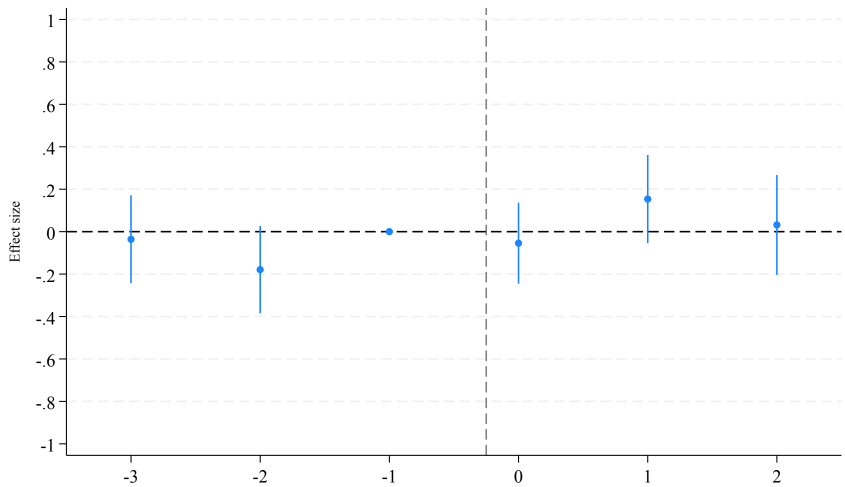
(a) Student response rate



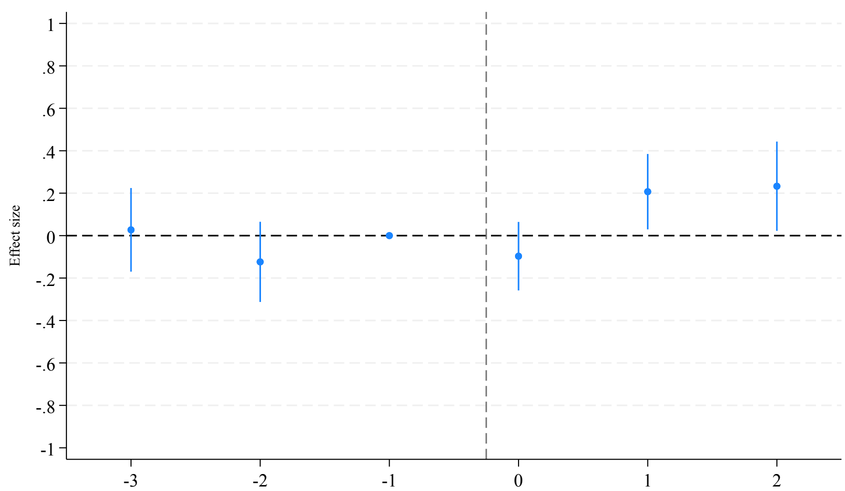
(b) Staff response rate

Notes: This figure reports event study estimates of the settlement's effects on survey response rates for students (Panel a) and staff (Panel b). The regressions include school and year fixed effects and cluster standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed line indicates the start of the settlement period. Response rates are measured as the percentage of enrolled students or employed staff who completed the survey.

Figure A.11: Effects on Students' School Experience Indices



(a) Bullying index



(b) School climate index

Notes: This figure reports event study estimates of the settlement's effects on survey outcomes about students' school experience. The regressions include school and year fixed effects and cluster standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed line indicates the start of the settlement period. I construct the survey indices following Anderson (2008) and normalize them to have mean zero and standard deviation one. The bullying index in Panel (a) combines four survey questions about students' experiences with physical aggression (being pushed, shoved, slapped, hit, or kicked), relational aggression (being the subject of mean rumors or lies), sexual harassment (experiencing sexual jokes, comments, or gestures), and appearance-based teasing. The school climate index in Panel (b) includes six survey questions on adult respect and help, feelings of safety and belonging at the school, and overall satisfaction with their school experience.

Table A.1: School Characteristics Across Different Groups

	CA (excluding LA)	LA (excluding LAUSD)	LAUSD (excluding study sample)	Non-Beneficiary	Beneficiary
N of students enrolled	922.53	1099.58	708.42	1162.02	811.71
	(763.08)	(848.77)	(712.99)	(678.81)	(418.28)
% free lunch	52.28	57.75	80.64	83.65	87.48
	(26.77)	(28.48)	(18.69)	(7.89)	(5.53)
% english learner	13.69	13.41	15.40	15.14	26.49
	(12.82)	(10.12)	(9.00)	(8.48)	(6.88)
% Homeless	2.62	3.84	1.38	3.02	4.20
	(4.23)	(5.53)	(2.14)	(1.99)	(1.92)
% Foster	0.60	1.20	0.68	0.55	1.27
	(2.35)	(4.82)	(1.12)	(0.41)	(1.04)
N Schools	2429	404	192	121	49

Notes: Each column reports the mean with standard deviation in parentheses. The sample includes all public high and middle schools in California in academic year 2016/17. *CA* refers to all California schools outside Los Angeles County. *LA* refers to schools in Los Angeles County outside LAUSD. *LAUSD* refers to LAUSD schools not in the study sample. *Non-Beneficiary* and *Beneficiary* refer to the comparison and treated groups in the study sample, respectively. Data source: California Department of Education.

Table A.2: Effects on Main Outcomes Using Log-Transformed Outcomes

	(1) DiD	(2) N
Suspension rate (log)	-0.269***	1020
	(0.083)	
% Math met (log)	0.095	677
	(0.064)	
% ELA met (log)	0.102***	679
	(0.039)	

Notes: This table presents DiD estimates using log-transformed outcomes. For suspension rates, I use ln(suspension rate+1) to accommodate schools with zero suspensions. For math and ELA, I use $\ln(y)$ directly. All regressions include school and year fixed effects. Standard errors are clustered at the school level. Test score regressions are restricted to 2016–2019 due to data availability. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.3: Synthetic DiD Control Unit Weights

School ID	Susp.	Math	ELA	School ID	Susp.	Math	ELA
154	0.020481	0.008460	0.008517	170	0.007788	0.008074	0.008451
58	0.013781	0.007722	0.007873	126	0.007723	0.007871	0.008523
128	0.013631	0.007787	0.008057	123	0.007722	0.007802	0.008580
103	0.013436	0.008211	0.008365	99	0.007715	0.008566	0.008560
166	0.013426	0.008394	0.008215	96	0.007709	0.007592	0.008271
142	0.013244	0.008117	0.008449	65	0.007709	0.008288	0.008705
147	0.012509	0.008235	0.008467	20	0.007708	0.007140	0.008673
1	0.011785	0.008862	0.007688	114	0.007704	0.010102	0.007762
78	0.011317	0.009119	0.007830	8	0.007694	0.008887	0.008429
121	0.011235	0.008213	0.008487	155	0.007691	0.008525	0.008382
174	0.010967	0.008290	0.008313	37	0.007659	0.007728	0.008709
54	0.010864	0.009307	0.008389	113	0.007640	0.009375	0.008285
150	0.010646	0.008281	0.008256	138	0.007639	0.008405	0.008414
66	0.010233	0.008239	0.008728	131	0.007634	0.008385	0.008305
164	0.010088	0.008659	0.008118	117	0.007627	0.008497	0.008088
173	0.010015	0.008319	0.008356	95	0.007626	0.009037	0.008248
136	0.009882	0.007329	0.008737	91	0.007623	0.009237	0.008230
116	0.009839	0.009368	0.008123	90	0.007620	0.007254	0.008460
97	0.009814	0.008551	0.007952	87	0.007620	0.008591	0.008810
41	0.009768	0.008036	0.007438	81	0.007620	0.007857	0.008975
25	0.009649	0.009276	0.008391	70	0.007619	0.008472	0.007815
171	0.009633	0.008563	0.008579	68	0.007619	0.007721	0.007939
32	0.009601	0.005785	0.008368	62	0.007616	0.008958	0.008324
175	0.009596	0.008276	0.008251	51	0.007616	0.006451	0.009380
35	0.009543	0.008621	0.008186	7	0.007614	0.008789	0.008487
22	0.009384	0.009131	0.007922	16	0.007586	0.008417	0.008248
146	0.009233	0.007782	0.008747	43	0.007552	0.008513	0.008682
120	0.009226	0.007907	0.008072	82	0.007451	0.008128	0.008136
152	0.009144	0.007865	0.008274	98	0.007415	0.007713	0.008125
125	0.009125	0.008235	0.008268	24	0.007412	0.007815	0.008227
88	0.008877	0.007823	0.008493	161	0.007361	0.008514	0.008272
162	0.008808	0.008266	0.008546	124	0.007277	0.008112	0.008400
40	0.008798	0.005676	0.009134	56	0.007187	0.007844	0.008586
165	0.008724	0.008537	0.008151	157	0.007151	0.007004	0.008481
107	0.008690	0.008732	0.008200	64	0.007109	0.007830	0.008614
57	0.008645	0.007912	0.008688	144	0.007099	0.008432	0.008263
80	0.008618	0.009930	0.007634	149	0.007082	0.008165	0.008153
169	0.008609	0.008295	0.008695	71	0.006940	0.007185	0.008097
151	0.008573	0.008063	0.008741	27	0.006914	0.008960	0.008111
79	0.008505	0.006375	0.008161	14	0.006856	0.009365	0.007570
119	0.008482	0.008429	0.008175	85	0.006775	0.009825	0.007988
28	0.008452	0.006329	0.007704	160	0.006636	0.009026	0.008400
111	0.008399	0.008444	0.008304	4	0.006622	0.007265	0.007612
129	0.008239	0.008195	0.008361	63	0.006618	0.008476	0.008460
89	0.008238	0.008300	0.008243	42	0.006612	0.009062	0.007599
69	0.008234	0.007756	0.007815	110	0.006612	0.008060	0.008348
21	0.008227	0.007215	0.009287	72	0.006582	0.008210	0.008052
10	0.008219	0.008192	0.007573	5	0.006582	0.007935	0.008119
106	0.008174	0.008610	0.007984	53	0.006342	0.008348	0.008771
156	0.008151	0.008558	0.008266	137	0.006299	0.007925	0.008360
92	0.008088	0.007232	0.008336	148	0.005899	0.008107	0.008215
102	0.008081	0.008952	0.008104	17	0.005808	0.008444	0.008098
93	0.008054	0.009111	0.008587	153	0.005650	0.008439	0.008450
86	0.008041	0.008353	0.008133	34	0.005564	0.007591	0.006380
38	0.007959	0.008662	0.008530	55	0.005358	0.008066	0.007566
105	0.007953	0.008360	0.007978	31	0.004769	0.009727	0.007888
134	0.007937	0.008144	0.008195	3	0.004391	0.008246	0.008246
12	0.007915	0.008607	0.007705	109	0.003776	0.009553	0.008360
118	0.007911	0.008749	0.008592	122	0.002384	0.008424	0.008126
172	0.007848	0.008095	0.008398	141	0.002384	0.008096	0.008419
112	0.007846	0.008734	0.007646				

Notes: This table reports the control unit weights assigned by the SDID estimator for all comparison schools across the three main outcomes. Weights sum to 1. All control schools receive positive weight, indicating that SDID estimates do not rely heavily on any single comparison school.

Table A.4: Effects on Students' Outcomes Using a Further Away Comparison Group

	(1) Treated group mean (2017)	(2) DiD	(3) N
Suspension rate	1.705	-0.654*	738
		(0.334)	
% Math met	7.475	0.810	491
		(0.609)	
% ELA met	21.897	1.809**	491
		(0.892)	

Notes: Suspension rate is measured as the percentage of students suspended at least once during the academic year. The math and ELA outcomes are the percentage of students who met the state standards for the math standardized test or for the English Language Arts standardized test. Column (1) shows the treated group mean in the academic year 2016/17 (N=49). Column (2) shows DiD estimates from regressions that include school and year fixed effects. Standard errors are clustered at the school level. The sample excludes comparison schools located within 2 miles of any lawsuit-beneficiary school at the same school level (middle or high school) to address potential spillover effects. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.5: Effects on Students' School Experience Indices

	(1)	N
Bullying index	0.115	960
	(0.077)	
School climate index	0.147**	960
	(0.071)	

Notes: All regressions include school and year fixed effects. Standard errors are clustered at the school level. I construct the survey indices following Anderson (2008) and normalize them to have mean zero and standard deviation one. The bullying index in Panel (a) combines four survey questions about students' experiences with physical aggression (being pushed, shoved, slapped, hit, or kicked), relational aggression (being the subject of mean rumors or lies), sexual harassment (experiencing sexual jokes, comments, or gestures), and appearance-based teasing. The school climate index in Panel (b) includes six survey questions on adult respect and help, feelings of safety and belonging at the school, and overall satisfaction with their school experience. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.6: Effects on Individual Components of Student Survey Indices

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Physical aggression	Mean rumors	Sexual harassment	Appearance teasing	Bullying index		
Panel A: Bullying							
Estimate	0.803	1.123*	0.542	1.376**	0.115		
	(0.556)	(0.574)	(0.604)	(0.586)	(0.077)		
	Adult help	Adult respect	School support	Happy	Belonging	Safe	Climate index
Panel B: School Climate							
Estimate	0.801	1.416**	2.283***	1.603***	1.325**	1.178*	0.147**
	(0.744)	(0.629)	(0.802)	(0.620)	(0.522)	(0.614)	(0.071)

Notes: The number of observations is 960. All regressions include school and year fixed effects. Standard errors are clustered at the school level. Panel A reports effects on individual bullying components: physical aggression (being pushed, shoved, slapped, hit, or kicked), relational aggression (mean rumors or lies), sexual harassment (sexual jokes, comments, or gestures), and appearance-based teasing, along with the summary bullying index. Panel B reports effects on individual school climate components: adult help, adult respect, school support, happiness, belonging, and safety, along with the summary climate index. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.7: Effects on Individual Components of Staff Survey Indices

	(1)	(2)	(3)	(4)	(5)
	PD on bullying	Address bullying	Handling bullying index		
Panel A: Handling Bullying					
Estimate	4.095***	4.789***	0.381***		
	(1.504)	(1.315)	(0.113)		
	Bullying	Physical fighting	Discrimination	Bullying index	
Panel B: Bullying					
Estimate	2.683*	6.458***	4.880***	0.260***	
	(1.373)	(1.758)	(1.549)	(0.082)	
	Fair discipline	Effective discipline	Discipline index		
Panel C: Discipline					
Estimate	5.436***	3.920**	0.312***		
	(1.640)	(1.581)	(0.106)		
	Safe	Adult respect	Disruptive behavior	Student disrespect	School Climate index
Panel D: School Climate					
Estimate	7.049***	4.689***	2.372**	3.246**	0.370***
	(1.627)	(1.400)	(1.152)	(1.502)	(0.092)

Notes: The number of observations is 912. All regressions include school and year fixed effects. Standard errors are clustered at the school level. Panel A reports effects on individual components of the handling bullying index: professional development on bullying and addressing bullying. Panel B reports effects on individual components of the bullying perceptions index: bullying, physical fighting, and discrimination among students. Panel C reports effects on individual components of the discipline index: fairness and effectiveness of discipline. Panel D reports effects on individual components of the school climate index: safety, adult respect, disruptive behavior, and student disrespect. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.8: Effects on Main Outcomes Using Continuous Treatment (Funding Intensity)

	(1) Treated group mean (2017)	(2) Continuous DiD	(3) N
Suspension rate	1.705	-0.044**	1020
		(0.021)	
% Math met	7.475	0.072	679
		(0.045)	
% ELA met	21.897	0.206***	679
		(0.065)	

Notes: This table replicates the main specifications replacing the binary treatment indicator with a continuous measure of funding intensity, defined as the settlement allocation in year 1 as a percentage of each school's baseline budget (mean 13.5%, min 7.4%, max 17.8%). The coefficient on the continuous treatment interaction represents the effect of a 1 percentage point increase in funding intensity on each outcome. Column (1) shows the treated group mean in the academic year 2016/17. Column (2) shows DiD estimates from regressions that include school and year fixed effects. Standard errors are clustered at the school level. Test score regressions are restricted to 2016–2019 due to data availability. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

B Settlement Implementation Requirements

Authorized Expenditure Categories. Schools could spend additional funds on six types of services for high-need students: (1) significant increased investments in high-need students, including academic support and mental health and social-emotional support; (2) increasing A-G and Advanced Placement course access for high school students; (3) Linked Learning (an approach that incorporates rigorous academics, career and technical education, work-based learning, and student supports); (4) school climate initiatives, including restorative justice; (5) high school graduation and student recovery for drop-out prevention; and (6) parent and community engagement.

Tiered Support Structure. School tiers were determined by the total funding amount each school received, calculated using the school's duplicated student count and roughly corresponding to enrollment size. These tiered school climate bundles, known as Achievement through Support Teams (ATS), provide wellness, restorative, child welfare and attendance, dropout prevention, intervention, and recovery, and trauma-informed supports to schools.

Tier 1 schools needed to hire one Restorative Justice Teacher Advisor. Tier 2-3 schools needed either a Pupil Services and Attendance Counselor or a Psychiatric Social Worker plus a Restorative Justice advisor (with schools choosing the Psychiatric Social Worker also receiving Resilience Classroom Curriculum). Tier 4 schools required all three positions (PSA Counselor, Psychiatric Social Worker, and Restorative Justice advisor) plus the mandatory Resilience Classroom Curriculum. The Resilience Classroom Curriculum includes resiliency screenings, trauma-informed teacher professional development, and after-school family resiliency education.

C Survey Items Used in Analysis

This appendix lists the survey questions used to construct the outcome indices analyzed in this study.

C.1 Student Survey Questions

Bullying Experiences Index

How many times have other students from your school... (0 Times, 1 Time, 2 or 3 Times, 4 or More Times)

- Pushed, shoved, slapped, hit, or kicked you when they weren't just kidding around?
- Had mean rumors or lies been spread about you?
- Had sexual jokes, comments, or gestures made to you?
- Been teased about what your body looks like?

School Climate Index

How strongly do you agree or disagree with the following statement about your school?

- I am happy to be at this school
- I feel like I am part of this school
- Adults at this school treat all students with respect
- I feel safe in this school
- If I told a teacher or other adult at this school that another student was bullying me, they would try to help me
- This school is a supportive and inviting place for students to learn

C.2 Staff Survey Questions

Bullying Perceptions Index

How much of a problem at this school is...

- Harassment or bullying among students?
- Physical fighting between students?
- Racial/ethnic conflict among students?

Professional Development Index

Please indicate how much you agree or disagree with the following statement about this school

- I have received professional development on preventing bullying
- I address bullying that occurs in my school

Disciplinary Effectiveness Index

Please indicate how much you agree or disagree with the following statement about this school

- This school handles discipline problems fairly
- This school effectively handles student discipline and behavioral problems

School Climate Index

Please indicate how much you agree or disagree with the following statement about this school

- I feel safe on school grounds during the day
- Adults at this school treat all students with respect

How much of a problem at this school is...

- Disruptive student behavior
- Lack of respect for staff by students