Can We Save Failing Schools? Evidence From Los Angeles*

Antonia Vazquez

December 21, 2025

Click here for the most recent version of this paper

Abstract

Can investing in failing schools help them improve? This paper studies this question using a natural experiment based on a 2017 lawsuit settlement that allocated substantial resources 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 of 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, all effects statistically significant. Settlement schools achieved a 0.75 percentage points reduction in suspension rate relative to unfunded schools, reflecting a marked improvement in performance 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.

*LBJ School of Public Affairs, University of Texas–Austin, USA; antonia.vazquez@utexas.edu. I am grateful to my dissertation committee members— Mary Evans, Raissa Fabregas, Andrew Waxman, and Manuela Angelucci— for their insightful guidance and support throughout this work. I especially thank Eric Chyn, Christopher Campos, and Scott Carrell for their substantial comments and feedback on this paper. I also thank Emileigh Harrison, Yumin Hong, Paula López Villalba, Richard Murphy, Francisco Pardo, Ljubica Ristovska, and seminar participants at the LBJ School Sack Lunch Seminar for helpful comments and suggestions.

1 Introduction

A longstanding debate has centered on how to address poor school performance: whether to close or invest more in low-performing schools. Supporters of closures argue that chronically underperforming schools cannot be fixed and that closing them redirects scarce resources toward more effective alternatives. This view has influenced policy considerably; between 2000 and 2022, U.S. districts closed approximately 1,450 schools each year on average.[1] In contrast, advocates for investment contend that low-performing schools fail due to systematic underresourcing and that substantial, sustained funding can turn them around.

While recent research demonstrates that increased school spending improves student outcomes (Jackson and Mackevicius, 2024), this evidence derives mainly from broad-based finance reforms or district-level policies affecting typical schools rather than targeted interventions in chronically low-performing schools. These schools face distinct challenges: low-income districts face funding gaps at more than twice the rate of higher-income districts,[2] chronically low-performing schools experience enrollment declines at more than twice the rate of other public schools (Goulas, 2024), and they suffer worse school climates, including higher rates of student discipline problems.[3] Moreover, existing research has focused on cognitive outcomes while overlooking other dimensions of school quality, such as school climate, teacher effectiveness, and counseling support, which may be particularly important for low-performing schools and are strongly correlated with overall school performance (Carrell et al., 2025).

This paper addresses these gaps by examining how investment affects the lowest-performing schools. I study a natural experiment arising from a 2017 lawsuit settlement in the Los Angeles Unified School District (LAUSD) that delivered substantial resources to the district schools struggling the most. Specifically, fifty 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. The settlement required treated schools to improve or provide more services for high-need students broadly defined, hire additional staff to improve school climate, and invest in teachers' professional development. Using a difference-in-differences (DiD) design, I compare outcomes between the fifty funded secondary schools and nearby public, noncharter secondary schools that did not receive funding. To interpret the results causally, this

[1] National Center for Education Statistics, "Fast Facts: School Closures," available online.

[2] Halpin, John, James Morton, Rebecca Goldstein, and Erin Russ. Closing America's Education Funding Gaps." The Century Foundation, July 28, 2020, available online.

[3] "Low-Performing Schools." Education Week, September 2004, available online.

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 councelors 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, I observe only two pretreatment periods for academic outcomes, which cautions against a causal interpretation of these estimates.

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 5.7% decline from baseline), experienced a 0.9 pp decrease in Black enrollment (a 6.2% 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 following Horowitz and Manski (2000) that assesses how sensitive my estimates are to different assumptions about the characteristics of students who left. I consider various scenarios regarding these students' suspension propensity and find that, across all reasonable bounding assumptions, the sign and magnitude of the results remain substantively unchanged and statistically significant. This suggests that the observed improvements in the disciplinary climate are not simply a mechanical result of the altered student composition but instead reflect genuine changes in school practices.

To understand whether the observed reduction in suspension rates reflects behavioral improvements, shifts in educator responses, or changes in administrative practices, 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 (0.36 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.

This research contributes to several strands of literature. First, it addresses a gap in the school finance literature by examining whether increased spending can improve outcomes in the lowest-performing schools. While recent research demonstrates that increased school investment improves student outcomes on average (Jackson and Mackevicius, 2024), this evidence derives mainly from general finance reforms or district-level policies rather than targeted interventions in failing schools. Moreover, 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 (Carrell et al., 2025). Together, these limitations make it unclear whether the positive effects documented in the literature extend to the most underperforming institutions across all relevant dimensions of school quality.

Existing research on school spending has documented improvements in test scores, graduation rates, college attendance, and adult earnings, with particularly strong effects for low-income students (Abott et al., 2020, Baron, 2022, Candelaria and Shores, 2019, Hyman, 2017, Jackson et al., 2016, Lafortune et al., 2018). Yet these studies examine broad expenditure increases from school finance reforms or close elections that affected diverse school populations, and they focus almost exclusively on test scores and longer-term academic outcomes. I directly examine a reform targeting the lowest-performing schools and expand the analysis beyond cognitive outcomes to include school climate and disciplinary outcomes, dimensions that may be particularly important in failing schools. This distinction matters because these schools may respond differently to additional resources due to more severe baseline challenges or distinct intervention needs, and improvements in school environment may be a critical pathway for long-term success.

Beyond targeting, the composition of spending matters for effectiveness. Researchers have found larger effects from operational spending on personnel than from capital expenditures (Abott et al., 2020, Baron, 2022, Biasi, 2024), suggesting that the specific nature of investments and institutional context are crucial. I provide direct evidence on whether investment with targeted and discretionary components can improve outcomes in failing schools. Crucially, I examine both disciplinary outcomes and the mechanisms through which resources operate in these challenging contexts. This analysis is essential because demonstrating that "money matters" in general may not suffice to conclude that investment can help the schools that need it most.

The closest prior evidence on interventions to improve failing 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 changes such as replacement of principals and at least 50% of the schools' staff. While recent meta-analytic evidence finds these interventions generated moderate positive gains in math achievement (Schueler et al., 2022), they differed fundamentally from resource-based investment in that they were highly prescriptive and disruptive. The Los Angeles settlement preserved school autonomy, providing substantial funding while allowing schools considerable discretion in addressing their challenges. This distinction is important because local actors may better understand their specific problems and have greater buy-in for solutions they help design. My analysis addresses a gap by testing whether this gentler, resource-focused approach can achieve improvements without the disruption of imposed turnaround models.

This paper also contributes to the school accountability literature by examining whether investment-

based approaches ("carrots") can improve low-performing schools, contrasting with the punishment-based mechanisms ("sticks") that have dominated accountability policy. While extensive research documents how accountability pressure generates positive responses, it also produces problematic unintended consequences, including teaching to the test, strategic student exclusion, and various gaming behaviors (Booher-Jennings, 2005, Cullen and Reback, 2006, Dee and Jacob, 2011, Figlio, 2006, Figlio and Rouse, 2006, Figlio and Getzler, 2006, Jacob, 2005, Jacob and Levitt, 2003, Neal and Schanzenbach, 2010, Reback, 2008, Reback et al., 2014). Less is known about whether positive incentives can achieve improvements without these negative side effects. The Los Angeles settlement provided schools substantial resources to help them improve rather than threatening them with sanctions, offering insights into whether "carrots" can succeed where "sticks" have yielded mixed results.

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.[4] 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.[5] 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 socioeconomi-

[4] Los Angeles Unified School District, "2024 Fingertip Facts," available online.
[5] Public Policy Institute of California, "Targeted K-12 Funding and Student Outcomes," available online.

cally disadvantaged, and 2.7% were foster or homeless youth.[6]

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 receiving *supplemental* grants when these students are present, and *concentration* grants when they make up more than 55% of enrollment.[7] 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.[8]

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.[9] 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

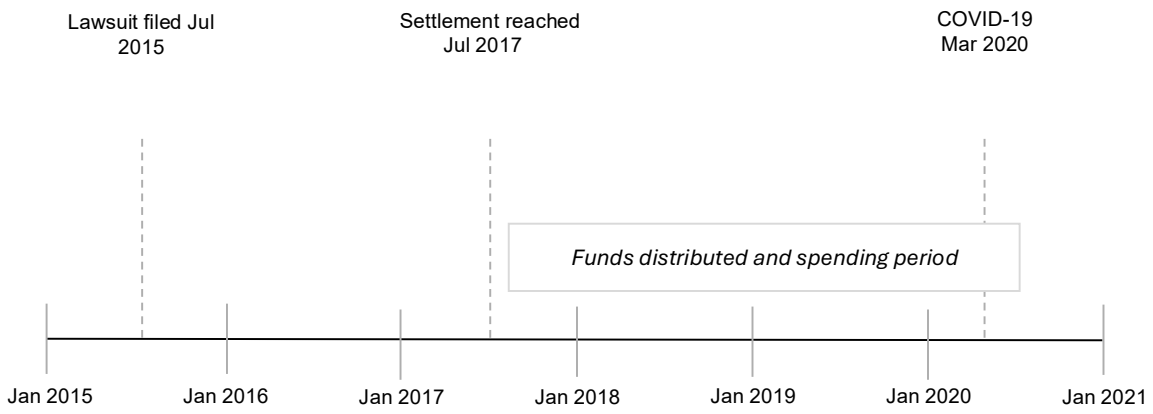
[6] Ed-Data, "Los Angeles Unified School District," available online.

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

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

[9] 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.

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), 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

My 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. These criteria yield a final sample of 171 schools (121 control, 50 treatment) observed over 6 academic years (2014/15 to 2019/20).[10]

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.[11] 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 2017/18.

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.

[10] In my analysis, I have 49 treated schools because one school merged with another after the settlement.

[11] Complete school-level budget data for 2019 and 2020 were not publicly accessible. After submitting multiple unsuccessful requests for these data to the district, I opted to focus on staffing data as the most reliable available indicator of resource allocation changes.

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.

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 or nearly met standards as my measure of academic achievement.[12] 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 achieves 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 achieves 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

[12] 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.

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 Appendix C for details on survey questions.

To analyze these multidimensional survey outcomes systematically, I construct summary indices following the methodology outlined in Anderson (2008). This approach employs 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. For 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. See Appendix C for a list of the questions I used to construct each index.

3.3 Summary statistics

Table 1 compares pretreatment characteristics (measured in 2016/17) between the 50 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 11.2 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 reporte worse school climate (-0.73 versus -0.21, a difference of 0.52 standard deviations).

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:

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

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, Lawsuit_s is an indicator equal to 1 for schools that benefited from the lawsuit, Post_t is an indicator equal to 1 for years 2018 and after (postsettlement), $\text{Lawsuit}_s \times \text{Post}_t$ is the interaction term capturing the treatment effect, δ is the DID estimator,

Table 1: Summary statistics

	Lawsuit beneficiary					Comparison group				
	Mean	SD	Min	Max	N	Mean	SD	Min	Max	N
Demographics										
N of students enrolled	833.59	407.55	288	1961	49	1187.20	682.70	189	2953	121
% Hispanic	82.09	16.16	28	99	49	80.42	19.17	20	99	121
% Black	14.42	16.39	0	69	49	7.51	12.19	0	71	121
% White	1.21	0.89	0	5	49	4.98	7.32	0	35	121
% Asian	0.89	2.18	0	14	49	3.20	5.16	0	26	121
% less than 2% races	1.39	1.39	0	5	49	3.90	6.06	0	45	121
% free lunch	87.41	5.42	74	99	49	83.61	7.85	53	97	121
% english learner	26.35	6.93	9	42	49	15.12	8.49	0	41	121
% Homeless	4.21	1.88	1	10	49	3.03	1.98	0	12	121
% Foster	1.29	1.02	0	4	49	0.56	0.40	0	2	121
Employees										
Certified staff per 1,000 st	63.56	11.90	48	99	49	54.81	6.47	44	77	121
Classified staff per 1,000 st	37.80	14.54	16	83	49	27.58	9.65	6	61	121
Main outcomes										
Suspension rate	1.70	2.11	0	12	49	0.64	1.11	0	9	121
% near or met math standard	28.30	8.89	9	48	49	44.64	8.41	25	77	121
% met ELA standard	21.90	8.42	5	41	49	32.48	7.61	18	50	121
% not meet ELA standard	44.37	16.40	8	78	49	25.65	13.03	0	55	121
% not meet math standard	69.56	9.94	45	91	49	46.88	12.40	2	71	121
Survey indexes outcomes: students										
Bullying experiences	-0.51	1.11	-2.71	1.42	44	-0.42	1.04	-2.85	2.00	116
School climate	-0.73	0.68	-2.01	1.02	44	-0.21	0.94	-1.88	3.95	116
Survey indexes outcomes: staff										
Handling bullying	-0.35	1.06	-1.96	3.09	45	0.43	1.02	-1.86	3.34	107
Bullying among students	-0.49	1.12	-1.97	2.02	45	0.36	0.98	-1.42	2.91	107
Discipline	-0.35	0.82	-1.65	2.70	45	0.53	1.11	-1.41	3.79	107
School climate	-0.57	0.84	-1.91	1.47	45	0.47	1.03	-1.40	3.70	107
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. 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.

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).

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 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.

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 through negotiations between 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. Selection was based on levels of these characteristics, not on whether schools were improving or deteriorating over time. Schools themselves had no role in this process, they were not party to the lawsuit, could not advocate for their inclusion, and learned of their selection only after the settlement was finalized, when they received notification from the district. This means schools could not have strategically altered their behavior in anticipation of receiving funds. Because selection depended on where schools stood at a point in time rather than on their trajectories, there is little reason to expect systematic differences in trends between the two groups absent the intervention.

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:

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

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

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, Lawsuit_s is an indicator equal to 1 for schools that benefited from the lawsuit, $\mathbf{1}_{t=2018+k}$ 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 pretreatment coefficients (δ_{-3} and δ_{-2}) test 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 1% level and represents a 4.9% increase relative to the baseline level of 63 certified staff per 1,000 students in the treated schools. This substantial increase encompassed teachers providing direct instruction, administrative personnel in certified positions, and specialized pupil services staff such as counselors and psychologists, all critical components of school capacity who require specialized credentials. 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. This category includes paraprofessionals, clerical staff, and other support personnel 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.

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

Table 2: Effects on school spending outcomes

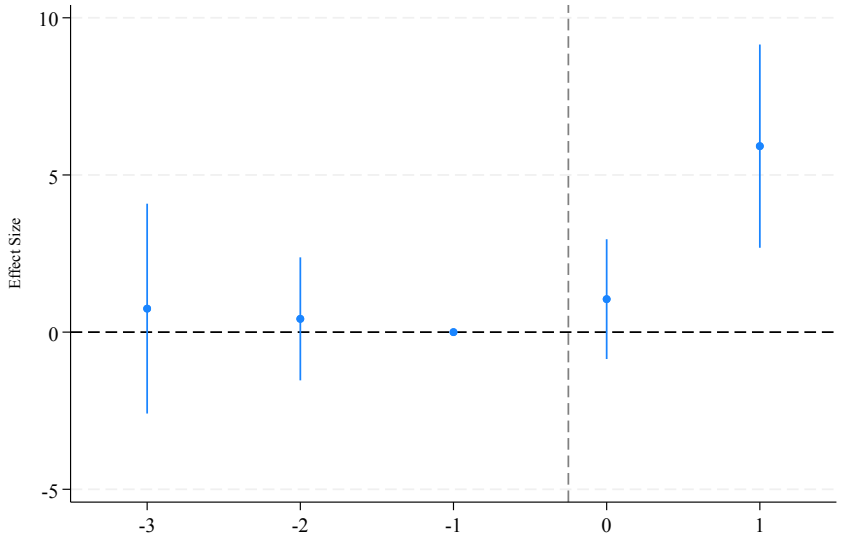
	(1) Treated group mean (2017)	(2) DiD	(3) SDiD	(4) N
Certified staff per 1,000 st	63.564	3.093**	3.410***	850
		(1.199)	(1.119)	
Classified staff per 1,000 st	37.798	1.828*	1.709*	1020
		(0.942)	(0.904)	

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). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

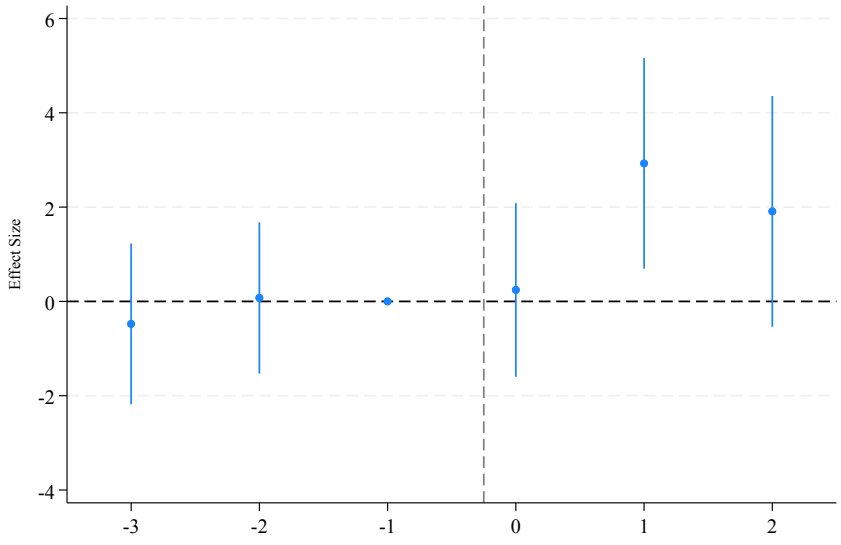
to student behavioral outcomes through two pathways. First, restorative justice interventions have been shown to reduce disciplinary incidents by emphasizing reparative rather than punitive responses to student behavior (Adukia et al., 2025, Augustine et al., 2018). Second, research demonstrates that 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%.

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 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

Figure 2: Effects on spending outcomes



(a) Certified staff



(b) Classified staff

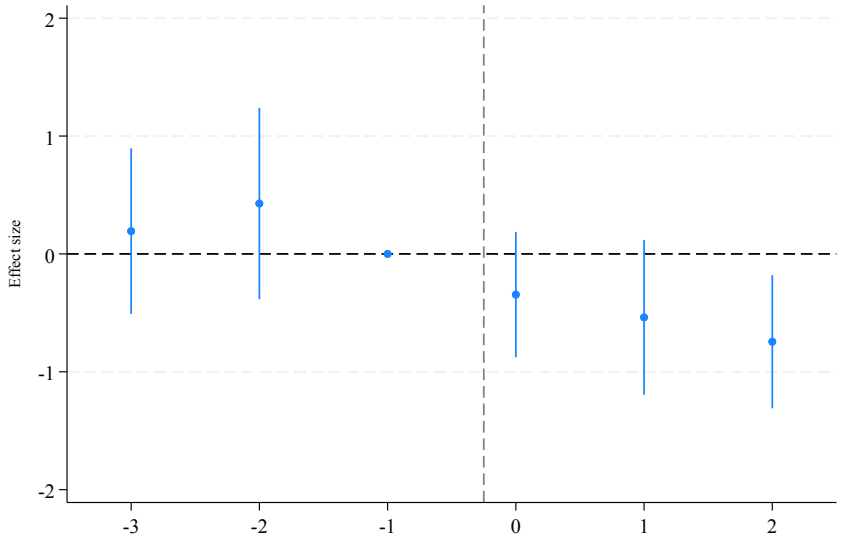
Notes: This figure reports event study estimates of the settlement's effects on school staffing: Panel (a) for certified staff, which includes teachers, administrators, and pupil services personnel requiring credentials, and Panel (b) for classified staff, which includes paraprofessionals, clerical, 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.

Table 3: Effects on students' outcomes

	(1) Treated group mean (2017)	(2) DiD	(3) SDiD	N N
Suspension rate	1.705	-0.750**	-0.575*	1020
		(0.313)	(0.321)	
% Math nearly met or met	28.299	1.512	2.184**	679
		(1.033)	(1.019)	
% 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 and ELA outcomes are the percentage of students who met or nearly met the state standards for the math standardized test, and 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). * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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.

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 5% level; see Table 3, Column (3)). The convergence of estimates across specifications provides strong causal evidence that the settlement meaningfully improved disciplinary climate in treated schools.

Next, I examine effects on academic outcomes, measured by the percentage of students who nearly met or 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.1), 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 nearly met or met standards by 1.5 pp, though this effect is not statistically significant. 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 2.2 pp increase in math proficiency (significant at the 5% 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.

5.3 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.[13] 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 4, A.2, and A.3 display the 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 5.7% 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

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

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.

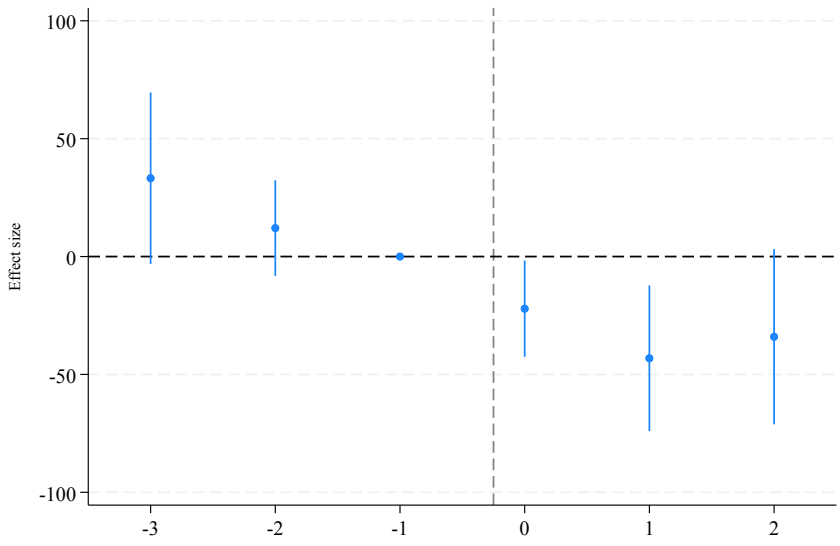
Table 4: Effects on enrollment and student body composition

	(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.

Figure 4: 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.

6 Understanding Suspension and Enrollment: Bounding Evidence

These 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 following 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.[14]

[14] I also considered the more extreme and conservative case in which I assume that the departures were among high-suspension-risk students, with suspension rates twice their demographic baseline (Hispanic students at 2.19% ($2 \times 1.10\%$) and Black students at 9.34% ($2 \times 4.67\%$)), resulting in approximately 2.22 expected suspensions among the 48 departing students per school. Even under this scenario, where the departed students were at the highest risk of suspension, the estimated effect remains negative, -0.485, though not statistically significant.

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.4 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, 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 5 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

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

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 (0.36 SD), consistent with mechanism four. All effects are statistically significant at the 1% level.

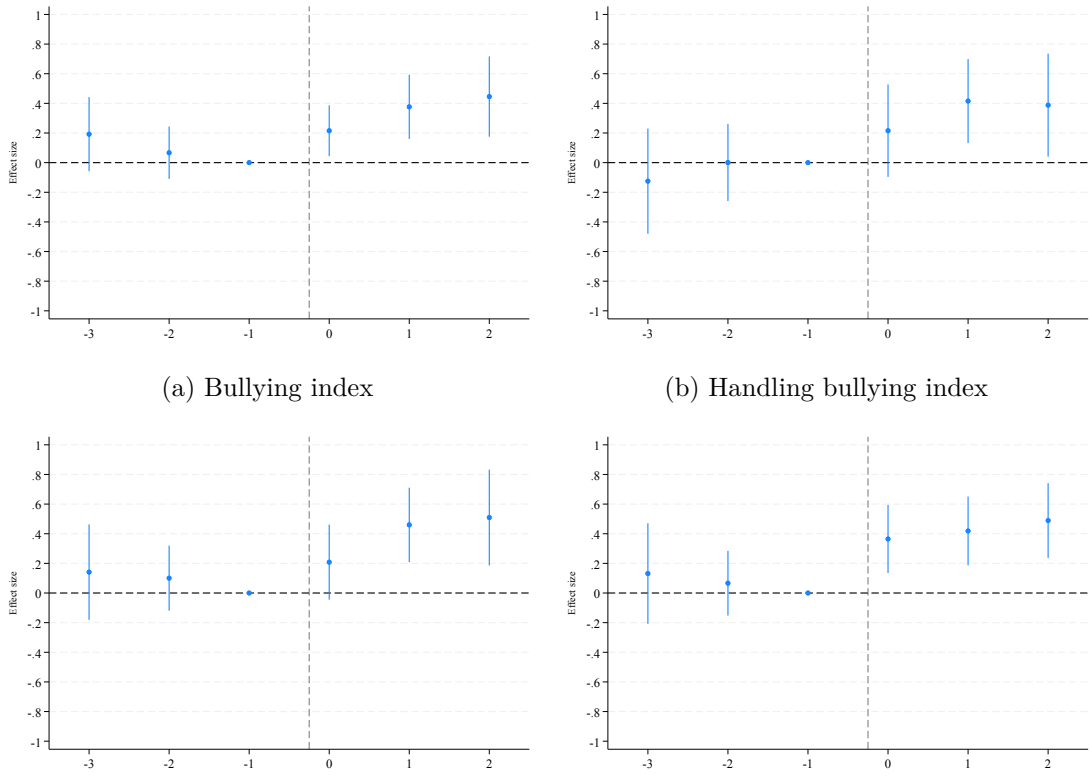
Table 6: Effects on staff's school experience indices

	(1)	N
Bullying index	0.260***	912
	(0.082)	
Handling bullyng index	0.381***	911
	(0.113)	
Discipline index	0.312***	912
	(0.106)	
School climate index	0.358***	912
	(0.095)	

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. 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table A.1 and Figure A.5 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.

Figure 5: 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. 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.

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

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 failing 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%, translating to roughly 852 fewer arrests over the three-year post-intervention period. To quantify the economic implications, I draw on per-offense cost estimates from McCollister et al. (2010), which incorporate victim losses, criminal justice system expenditures, and intangible costs. Because the majority of juvenile arrests involve non-violent offenses, I use a weighted average cost of approximately $17,500 per arrest (in 2024 dollars), yielding estimated social cost savings of roughly $15 million. While this is a back-of-the-envelope calculation that applies elasticities from a different context and relies on assumptions about offense composition, 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.

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.

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, "What Works and For Whom? Effectiveness and Efficiency of School Capital Investments Across The U.S.," 2024.

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

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.

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

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.

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

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.

Goulas, Sofoklis, "Underachieving and Underenrolled: Chronically Low-Performing Schools in the Post-Pandemic Era," Technical Report, Thomas B. Fordham Institute September 2024.

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 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.

_ **, 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.

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.

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.

McCollister, Kathryn E., Michael T. French, and Hai Fang, "The Cost of Crime to Society: New Crime-Specific Estimates for Policy and Program Evaluation," *Drug and Alcohol Dependence*, 2010, *108* (1-2), 98–109.

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.

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.

_ **, Jonah Rockoff, and Heather L. Schwartz**, "Under Pressure: Job Security, Resource Allocation, and Productivity in Schools under No Child Left Behind," *American Economic Journal: Economic Policy*, 2014, *6* (3), 207–241.

Schueler, Beth E., Catherine Armstrong Asher, Katherine E. Larned, Sarah Mehrotra, and Cynthia Pollard, "Improving Low-Performing Schools: A Meta-Analysis of Impact Evaluation Studies," *American Educational Research Journal*, October 2022, *59* (5), 975–1010.

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

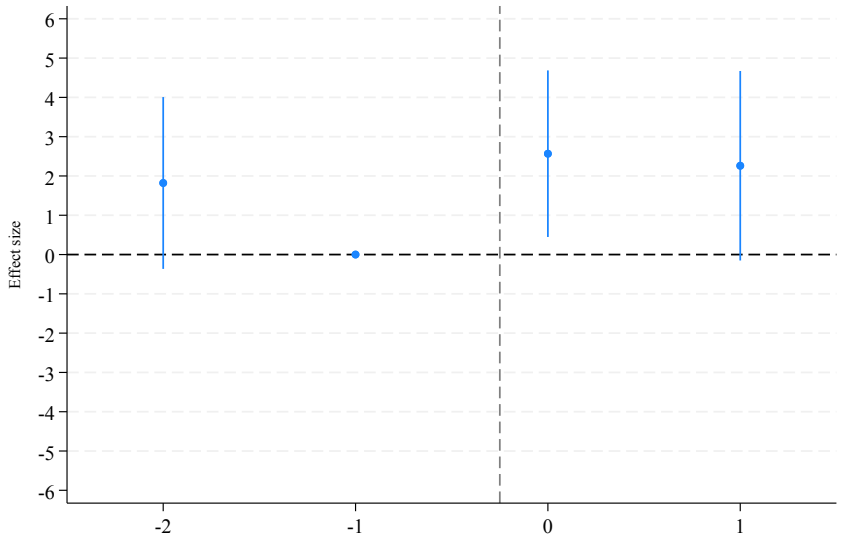
Can We Save Failing Schools? Evidence From Los Angeles

Antonia Vazquez

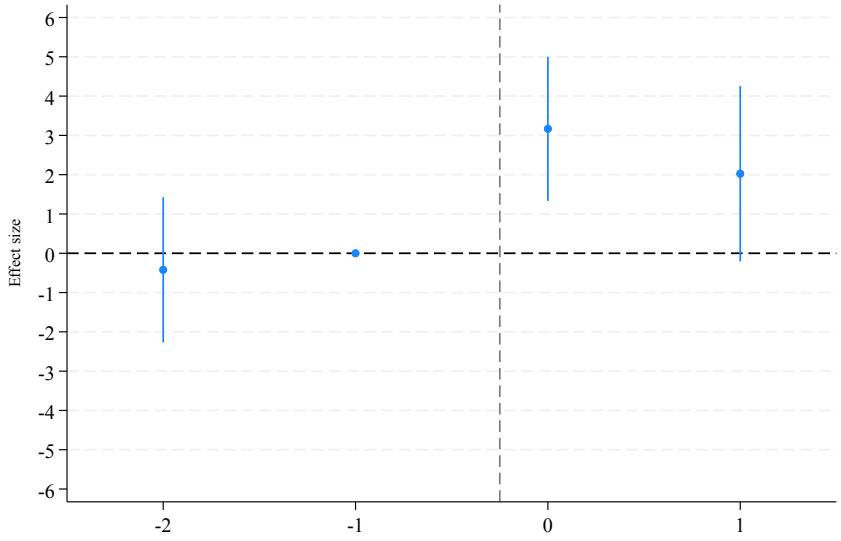
Table of Contents

A Appendix Figures and Tables

Figure A.1: Effects on students' performance outcomes



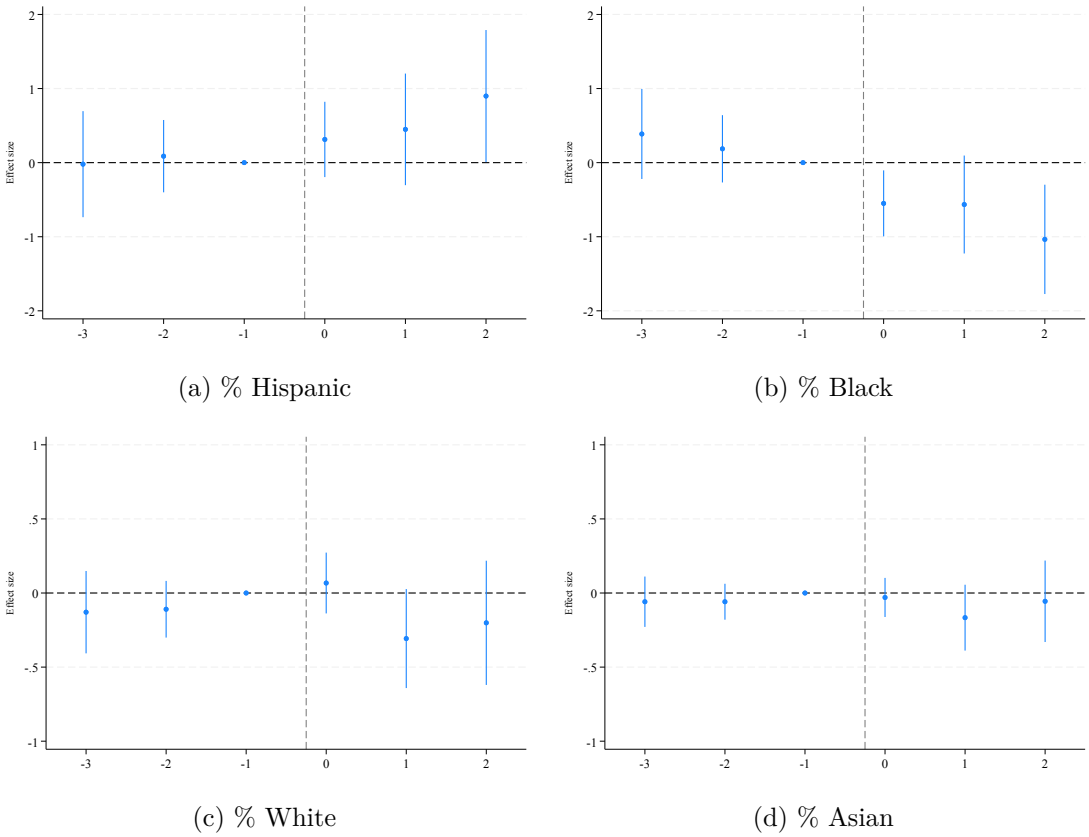
(a) Math



(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 or nearly met the state standards for the math standardized test (Panel (a)) and the percentage of students who met the state standards for the ELA standardized test (Panel (b)), respectively. 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.2: Effects on racial composition



(a) % Hispanic

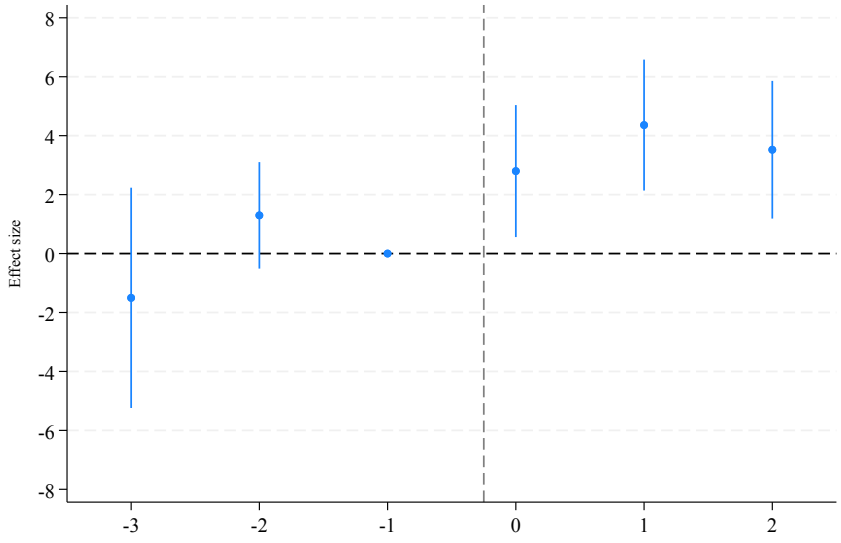
(b) % Black

(c) % White

(d) % Asian

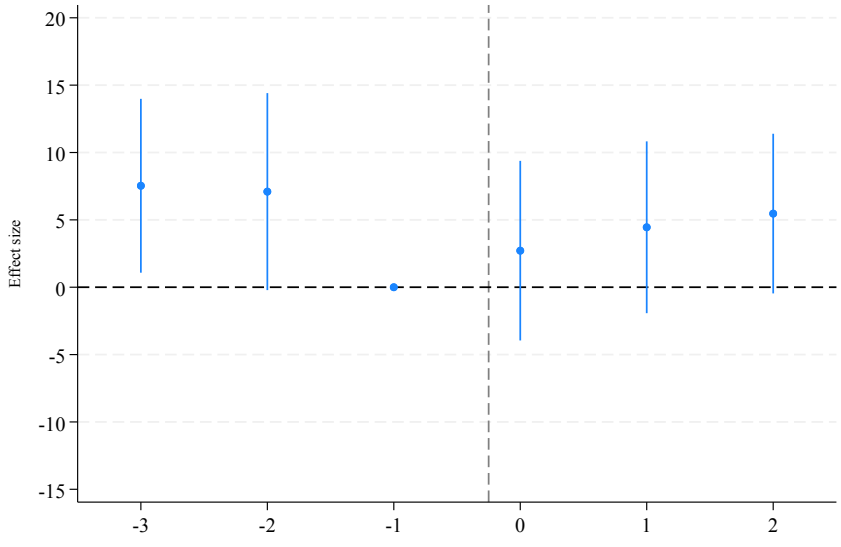
Notes: This figure reports event study estimates of the settlement's effects on schools' racial composition. The regressions include school and year fixed effects and cluster standard errors at the school level. The x-axis shows years relative to the settlement implementation (year 0). The vertical dashed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

Figure A.3: Effects on percentage of students eligible for free lunch

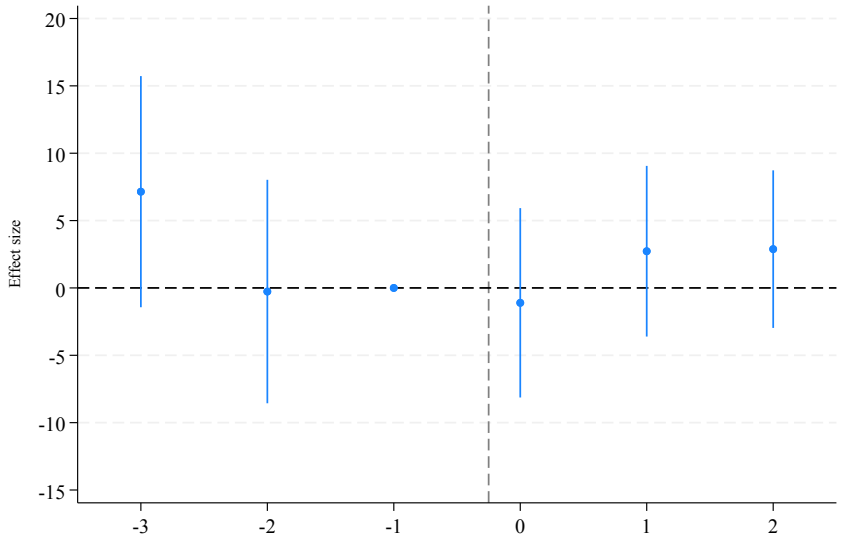


Notes: This figure reports event study estimates of the settlement's effects on the percentage of students eligible for free lunch. The regressions include school and year fixed effects and cluster standard errors at the school level. The x-axis shows years relative to the settlement implementation (year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

Figure A.4: Effects on survey response rates



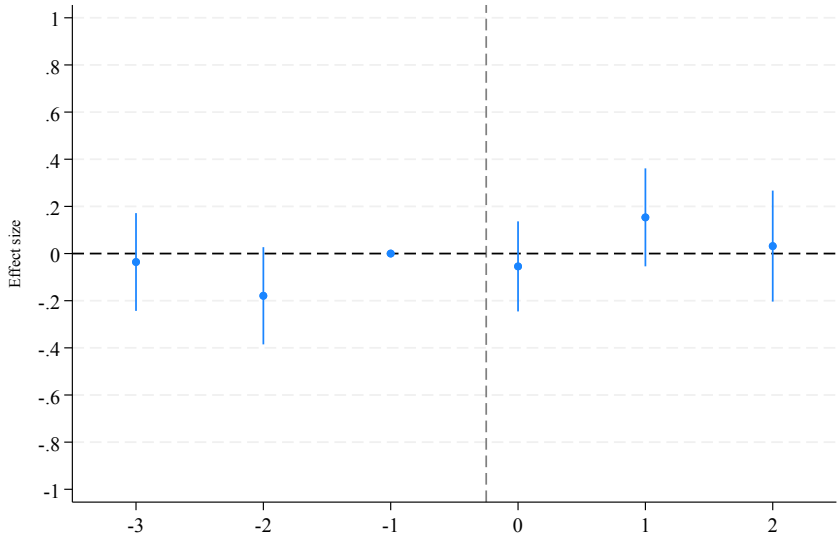
(a) Student response rate



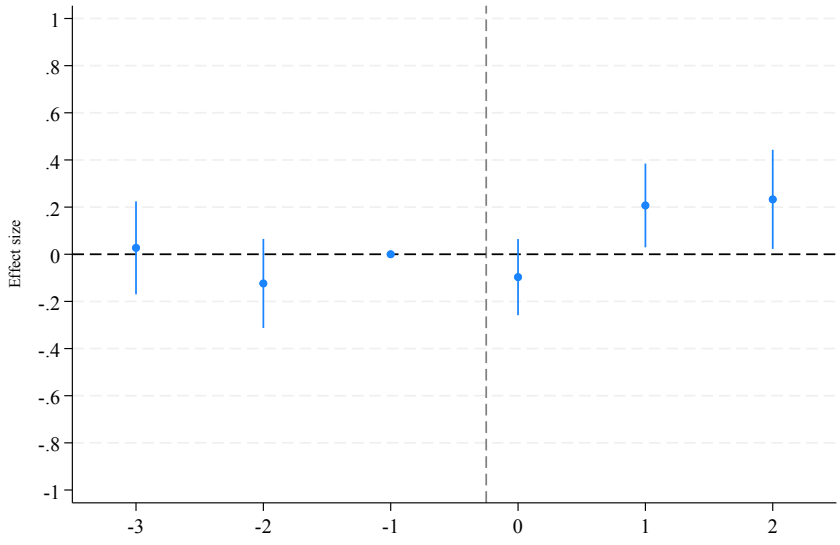
(b) Staff response rate

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.5: 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: 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. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

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