Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 24-1006

# The Impact of Additional Funding on Student Outcomes: Evidence from an Urban District Using Weighted Student Funding and Site-based Budgeting

Christopher A. Candelaria
Vanderbilt University

Angelique N. Crutchfield
Vanderbilt University

Dillon G. McGill
Vanderbilt University

This study uses a concurrent embedded mixed-methods design to assess the impact of additional funding on student outcomes in a large, urban school district in the Southeastern United States. The district implemented student-based budgeting (SBB), which allocates dollars to schools based on student characteristics using a weighted student funding (WSF) formula and provides flexibility to principals to allocate those dollars under site-based budgeting. Using simulated instruments in a difference-in-differences framework, we estimate the impact of additional funding on student outcomes provided by WSF. Student test scores in math and ELA increased by 0.14 and 0.12 standard deviations, respectively. Our qualitative analysis suggests that the flexibility given to principals was a key mechanism that improved student outcomes.

VERSION: August 2024

Suggested citation: Candelaria, Christopher A., Angelique N. Crutchfield, and Dillon G. McGill. (2024). The Impact of Additional Funding on Student Outcomes: Evidence from an Urban District Using Weighted Student Funding and Site-based Budgeting. (EdWorkingPaper: 24-1006). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/kbq5-0b25

# The Impact of Additional Funding on Student Outcomes: Evidence from an Urban District Using Weighted Student Funding and Site-based Budgeting

Christopher A. Candelaria, Angelique N. Crutchfield, and Dillon G. McGill

Vanderbilt University

Draft Date: July 2024

## Author Note

The authors acknowledge generous funding from the William T. Grant Foundation (OR-201088) in supporting their work. Correspondence concerning this article should be addressed to Christopher A. Candelaria https://orcid.org/0000-0002-5903-6214, Department of Leadership, Policy, and Organizations, Peabody College, Vanderbilt University, 230 Appleton Place, PMB 414, Nashville, TN 37203-5721. Email: chris.candelaria@vanderbilt.edu. Phone: (615) 322-0722.

**Abstract**

This study uses a concurrent embedded mixed-methods design to assess the impact of additional funding on student outcomes in a large, urban school district in the Southeastern United States. The district implemented student-based budgeting (SBB), which allocates dollars to schools based on student characteristics using a weighted student funding (WSF) formula and provides flexibility to principals to allocate those dollars under site-based budgeting. Using simulated instruments in a difference-in-differences framework, we estimate the impact of additional funding on student outcomes provided by WSF. Student test scores in math and ELA increased by 0.14 and 0.12 standard deviations, respectively. Our qualitative analysis suggests that the flexibility given to principals was a key mechanism that improved student outcomes.

*Keywords*: educational finance, weighted student funding, site-based budgeting

**The Impact of Additional Funding on Student Outcomes: Evidence from an Urban District Using Weighted Student Funding and Site-based Budgeting**

In the education research landscape, school funding has consistently emerged as a contentious focal point (Jackson & Persico, 2023; McGee, 2023). Historically, the debate has centered on whether school resources were directly correlated with student outcomes. Past studies presented an ambivalent picture, suggesting no clear causal connection between school resources and achievement (Jackson, 2020). However, a resurgence of school finance research now underscores the significance of financial resources on students' academic trajectories.

Recent findings demonstrate that additional school funding translates into improved student outcomes, from test scores and graduation rates in the short term to future earnings and intergenerational mobility in the long term (Biasi, 2023; Candelaria & Shores, 2019; Lafortune, Rothstein, & Schanzenbach, 2018; Jackson, Johnson, & Persico, 2016). These studies primarily focus on spending across districts and the impacts of school finance reforms on student outcomes. Yet, important questions remain unanswered: How are these funds distributed across schools within districts, and how do varying levels of funding at the school level influence student outcomes?

Although school finance reforms have generally increased spending equity across districts within states (Shores et al., 2023), funding allocations from districts to schools may still be inequitable—schools serving more disadvantaged students could receive inadequate per-pupil funding if district resource allocation policies do not align with diverse student needs. Research on sub-district allocation indicates that resources within many districts are not equitably distributed across schools (Knight, 2019; Shores & Ejdemyr, 2017). If district-to-school funding allocations systematically disadvantage schools with higher proportions of marginalized students, this inequity could significantly hinder their educational opportunities and negatively affect student outcomes (Condron & Roscigno, 2003; Houck, 2011; Iatorala & Stiefel, 2003; Rubenstein et al., 2007).

This paper extends existing school finance research by examining how a district-level funding policy known as student-based budgeting (SBB) influences outcomes at the school level in the context of a large, urban district. While SBB is often characterized as a single policy, it consists of two components: weighted student funding (WSF) and site-based budgeting (Levin et al., 2019). WSF is an allocation system that provides a base amount of funding for each student and supplements it with additional, additive funding based on student characteristics chosen by district policymakers. These characteristics, which often include economic disadvantage, English language learner status, and special education needs, receive proportional weights relative to the base funding amount. Site-based budgeting gives principals budgetary flexibility to allocate resources based on the unique characteristics and educational needs of students in their schools. Given their proximity to students, principals can, in theory, better align resources to student needs than district leaders (Ladd, 2008; Malen et al., 2017).

While a primary goal of SBB is to promote equitable outcomes for students (Chang, 2018), analyses of how weighted student funding influences these outcomes remain limited. Much of the existing literature relies on district-level data, and few studies attempt to estimate a causal effect linking student outcomes to this funding approach. Although causal evidence suggests increased principal autonomy in budgeting improves student outcomes (Jackson, 2023), the causal evidence for increasing funding for high-needs schools via district-level weighted student formulas is scarce. Additionally, no research has explored the mechanisms that lead to the observed outcomes in these studies, which is surprising given the central role site-based budgeting plays in WSF districts (Roza et al., 2021).

In this study, we use a concurrent embedded mixed-methods design (Creswell & Creswell, 2017) and data from a large, urban district, the Southeastern Unified School District (SEUSD), to estimate the causal impact of additional SBB funding and explore the mechanisms that explain the results. This paper is one of the first studies to examine this funding policy in a causal framework and one of the few in the finance literature that embeds interview data into a quantitative study to characterize how a district finance system operates on the ground. Using the district as a case study, the primary goal of this paper is to assess the extent to which the increased funding for specific schools via the WSF component of SBB improved student outcomes, especially among disadvantaged subpopulations, and explore the mechanisms that may contribute to the student outcome effects. Specifically, we address the following research questions:

1. *Impact of additional funding*: Comparing schools that were predicted to receive more funding relative to schools that received less funding before and after the introduction of SBB, to what extent do student outcomes—test scores, attendance, and discipline outcomes—change because of additional funding?
2. *Heterogeneous impacts on student outcomes*. To what extent do student outcome effects vary by student subgroups, including economically disadvantaged (ED) students, English language learners (ELLs), and special education (SPED) students?
3. *Assessing Mechanisms*. How do the experiences of school leaders with WSF and site-based budgeting explain the results we observe?

For research question 1, we investigate the impacts of additional funding generated through the WSF formula on student results using a synthetic difference-in-differences (SDID) approach. Even though every school was subject to WSF in the 2015-16 academic year, we use differences in treatment intensity to identify causal impacts. This variation in intensity is derived from certain schools getting more funds due to their student demographics matching the district's priority categories in the weighted funding formula. Consequently, our study design contrasts schools with a high funding allocation (i.e., high-dosage schools) to those with hold-harmless funding (i.e., low-dosage schools) before and after the new funding system's introduction. We find that test score outcomes in math and ELA improved for high-dosage schools, but there is no evidence of an effect on attendance and discipline outcomes.

For research question 2, we analyze how student outcome effects vary by student subgroup. Specifically, we examine whether and to what extent economically disadvantaged students, English language learners, and students with disabilities improve because of additional funding. We focus on these subgroups as they receive additional funding through the district's WSF formula. Our findings suggest that economically disadvantaged students and English language learners experienced gains in test scores, but there were no effects on attendance or discipline outcomes. We also did not find any statistically significant effects among special education students.

For research question 3, we interview principals to explore the underlying mechanisms of the district's finance system that may explain the quantitative findings. Given a key characteristic of WSF is to empower school leaders to allocate WSF resources across their schools flexibly, we seek to understand principals' perspectives about WSF and site-based budgeting and whether their experiences help explain the quantitative results we observe. Namely, we ask principals about the reform's implementation and how it has impacted budgeting and resource allocation practices at their schools. Findings indicate that while principals across all schools attempt to align resources to students' needs, schools receiving additional WSF funding benefit from the additional flexibility increased funding provides. On the other hand, schools with no change to their budgets in the post period face financial constraints that limit their ability to allocate resources to best meet student needs, although principals suggest additional Elementary and Secondary School Emergency Relief (ESSER) funding provided in later years alleviated these constraints.

## Student-based Budgeting Background and District Context

Approximately 30 school districts across the country, serving around 5 million students, have adopted an SBB policy to guide resource allocation and budgeting (Roza et al., 2021). While specific policies, such as the student characteristics that receive additional funding and the portion of the district budget governed by the formula, may vary, districts typically give school personnel some budgeting flexibility (Roza et al., 2021). This approach aims to better meet student needs, ensuring a more equitable distribution of resources by providing increased funding for subgroups such as English language learners and special education students (Goertz & Odden, 1999). It also empowers school leaders, who are best positioned to tailor resources to student needs (Ouchi, 2006).

Although the funding schools receive under the SBB is partially a direct function of the characteristics of the students they serve (via the WSF component of SBB), research studying resource allocation equity following SBB's implementation is mixed. Scholars have found positive relationships between the proportion of schools' student populations that are considered high need and school funding amounts in the San Francisco and Oakland Unified School

Districts (Chambers et al., 2010), in the Houston Independent School District, in Cincinnati Public Schools (Miles & Roza, 2006), and in Hawaii (Levin et al., 2013). More recent multi-district studies find a more complicated relationship, with different districts allocating resources equitably for some student groups but not others (Levin et al., 2019). In some cases, discrepancies are related to how researchers define and measure equity, while in others, it appears equity is a product of the finance system's design features (Atchison & Levin, 2023). Despite evidence that WSF formulas embedded in SBB systems can mechanically improve the equitable allocation of resources in some districts, the effect on student outcomes is not well understood.

**Student-Based Budgeting Effects on Outcomes**

Studies of the effects of SBB on student outcomes are almost exclusively descriptive, and their findings are often inconsistent. The inconsistencies are likely due to the challenge of isolating the effect of SBB as a funding policy, given contemporaneous policy changes often accompany it. Limited access to student- or school-level data and policy design differences between districts also create challenges.

Descriptive analyses from various districts suggest improved student outcomes after SBB implementation. After the implementation of SBB, higher achievement scores were observed in Houston, Seattle, and Edmonton compared to relatively similar districts (Ouchi, 2006; Ouchi & Segal, 2003). Similarly, an SBB-style reform implemented statewide in California was also associated with improved test scores, on average (Derby & Roza, 2017). Using a sample of 18 districts across the United States, Tuchman, Gross, & Chu (2022) compared student test scores within-district using an event study framework and found some evidence that in the post-period, districts were able to decrease achievement gaps between students subgroups; however, as the authors point out, these districts were already closing achievement gaps in the years leading up to the policy's implementation, obfuscating whether SBB or some other cause led to the improvements in student outcomes.

Very few studies have used quasi-experimental methods to identify the causal effects of SBB or SBB-related reforms, and they reach somewhat different conclusions. Using a matched difference-in-differences method to find a causal link between SBB and test scores in the Houston Independent School District, Stroub (2018) found mixed results. While Stroub (2018) did find that school-level test scores moderately improved after the implementation of SBB, the difference-in-differences analysis comparing HISD to coarsened-exact matched schools with a similar pre-trend showed that SBB was unrelated to school-level pass rates on the Texas Assessment for Academic Skills.

Given a key theoretical benefit of SBB is to improve equity in outcomes via more equitable resource allocation, studies that use state- or district-level outcome data are unable to observe changes that occur at the school level. School-level data allows researchers to examine intra-district variation, as we do in this study. However, even aggregated school-level data can mask underlying trends in student-level outcomes. Student-level data usually offers more detail about students, including subgroup membership, which is important for heterogeneity analysis; this is especially important in equity analyses, where identifying what subgroups benefit most from reforms is central to the research question (Downes, 2004).

**Student-Based Budgeting Implementation in SEUSD**

The Southeastern Unified School District (SEUSD) is a large and diverse district that serves more than 85,000 students across approximately 150 schools. According to the state's education agency, the district has a school population of over 60 percent black or Latino students, with approximately 40 percent of students qualifying as economically disadvantaged.

Prior to adopting SBB, SEUSD used a traditional funding model that used school-level student counts to distribute full-time equivalent staff positions across schools (Miles & Roza, 2006). This traditional allocation model, which is the most common across the U.S., need not take student characteristics into account. Rather, whether schools receive an assistant principal, for example, is contingent upon enrolling enough students to qualify for one (e.g., having more than 499 students).

Given its diverse student population with varying student needs, SEUSD sought a way to more equitably use resources to better serve student needs and undertook a two-year pilot program to provide a select group of 17 principals with school budgetary flexibility beginning in 2013-14. This pilot gave principals greater discretion over a portion of their budget but did not yet give these schools additional weighted funding based on student characteristics. The insights gained during the pilot years informed the SBB design team, which consisted of six principals, three teachers, and six central office staff who helped design and build the SBB formula. In preparation for the full-district rollout, the district made significant investments in training for school leaders and established robust communication mechanisms between the central office and individual schools.

SEUSD transitioned to student-based budgeting in the 2015-2016 school year, giving principals control over $400 million of the district's budget. SBB funding flowed to zoned and magnet elementary, middle, and high schools through the SBB formula. Special education schools, pre-K centers, virtual schools, and non-traditional schools (including adult schools and alternative schools) were excluded from SBB dollars. SBB weights included a school-level weight (i.e., elementary, middle, secondary), prior academic performance weight (a proxy for economic disadvantage), English language learner weight, and special education weight.

## Data, Sample, and Methods

We use a concurrent embedded mixed-methods design in this study (Creswell & Creswell, 2017) to identify the causal impact of additional funding on student outcomes in high-dose schools, assess to what extent there are heterogeneous effects for specific subgroups of students, and explore the mechanisms leading to the results we observe. The quantitative causal analysis is the primary method guiding this study, with qualitative data embedded to explore SBB implementation details that may lead to differential effects between school types. We began concurrently collecting quantitative administrative data from the district office and conducting qualitative interviews of school leaders in 2021, and initial analyses were conducted separately on the quantitative and qualitative data. In the second stage of analysis, we elucidate how principals experienced the dosage their schools received, which helps explain the causal results we identify. Below we discuss the data and methods used in both the causal and qualitative analysis.

### Pre-Implementation Pilot Study Considerations

The district pilot study that took place in 2013-14 consisted of 17 schools, which spanned the elementary, middle, and high school levels. In the 2014-15 school year, the district gave all middle and high schools budgetary flexibility but no additional funding. While we do not have any data associated with the pilot study, we are able to identify all schools involved.

Because pilot study schools received part of the full SBB treatment—specifically, budgetary flexibility—before implementation in 2015-16, we limit our analyses to elementary schools not selected for the pilot. By focusing on elementary schools, we leverage an uncontaminated sample of schools for causal inference. After excluding six elementary schools that were part of the pilot study, we have 61 elementary schools in our analytic sample.

**Administrative Data**

To answer research questions 1 and 2, we use school- and student-level administrative data from 2011-12 to 2021-22, made available through a data-sharing agreement with the district. Our student-level data include state test scores, disciplinary events, attendance records, and demographic characteristics. Because SBB is a school-level policy, the primary unit of variation is at the school-by-year level. Therefore, we aggregate our student-level data to the school level. Our school-level finance data include each school's actual per-pupil base funding in the year before the introduction of SBB (in 2014-15) and projected per-pupil base funding based on school characteristics before the implementation of SBB.

***School-level outcomes***

State test scores in elementary school are available for $3^{rd}$ and $4^{th}$ graders in Math and English Language Arts (ELA). Although some schools in the district added a $5^{th}$ grade over our sample period, we only use $3^{rd}$ and $4^{th}$-grade scores to keep estimates comparable over time. To put test scores in a meaningful metric, we convert student scale scores to the scale used by the National Assessment of Educational Progress (NAEP) using the procedure described by Reardon et al. (2021). The scores are then standardized to a base cohort group in 2014-15, which is the year before SBB implementation. This transformation allows for comparisons to other reforms in the literature, which is policy-relevant for districts trying to improve student achievement. Test score data are unavailable for the 2015-16 school year due to a testing anomaly affecting the district. Due to the COVID-19 pandemic, we are also missing testing data from 2019-2020.

Student attendance outcomes are constructed from three variables. First, we obtain each student's total days of instruction within a school. Second, we obtain a student's total number of absences. Third, we obtain a student's total number of unexcused absences. We compute a student's absence rate as the percentage of total absences relative to the total number of instructional days. Similarly, we compute a student's unexcused absence rate as the percentage of unexcused absences relative to the total number of instructional days. We average these measures in each school and year.

Student disciplinary files record all disciplinary events across the year, including the number of times the student was disciplined each year and the type of action (in-school or out-of-school suspension). For each student and year in our administrative files, we create an indicator variable that takes value 1 if the student was given an in-school suspension (ISS) at least one time over the course of the school year and value zero, otherwise. We do the same procedure for out-of-school suspensions (OSS). We compute the proportion of these infractions in each school and year.

***Dosage-Based Measure of Treatment***

We use the district-provided school finance data to identify treatment and control schools. Because the data include schools' per-pupil actual base funding in academic year 2014-15 and the projected budget amount for the 2015-16 academic year (the first year of SBB), we leverage these numbers to simulate the fiscal impact transitioning to SBB would have on school-level budgets. We created our measure of dosage using the raw difference in per-pupil funding between the two years. If the difference is negative, the school belongs to the low-dosage group (i.e., $D = 0$). Practically speaking, schools in this category did not actually lose funding, as the district included hold-harmless safeguards; these schools would receive no less than the previous year's funding amount. Among schools for which the difference is positive, we assign them to the high-dosage group (i.e., $D = 1$). Of 61 elementary schools in our sample, 40 are high-dose schools, and 21 are low-dose schools.

Figure 1 visually shows the distribution of the funding difference between academic years 2014-15 and 2015-16. The negative predicted funding differences range from $-\$2{,}100$ to $-\$8$ dollars per pupil, but again, this is a hold-harmless group that did not lose funding. Among the high-dosage group, values range from \$43 to \$862 dollars per pupil, with the mean amount at \$312 per pupil.

We consider this simulated dosage measure to be a valid instrument for several reasons. First, the funding amounts that were projected for 2015-16 were not based on actual enrollment data for that school year. Instead, they were based on enrollment data from the 2014-15 academic year. Therefore, we are not mechanically incorporating any student sorting across schools. Second, funding amounts were not publicly broadcast to families with the introduction of SBB. Families would have no access to the 2014-15 per-pupil funding amounts nor the projections in 2015-16.

Table 1 summarizes our analytic sample based on baseline characteristics averaged across pre-treatment years 2011-12 and 2014-15. With respect to demographics, we observe higher proportions of students that were ever classified as economically disadvantaged (ED), via free- and reduced-price lunch participation (FRPL), or English language learners (ELLs) in high-dosage schools (i.e., $D_s = 1$) relative to low-dosage schools. Specifically, 86 and 32 percent of students are ED and ELL, respectively, in high-dosage schools. The corresponding percentages in low-dosage schools are 76 and 14 percent, respectively. Based on a measure of economic disadvantage according to direct certification (DC), shares are more similar between low- and high-dosage schools. Standardized test scores in the NAEP scale are also lower in high-dosage schools for both math and ELA. There are no substantive differences with respect to ISS rates across low- and high-dose schools, but students in low-dosage schools appear to have a lower likelihood of receiving an out-of-school suspension. Finally, we observe that students in low-dosage schools have higher absence rates than their high-dosage counterparts.

**Identification Strategy**

To determine the effect of differential dosages of funding, we use a difference-in-differences research design. Although all schools adopted SBB in academic year 2015-16, we leverage variation in the dosage of treatment to identify causal effects. As previously discussed, the dosage variation we use is based on our simulated instrument: some schools were predicted to receive more funding relative to others because student demographics in these schools aligned with priority categories in the district's weighted student funding formula. Our comparison group contains the low-dosage schools, or the hold-harmless group. Our research design, therefore, compares high-dosage schools with low-dosage schools before and after the introduction of SBB. Importantly, to ensure a consistent sample of schools, we include only schools that are in all years of our sample period, spanning 2011-12 to 2021-22.

We operationalize our difference-in-difference design by estimating the following equation:

$$y_{st} = \alpha + \theta_s + \mu_t + \tau(D_s \times Post_t) + \varepsilon_{st}, \quad (1)$$

where $y_{st}$ denotes an outcome for school $s$ in year $t$; $\alpha$ is a constant; $\theta_s$ is a school fixed effect, which accounts for time-invariant school level factors; $\mu_t$ is a year fixed effect to account for secular shocks that are common across the low and high-dosage schools in the district; $D_s$ is a binary indicator variable that takes value one among schools in the high-dosage group; $Post_t$ is a binary indicator variable that takes value one beginning in the 2015-16 academic year; and $\varepsilon_{st}$ is assumed to be a mean-zero random error term. We assume that errors are serially correlated within schools over time.

Our parameter of interest is $\tau$—the causal estimate of SBB for the high-dosage group after the introduction of SBB—assuming outcomes for schools in low-dosage schools serve as a valid counterfactual for what would have happened if treatment had not occurred. We provide estimates of $\tau$ for our entire treatment period (i.e., 2015-16 to 2021-22) and before the COVID-19 pandemic (i.e., 2015-15 to 2018-19), given schools received an influx of dollars from the federal government that could influence the total amount of money schools receive. This added pandemic funding could also impact the amount of flexibility principals have in their budgets, which we explore in our qualitative analysis.

To estimate our difference-in-differences model in Equation (1), we use the synthetic difference-in-differences (SDID) procedure (Arkhangelsky et al., 2021). Our initial OLS estimation of Equation (1) on the outcome variables did not reveal strong evidence of parallel pre-trends prior to the adoption of WSF, which is crucial for assuming that low-dosage schools can serve as a valid counterfactual for high-dosage schools in the difference-in-differences framework; therefore, we chose to implement the SDID procedure. Similar to OLS, the SDID approach minimizes the sum of squared residuals; however, it augments the OLS objective function through the inclusion of estimated school and year weights, reflected as $\hat{\omega}_s^{sdid}$ and $\hat{\lambda}_t^{sdid}$ in the following equation:

$$\left(\hat{\tau}^{sdid}, \hat{\alpha}, \hat{\theta}, \hat{\mu}\right) = \underset{\tau,\alpha,\theta,\mu}{\arg\min} \left\{ \sum_{s=1}^{N} \sum_{t=1}^{T} (y_{st} - \alpha - \theta_s - \mu_t - \tau(D_s \times Post_t)^2 \hat{\omega}_s^{sdid} \hat{\lambda}_t^{sdid} \right\}. \quad (2)$$

The SDID procedure selects school weights to minimize the difference in outcomes between the low-dosage schools (i.e., comparison group) and the average of the high-dosage schools (i.e., treatment group) in the pre-treatment period, helping to ensure approximate parallel trends before the introduction of WSF. Additionally, the procedure chooses pre-treatment year weights to minimize the difference in outcomes between pre-treatment periods and the average of post-treatment periods among the low-dosage schools. We estimate our model in Stata using the “sdid” command and use a clustered bootstrap to estimate standard errors (Clark et al., in press).

***Analyzing Heterogeneity Among Student Subgroups***

Recognizing that not all students are impacted equally, and recognizing that a central purpose of SBB is to improve outcomes for marginalized students, we also examine the differential effects of SBB based on distinct student subgroups: economically disadvantaged (ED) students; English language learners (ELL); and special education (SPED) students. This focus ensures our conclusions account for the diverse SEUSD student population and aligns with groups in the district’s weighted student funding formula. To estimate outcome effects by student subgroup, we estimate Equation (2) by restricting our analyses to each student subgroup (that is, ED = 1, ELL = 1, and SPED = 1) and each of their counterparts: non-ED students, non-ELLs, and non-SPED students.

***Robustness and Sensitivity Checks***

We assess the validity of our difference-in-differences design by examining pre-treatment dynamics, identifying whether there are changes in the composition of students, and estimating a generalized difference-in-differences equation via OLS. We begin by estimating event studies of our results based on the SDID approach, which we describe in Online Appendix A. If the event studies show that high- and low-dose schools were trending similarly before the introduction of SBB, then we have suggestive evidence of the parallel trends assumption needed for difference-in-differences designs. Next, we examine whether the composition of students changed because of SBB. Demographic shifts suggest that our results might be attributable to differences in the student population composition and not the impact of SBB alone. Finally, we estimate a generalized difference-in-differences equation, one that adjusts for pre-trends, using OLS as an alternative to the SDID approach. We discuss this specification in Online Appendix B.

Our next set of robustness checks applies only to test score outcomes. First, we test for differences in test score missingness for all students and demographic characteristics between low- and high-dosage schools because certain groups of students may have differentially opted out of taking the end-of-year state standardized assessment. Second, we assess the sensitivity of the NAEP standardized scale by estimating our difference-in-differences model using test scores that are standardized within the state in which the district is located and at the district level.

**Qualitative Data and Analysis**

We draw on interview data collected during the early stages of the research project to understand what mechanisms led to the results we identified in the quantitative analysis. In addition to weighted student funding, SBB provides budgetary flexibility to site-level personnel. This flexibility is key to SBB's theory of change, as school leaders are closer to students' needs, giving them more information to make better resource allocation decisions. Given principals are the key decision-makers at the school level and they report making most decisions in SBB-like systems (Jochim & Silberstein, 2020), their experiences are critical to understanding how SBB functions in practice. Once we estimated the effects using administrative data, we again turned to our interview data to explore why we observed differences in student outcomes at the school level.

We invited all principals within SEUSD to participate in the study, and 26 principals across all grade level tiers agreed to participate (roughly 17% of all SBB schools). The qualitative analysis for this paper uses interview data from all 11 elementary school principals who participated to provide insight into possible mechanisms that influence the results we observe at the elementary tier in our quantitative analysis. The 11 principals represent 18% of the 61 schools in our sample. The interviews were conducted between Summer 2021 and Summer 2022 via videoconferencing software. With permission, the sessions were recorded to ensure accurate transcription, and each principal was given a pseudonym to maintain confidentiality. Each interview lasted approximately 45 minutes, and each participating principal was compensated with a $50 gift card.

In these semi-structured interviews, we asked principals to describe how they constructed their budgets, what areas they prioritized and why, how federal ESSER funding impacted their decision-making, and their overall perceptions of SBB as a policy. We began each interview with open-ended questions about the general process principals take to construct and submit their budgets and what they prioritized before asking specific questions about their perceptions, district supports, and ESSER, to capture principals' experiences as they perceive them before following up with specific questions about topics of interest that we generated deductively from prior literature on SBB's theory of change. To ensure consistency across interviews, all authors were provided with a guide that outlined the core interview questions but were encouraged to ask respondents to expound on answers.

Of the 11 elementary school principals, 6 represent high-dose schools (around 15%) and 5 represent low-dose schools (roughly 25%). Principals of high-dosage schools, as seen in Table 2, exhibit a wide range of years in the principalship (M=6.67, SD=6.71), from early-career principals with as few as 3 years to more seasoned veterans with 20 years of principal experience. Most principals in high-dosage schools had 4 or fewer years of experience. However, they had more extensive total years of experience in education (M=21.33, SD=8.5) and typically served as teachers or assistant principals for many years before reaching the executive principal level. One commonality among principals in high-dosage schools is that most lack prior budgeting experience before working with SBB in SEUSD.

The low-dosage school principals consistently possess greater years of experience, both in their roles as principals (M=12.6, SD=6.35) and within the broader educational field (M=22.6, SD=9.4). Most have dedicated over a decade to educational leadership, and their cumulative years in education frequently surpass those in high-dosage schools. Notably, this group also features a higher proportion of principals with prior budgeting knowledge before working with SBB in SEUSD.

***Coding Strategy***

Our qualitative data analysis was conducted in two stages to ensure a thorough examination of the interview data.

**Stage One: Initial Coding and Analysis.** The initial stage of qualitative analysis was conducted concurrently with the quantitative analysis. Two graduate research assistants, including one of the authors, used the qualitative software program Dedoose for coding the Summer and Fall 2021 interviews and the software program NVivo 14 for the Summer and Fall 2022 interviews. We adopted a flexible coding approach described by Deterding and Waters

(2018). After loading the interview transcripts into the qualitative software programs, initial coding was conducted using index codes that aligned directly with our interview protocol, such as “approach to resource allocation,” “benefits of SBB,” and “impact on student outcomes” (Deterding & Waters, 2018). This was followed by detailed line-by-line coding on a select sample of interviews to generate a deeper understanding of participant responses under each index code.

To develop our analytic codes, we sorted the line-by-line codes based on their salience. Examples of such analytic codes include “flexibility in addressing student needs,” “reducing resources in low-need schools,” and “improved growth/achievement.” To ensure the reliability and validity of our coding, we employed double-coding practices on selected transcripts, effectively establishing inter-rater reliability (Miles, Huberman, & Saldaña, 2014). Discrepancies in coding were resolved through discussions within the larger research team. Analytic memos were drafted post-coding to document significant observations about the school, principal, and emergent themes, serving as evidence for verifying the credibility of findings in the subsequent stage.

**Stage Two: Comparative Analysis and Pattern Identification.** After estimating the causal effects for this study, the second stage of analysis focused exclusively on interviews with elementary school principals (n=11). We constructed matrices to categorize principal responses by school dosage, enabling direct comparisons (Miles & Huberman, 1994). This stage involved pattern recognition across responses, with the aid of stage one analytic memos, to ensure that only relevant responses linking student body composition or needs to resource allocation strategies were included in the findings. For example, responses that discussed adjustments like hiring additional ELL teachers to reduce class sizes in schools with high numbers of EL students were specifically analyzed.

**Enhancing Credibility and Integrity of Qualitative Findings.** The credibility of our qualitative findings was further bolstered by our research design and the timing of the coding procedures. Our interview questions were deliberately open-ended and designed to generate expansive responses. Importantly, all interviews were conducted before estimating causal effects to prevent any bias that might arise from prior knowledge of quantitative outcomes, ensuring that the interviews did not lead principals to provide tailored responses based on the outcomes in high- or low-dosage schools. The first-stage qualitative memos also acted as a methodological triangulation tool, corroborating the quantitative results and providing potential explanations for the causal relationships observed.

## Results

### Impact of additional funding

We formally estimate our difference-in-difference model in Equation (1) using the SDID objective function in Equation (2). We present results for test scores, attendance, and discipline outcomes in Table 3. Columns (1) and (2) present results before the COVID-19 pandemic, and columns (3) and (4) present results through the COVID-19 pandemic. Our preferred specifications are columns (2) and (4), which adjust for student race and ethnicity and the logarithm of enrollment.

#### *Test score outcomes*

In Panel A of Table 4, we find that math and ELA test scores, standardized in the NAEP scale, increased after the introduction of SBB. As previously mentioned, there were no testing data during the first year of SBB, so the average pre-pandemic effect reflects three academic years, 2016-17 to 2018-19. In math, test scores increased by 0.14 standard deviations during the pre-pandemic years (column 2). The effect remained statistically significant across the entire time span through 2021-22 but dropped to 0.12 standard deviations. In ELA, test scores increased by 0.12 standard deviations before the pandemic and remained stable throughout the entire period of study. Due to the deleterious effect of the COVID-19 pandemic on student learning, the causal warrant of results is stronger during the pre-pandemic years; however, we provide the estimate across all years, as it provides suggestive evidence that additional funding sustained outcomes over time.

These learning effect sizes are meaningful and achieved at a moderate cost. An effect size of 0.14 in math corresponds to approximately 16 to 27 percent of an annual gain in learning, while an effect size of 0.12 in reading equates to about 20 to 33 percent of an annual gain (Hill et al., 2008). Given that high-dosage schools received an additional $312 per pupil, the financial cost to achieve these learning gains is relatively low (Kraft, 2020). However, transitioning from a traditional district financing system to a Student-Based Budgeting (SBB) model entails significant changes, increasing the effective costs in terms of necessary training and ongoing support for principals.

***Attendance and Discipline Outcomes***

Panels B and C report our results on attendance and discipline outcomes for elementary school students. Unlike test scores, which only include grades 3 and 4, our non-test outcomes are provided for grades K through 4. With respect to attendance, we find no evidence of an impact of additional funding when examining the total absence rate and the unexcused absence rate. The imprecise estimates suggest that total absences increased by 0.04 percentage points but that unexcused absences fell by 0.13 percentage points before the COVID-19 pandemic. Similarly, we find no statistically significant effects on student discipline outcomes in grades K through 4. Qualitatively, the results show a decrease in the proportion of students who received an in-school or out-of-school suspension. In-school suspensions decreased by 0.23 percentage points, and out-of-school suspensions decreased by 0.79 percentage points due to additional funding before the pandemic.

**Heterogeneous impacts on student outcomes**

Table 4 reports the results of our heterogeneity analysis. We examine whether there are differential effects by economic disadvantage, English language learner status, and special education status in the years before the COVID-19 pandemic. We report two measures of economic disadvantage: one based on FRPL; the other, on direct certification. Each row denoted by (A) or (B) reflects the estimate of a different model. For example, ED=0 refers to students who were never classified as non-economically disadvantaged students, and ED=1 refers to students who were ever classified as economically disadvantaged students.

We find that economically disadvantaged students and English Language Learners in high-dosage schools performed better on math and ELA tests. Economically disadvantaged students had test score increases between 0.13 and 0.16 standard deviations in math and 0.12 standard deviations in ELA. Those ever classified as English language learners had higher learning gains: 0.21 and 0.27 standard deviations in math and ELA, respectively. Given that students in these categories receive more funding, the results suggest that principals in the highest-needs schools use their funds (and flexibility) to improve academic outcomes for the groups that need it most. We consider this hypothesis in our qualitative analysis.

We did not observe statistically significant outcomes in math or ELA for special education students, nor did we find significant effects among non-test score outcomes for any subgroup. Although the effects for special education students—who also receive additional funding—are positive, they are smaller compared to those for economically disadvantaged and English language learner subgroups. Furthermore, among all non-test-based outcomes, we found no statistically significant heterogeneous group effect, except for a marginally significant decrease in the proportion of out-of-school suspensions for English language learners.

**Robustness Checks**

***Assessing the validity of the difference-in-differences design***

We assess the validity of our quasi-experimental design by examining event studies, assessing student composition changes, and estimating a generalized difference-in-differences equation using OLS. Figure 2 shows event studies for our test score outcomes (Panel A), attendance outcomes (Panel B), and discipline outcomes (Panel C). Across test scores and attendance outcomes, we do not find strong evidence of pre-treatment effects. One concern, however, may be the estimate in 2011-12 for math and ELA scores, which appears to generate a downward trend in relation to the other pre-treatment estimates. In Online Appendix Figure C1, we remove this year from estimates, and the post-treatment effect sizes are nearly identical. We also note a statistically significant estimate for out-of-school suspensions in the 2012-2013 academic year, but this appears to be an anomaly in the context of the other pre-treatment years. Based on our event study analyses, we have suggestive evidence that the parallel trends assumption is satisfied.

Compositional changes in student demographics do not appear to weaken the causal warrant of our results. We present our estimates of compositional changes in Online Appendix Table C1. Appendix Table C1 shows a statistically significant decrease among English language learners in high-dosage schools—2.1 percentage points—and a statistically significant increase in Asian students—0.73 percentage points—in grades K to 4. The decrease in English language learners is somewhat concerning, reflecting a 6.6 percent decrease relative to pre-treatment years. However, when we focus on tested grades (i.e., grades 3 and 4), where our results are statistically significant in math and ELA, we only find a small decrease in the proportion of Pacific Islander students—0.14 percentage points. These results suggest that changes in student composition do not drive our primary test score results.

Lastly, we find qualitatively similar results when we estimate a generalized difference-in-differences model that adjusts for linear pre-trends at the school level via OLS. As shown in Online Appendix Table C2, results are qualitatively similar to our main results in Table 3. Test score results are statistically significant and slightly larger in magnitude; attendance and discipline results remain statistically insignificant. Event study figures for these models, which appear in Online Appendix Figure C2, also show no substantial evidence of pre-treatment effects or trends, aside from out-of-school suspensions, which was also a concern with the SDID approach.

***Robustness of Test Score Outcomes***

For test score outcomes, we examine whether test score data are differentially missing between high- and low-dosage schools and whether results are stable across scaling changes to test scores. To assess missingness, we match demographic data on all students to their test scores. If a test score cannot be matched to a student in a school, we code the missing observation with a value of 1 and 0, otherwise. We then calculate the proportion of missing scores for all students and by student subgroup. Online Appendix Table C3 shows no evidence of differential missing between high- and low-dosage schools.

Next, we assess how our point estimates change when we standardize test scores relative to the student-level means and standard deviations at the state and district levels. To standardize scores in the state distribution, we standardized student scale scores by year and grade using the average student scale score and standard deviation of student scale scores in the state. We then average the standardized scores at the school level to create our outcome measure. We use the same process to standardize test scores at the district level, but we leverage the average student scale score and standard deviation from the district. Online Appendix Table C4 shows that our point estimates are similar to our benchmark results in Table 3, regardless of the transformation. We privilege the NAEP transformation, however, because we are able to put scores on a common scale and standardize them relative to the year before SBB took effect. Moreover, because the district had changed tests during our sample period, the NAEP scale transformation allows us to have a consistent outcome measure over time.

**Assessing Mechanisms**

This section examines potential mechanisms that may help explain the results we observe in our quantitative analysis. Overall, interview data suggest that although principals of both low- and high-dosage schools are aware of the unique needs of their students and aim to target resources to fill those needs, high-dosage schools benefit from greater budgeting flexibility due to the additional funding they receive via the weighted-student funding formula. Principals of high-dosage schools describe more easily adapting to student needs and investing in targeted personnel and programs, leading to enhanced student outcomes. In contrast, principals in low-dosage schools describe grappling with financial constraints, limiting their ability to cater to the needs of their student body. Principals of both school types celebrated the additional funding received because of ESSER, but low-dosage school principals claim the funds were especially helpful in allowing them to target funding in flexible ways not afforded to them under the WSF component of SBB alone.

***Responding to student needs in high- and low-dose schools***

When we asked principals what influenced their budgetary decision-making process, high- and low-dosage school principals consistently highlighted the characteristics and needs of their students and the families they serve. For example, Principal Clark, who leads a low-dosage school, reflected, "Our school is predominately Hispanic and Latino, and our poverty rate is very high as well. What can we do? What do we need that's probably different from a school someplace else? And so that's how we decide how we're going to spend our money." Principal Fulton, who leads a high-dosage school, takes this a step further when explaining that supporting students goes beyond knowing what administrative categories they belong to. She stated, "You've got the trauma lens, you've got the immigrant lens, you've got the non-English speaking lens. You've got the domestic abuse lens. You really have to see it from every angle in order to meet that student's needs." This is echoed by Principal Miller, who conveys that she "very intimately" understands "even the unspoken things that [her] students may need." She "considers all of that" when she makes decisions. Principals do not solely focus on overt signs of struggle or success; they also pick up on the subtleties and the hidden challenges that might otherwise remain unnoticed.

Even principals of low-dosage schools that serve a relatively high-achieving, high-income population reflect on the unique needs of their student body. Take Principal Porter, who leads a low-dosage school with a low poverty rate: "We really want to focus on individual student growth, for every student, regardless of where they are. We do have a lot of high-performing students. And so, we have to add in some enrichment personnel, to really help drive those top 99 percentile students above and beyond." These principals often noted that their students have unique needs that require supports that are quite distinct from schools serving students who are behind academically. Another, Principal Jones, stated,

> We are historically a high-achieving school. We have a lot of kids that will score in the top two quintiles on our math assessments. So for us, it's a big focus on, how do we continue that that growth when we do have those high-achieving students, and making sure we have the supports to keep them moving forward.

At first blush, principals across high- and low-dosage schools took a needs-based approach when budgeting, centering their knowledge of their school community when allocating resources to ensure their students received the resources they needed to be successful.

In addition, principals cited their proximity to students as a significant benefit of the SBB budgeting process. They emphasized that their decisions would be more impactful for students than those made by the central office. For example, Principal Baker, who leads a high-dosage school, stated of her leadership team:

> We truly know our school and the student population that we serve. We know the needs of our families. So being able to arrange budget dollars to meet the needs of a wide variety of different demographics and backgrounds…We know the struggles that our students go through on a daily basis, even before they come to school.

The local decision-making aspect associated with site-based budgeting suggests that those closest to the students can effectively target resources where they are most needed.

***Equity in funding leads to disparity in flexibility***

Although principals from both low- and high-dosage schools claim the unique challenges their student bodies face influence their budgetary decision-making, they differ in the resources available to them, which generated opportunities for high-dosage school principals and constraints for low-dosage school principals. High-dosage school principals, in working with student subgroups that have historically been identified as requiring additional resources, consistently emphasize the benefit that the flexibility in their budgets add. Low-dosage principals cited strategic investments as well, but with a caveat that after mandatory staffing needs were met, little discretionary funding was left over.

High-dosage principals consistently made confident, declarative statements about the flexibility benefits that result from SBB. Principal Smith reflects, “If I see a specific need for my school, I can spend money to get those resources, which is nice.” Principal Davis mirrors this sentiment: “We’re able to make decisions that benefit our school instead of that being made for us...we get to determine the need and figure out how to meet that need.” Additional SBB funding serves as an instrument of empowerment, enabling educational leaders to respond proactively to the evolving needs of their schools. Pressed for examples of how they leverage this flexibility, Principal Fulton stated that because “literally every new kindergartner” new to her school was an EL student at the time of our interview, she targeted this need with additional staff:

> This year I heavily invested in EL. I invested a lot of more money than normal in EL, because I wanted every grade level to have their own EL teacher, their own EL support person. That was a very high request on my staff budget survey, and our EL population is growing.

Unlike principals of high-dosage schools, all 5 principals in low-dosage elementary schools at some point commented on a lack of adequate resources. Although the WSF formula targets funding to schools serving specific student populations, low-dosage schools can still enroll these populations, albeit in smaller proportions, leaving them with fewer resources to respond to those student needs. Principal Jones, a low-dosage principal, expressed this frustration in his discussion of the design of the WSF, while also acknowledging that the equity WSF brings is an appropriate goal:

> It can be frustrating that we're not going to have the same amount of funds as a similar sized school that has a higher, you know, EL population or low socioeconomic population. At the same time, we know that [high-dosage schools] need those resources, know that there's going to be additional support that is needed there. And that's where it should be going….It just makes you feel that things are more equitable overall.

Other low-dosage principals also grappled with this tension and explained in detail the various school-specific issues that lower levels of WSF funding exacerbates. Principal Porter, for example, operates a small school of high achievers, and given the state has staffing ratio requirements, the lower level of funding from WSF presents a challenge when assigning funding to staff. In a conversation about these challenges, she replies: "On the smaller [school] size, they assign 3.2 related arts teachers. Well, you can't…that doesn't work. You're going to have fourth person, so that's really where size impacts us or hurts us." She continues, recognizing the purpose of WSF is to increase equity while also recognizing the consequences for her school: "Equity is not equality, hence not everybody needs the same amount. So yes, poverty needs a greater weight. EL needs a greater weight. If my school and demographics are low in that, then that's it."

The other three low-dosage principals, Principals Grant, Hayes, and Clark, echoed Principals Jones and Porter, claiming they did not have enough funding. Principal Hayes stated, "Once you start paying out salaries and doing your mandatory stuff, you don't have as much money as you think." The struggle extends to basic supplies, as Principal Clark laments: "It's been very lean; we barely could buy paper…I don't get adequate funds to really meet all the needs of the students that I have." Principal Grant claims, "It's like bare bones, but just enough to get you by." Later, when asked about discretionary funding he reminded the interviewer, "the funds don't go that far."

Limited funding and budgetary requirements consistently emerge as dominant themes across experiences of low-dosage principals. The recurring comments about inadequate funds indicate a challenge where financial scarcity limits their capacity to invest in certain resources, tools, or personnel. Furthermore, the fixed costs associated with mandatory expenditures, such as salaries for required personnel, further constrict the already limited fiscal space for other identified needs. The difficulty of stretching limited resources to meet the diverse needs of students, as portrayed by these principal accounts, highlights the tradeoffs inherent in a SBB system and offers a potential explanation for the observed divergence in outcome trends between high-dosage and low-dosage schools.

***ESSER Funding Minimized Funding Constraints***

Due to COVID-19, schools received increased funding to help address student needs, which the principals in both groups expressed was beneficial; however, leaders of low-dosage schools in particular emphasized how ESSER allowed them to have the flexibility that SBB had not afforded them. For example, Principal Clark states in reference to ESSER, "I'll be honest, this is the first year coming, the year we're getting ready to move into, where I've had enough money to go in above and beyond to really feel that it is equitable to the clients that I serve. And that's because we've gotten a lot of COVID money." Like Principal Clark, when explicitly asked about ESSER funding, nearly all low-dose school principals described the flexibility it provided that was not present in their budgets before, and principals described how they leveraged this flexibility to target specific student and teacher needs in a variety of ways.

Principals across both groups of schools used ESSER funding to hire additional personnel (e.g., full-time translators, numeracy coach, after-school tutors) and pay for non-personnel expenses (e.g., technology, intervention programs, and supplies), as well as build the capacity of teachers through additional PD and EL certifications. Take, for example, Principal Hayes, who used the additional funding to pay for more interventionists and instructional staff. He states, "Because of that funding, I've been able to hire more interventionists. People that can come in and do what the extra funding is designed to do to remediate and catch up for the learning gaps that COVID created in students' education." He continues, "I've hired additional staff to lower class counts… You're able to really reach students and either remediate or accelerate their learning. It's a lot easier to do that when you have 14, 15 students compared to when you have 25, 26." By hiring interventionists or additional general education teachers, low-dose schools were able to provide the same services many high-dose schools had become accustomed to.

Instead of investing in personnel, some principals emphasized bolstering teacher capacities through professional development and certification. Principal Porter, for example, was in the process of launching a new literacy curriculum and used his funds to help teachers plan: "I have been able to pay teachers to come in, during the summer, to plan collaboratively. That jumped up, dramatically, through the use of ESSER funds." Principal Grant, leading another low-dose school, states,

> I invest in teachers so they can build themselves, which, in turn gets greater investment into your classrooms… I've used the money for management consultants providing

> professional development, working with our teachers. I think because as large as our district is, to have some extra hands on deck to help teachers is crucial.

Further illustrating her commitment to teachers' professional practice, she continues, "We have used our ESSER funds, because our EL population has grown, to build capacity with teachers by getting their EL certification." Sometimes verbally worrying about the longevity of the ESSER funds, these principals leveraged the influx of dollars to build staff instructional capacity rather than hire additional personnel.

Critically, the ESSER funding served as a financial catalyst for low-dosage schools, minimizing some of the constraints they had grappled with in previous years of SBB. This influx of resources provided these schools with the opportunity to make strides, further enabling them to address student needs more comprehensively.

## Discussion and Conclusion

Our study stands as one of the first to examine the causal impact of SBB, which is the combination of weighted student funding and site-based budgeting. Overall, we found that test score outcomes improved in both math and ELA by 0.14 and 0.12 standard deviations, respectively. However, we did not find strong evidence for changes in attendance and discipline. Our heterogeneity analysis shows that ED and ELL students in high-dosage schools had improved test-score outcomes, although we did not observe a similar pattern for special education students. These findings underscore the potential for SBB to address educational disparities among traditionally marginalized student groups in the highest-need schools. Our qualitative findings lend further support to these results.

Our qualitative examination of student-based budgeting and its influence on student outcomes has shed light on the distinctive challenges and advantages faced by both high-dosage and low-dosage schools. Facilitated by enhanced budgetary flexibility, principals in high-dose schools described various ways they strategically targeted student needs in tailored ways. Conversely, low-dosage schools emphasize the obstacles they encounter due to financial constraints, which can limit student growth within these schools compared to the growth experienced by students in high-dosage schools. The principal accounts from low-dosage schools underscore the profound impact of limited resources, describing the challenges they confront in adapting to the evolving needs of their students. This persistent financial struggle may offer insight into the observed disparities in outcomes between the two categories of schools.

While our study provides causal evidence of the impact of additional funding on student outcomes, we also acknowledge several limitations. First, we focus on a single urban district, so generalizability will be limited to other districts with similar demographics and SBB systems. Despite this limitation, the results can still inform policy discussion about the effectiveness of WSF as a tool for differentiating funding to support student success in combination with site-based budgeting, which is a tool that provides principals with the ability to leverage additional resources to target student needs. Second, we recognize that this study is limited to elementary schools. Because the middle and high schools in SEUSD were part of a pilot study that implemented the site-based budgeting component before WSF was implemented, we had to exclude these schools to avoid bias. Future work should examine schooling outcomes among older students to provide a comprehensive overview of SBB, especially since funding effects may be heterogeneous at the middle and high school levels. Finally, our results suggest that additional funding did not affect non-test score outcomes, but we only had access to a small set of non-test-based outcomes. Future research should examine the impacts of additional funding on social-emotional outcomes, which may complement test score gains.

**References**

Archangelsky, D., Athey, S., Hirshberg, D. A., Imbens, G. W., & Wager, S. (2021). Synthetic Difference-in-Differences. *American Economic Review*, *111*(12), 4088-4118.

Atchison, D., & Levin, J. (2023). Missed opportunities: Weighted student funding systems and resource equity. *Journal of Education Human Resources*, *41*(1), 142-171.

Biasi, B. (2023). School Finance Equalization Increases Intergenerational Mobility. *Journal of Labor Economics*, *41*(1), 1-38.

Candelaria, C. A., & Shores, K. A. (2019). Court-ordered finance reforms in the adequacy era: Heterogeneous causal effects and sensitivity. Education Finance and Policy, *14*(1), 31-60.

Chambers, J. G., Levin, J. D., & Shambaugh, L. (2010). Exploring weighted student formulas as a policy for improving equity for distributing resources to schools: A case study of two California school districts. *Economics of Education Review*, *29*, 283-300.

Chang, B. (2018). *Transforming School Funding for Equity, Transparency, and Flexibility: An Introduction to Student-Based Budgeting.* Watertown, MA: Education Resource Strategies.

Clark, D., Pailañir, D., Athey, S., & Imbens, G. (In Press). On Synthetic Difference-in-Differences and Related Estimation Methods in Stata. *Stata Journal*. Forthcoming.

Condron, D. & Roscigno, V. J. (2003). Disparities within: Unequal spending and achievement in an urban school district. *Sociology of Education, 76*(1), 18–36. https://doi.org/10.2307/3090259

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Derby, E., & Roza, M. (2017). *California's Weighted Student Formula: Does It Help Money Matter More? Rapid Response.* Washington, DC: Edunomics Lab at Georgetown University.

Downes, T. (2004). School finance reform and school quality: Lessons from Vermont. In J. Yinger (Ed.), Helping Children Left Behind: State Aid and the Pursuit of Educational Equity.

Goertz, M. E., & Odden, A. R. (1999). *School-based financing* (Vol. 20). Corwin Press.

Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical Benchmarks for Interpreting Effect Sizes in Research. *Child Development Perspectives*, *2*(3), 172–177. https://doi.org/10.1111/j.1750-8606.2008.00061.x

Houck, E. A. (2011). Intradistrict Resource Allocation: Key Findings and Policy Implications. Education and Urban Society, 43(3), 271–295. https://doi.org/10.1177/0013124510380234

Iatarola, P. & Stiefel, L. (2003). Intradistrict equity of public education resources and performance. *Economics of Education Review*, *22*(1), 69–78. https://doi.org/10.1016/S0272-7757(01)00065-6

Jackson, C. K. (2020). Does school spending matter? The new literature on an old question. In L. Tach, R. Dunifon, & D. L. Miller (Eds.), APA Bronfenbrenner series on the ecology of human development. Confronting inequality: How policies and practices shape children's opportunities (p. 165–186). American Psychological Association. https://doi.org/10.1037/0000187-008.

Jackson, C. K. (2023). When Does School Autonomy Improve Student Outcomes?. *Annenberg Institute at Brown University, EdWorkingPaper*, 23-808.

Jackson, C. K., Johnson, R. C., & Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *The Quarterly Journal of Economics*, *131*(1), 157-218.

Jackson, C. K., & Persico, C. (2023). Point column on school spending: Money matters. *Journal of Policy Analysis and Management*, *42*(4), 1118-1124.

Jochim, A., & Silberstein, K. (2020). Taking stock of principals' role in weighted student funding districts. *Edunomics Lab*.

Knight, D. S. (2019). Are School Districts Allocating Resources Equitably? The Every Student Succeeds Act, Teacher Experience Gaps, and Equitable Resource Allocation. *Educational Policy*, *33*(4), 615-649.

Kraft, M. A. (2020). Interpreting Effect Sizes of Education Interventions. *Educational Researcher*, *49*(4), 241–253. https://doi.org/10.3102/0013189X20912798

Ladd, H. F. (2008). Reflections on equity, adequacy, and weighted student funding. *Education Finance and Policy*, *3*, 402-423.

Lafortune, J., Rothstein, J., & Schanzenbach, D. W. (2018). School finance reform and the distribution of student achievement. *American Economic Journal: Applied Economics*, *10*(2), 1-26.

Levin, J., Chambers, J., Epstein, D., Mills, N., Archer, M., Wang, A., & Lane, K. (2013). *Evaluation of Hawaii's weighted student formula*. American Institutes for Research. https://www.air.org/resource/report/evaluation-hawaiis-weighted-student-formula

Levin, J., Manship, K., Hurlburt, S., Atchison, D., Yamaguchi, R., Hall, A., & Stullich, S. (2019). Districts' Use of Weighted Student Funding Systems to Increase School

Autonomy and Equity: Findings from a National Study. Volume 1--Final Report. Office of Planning, Evaluation and Policy Development, US Department of Education.

Malen, B., Dayhoff, J., Egan, L., & Croninger, R. G. (2017). The challenges of advancing fiscal equity in a resource-strained context. Educational Policy, *31*(5), 615-642.

McGee, J. B. (2023). Yes, money matters, but the details can make all the difference. *Journal of Policy Analysis and Management*, *42*(4), 1125-1132.

Miles, K. H., & Roza, M. (2006). Understanding student-weighted allocation as a means to greater school resource equity. *Peabody Journal of Education*, *81*(3), 39-62.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook*. Sage Publishing.

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook*. Sage Publications.

Ouchi, W. (2006). Power to the Principals: Decentralization in Three Large School Districts. *Organization Science (Providence, R.I.)*, *17*(2), 298–307.

Ouchi, W. & Segal, L. (2003). *Making public schools work: A Revolutionary plan to get your children the education they need.* Simon & Schuster. New York, NY.

Reardon, S. F., Kalogrides, D., & Ho, A. D. (2021). Validation Methods for Aggregate-Level Test Scale Linking: A Case Study Mapping School District Test Score Distributions to a Common Scale. *Journal of Educational and Behavioral Statistics*, *46*(2), 138–167.

Roza, M., Hagan, K., & Anderson, L. (2021). Variation is the norm: A landscape analysis of weighted student funding implementation. *Public Budgeting & Finance*, *41*(1), 3-25.

Rubenstein, R., Schwartz, A. E., Stiefel, L., & Hadj Amor, H. B. (2007). The distribution of resources across schools in big city school districts, *Economics of Education Review*, *26*(5), 532-545.

Shores, K. A., Candelaria, C. A., & Kabourek, S. (2023). Spending More on the Poor? A Comprehensive Summary of State-Specific Responses to School Finance Reforms from 1990–2014. *Education Finance and Policy*, *18*(3), 395-422.

Shores, K., & Ejdemyr, S. (2017). Pulling back the curtain: Intra-district school spending inequality and its correlates. Available at http://dx.doi.org/10.2139/ssrn.3009775.

Stroub, K. (2018). *Decentralization reform in HISD, part III: Evidence on the relationship between decentralization and achievement*. Houston, TX: Houston Education Research Consortium, 2018.

Tuchman, S., Gross, B., & Chu, L. (2022). Weighted student funding and outcomes: Implementation in 18 school districts. *Peabody Journal of Education*, *97*(4), 479-496.

## Tables

**Table 1.** Descriptive Statistics, averaged across pre-treatment years (2011-12 to 2014-15)

| | D=0<br>Low Dosage | D=1<br>High Dosage | Full Sample |
|---|---|---|---|
| *A. Demographic Characteristics* | | | |
| Ever Economically Disadvantaged (FRPL) | 0.76 | 0.86 | 0.83 |
| Ever Economically Disadvantaged (DC) | 0.60 | 0.58 | 0.58 |
| Ever English Language Learner (ELL) | 0.14 | 0.32 | 0.27 |
| Ever Special Education (SPED) | 0.21 | 0.18 | 0.19 |
| Female | 0.49 | 0.49 | 0.49 |
| White | 0.39 | 0.34 | 0.36 |
| Asian | 0.036 | 0.045 | 0.042 |
| Black | 0.43 | 0.34 | 0.37 |
| Hispanic | 0.14 | 0.27 | 0.23 |
| American Indian | 0.0016 | 0.0018 | 0.0017 |
| Pacific Islander | 0.0015 | 0.0015 | 0.0015 |
| *B. Test Scores (3rd and 4th Grade)* | | | |
| NAEP Scale Standardized Math Score | -0.0043 | -0.23 | -0.16 |
| | (0.91) | (0.86) | (0.88) |
| NAEP Scale Standardized ELA Score | -0.085 | -0.290 | -0.230 |
| | (0.98) | (0.91) | (0.93) |
| *C. Discipline* | | | |
| In-School Suspension (ISS) | 0.00091 | 0.00095 | 0.00094 |
| Out-of-School Suspension (OSS) | 0.049 | 0.028 | 0.034 |
| *D. Attendance* | | | |
| Absence Percentage Rate | 5.69 | 4.97 | 5.18 |
| | (9.09) | (7.25) | (7.83) |
| Unexcused Abs. Percentage Rate | 3.11 | 2.35 | 2.57 |
| | (8.47) | (6.39) | (7.06) |
| Number of Schools | 20 | 41 | 61 |

*Notes*: Standard deviations of continuous variables appear in parentheses below the means. The sample is limited to elementary schools open across all years of our sample period between 2011-12 and 2021-22. Given that the district switched between free and reduced-priced lunch (FRPL) and direct certification (DC) over our sample period, we report two measures of economic disadvantage.

**Table 2**. Descriptive Characteristics of Principals in Qualitative Sample

| SBB Type | Pseudonym | Prior Budgeting Experience | Years as Principal | Total Years in Education |
|---|---|---|---|---|
| D=1 (High Dosage) | Baker | Yes | 3 | 14 |
| | Taylor | No | 3 | 17 |
| | Davis | No | 3 | 13 |
| | Miller | No | 4 | 22 |
| | Fulton | No | 7 | 27 |
| | Smith | No | 20 | 35 |
| | | *Average (SD)* | 6.67 (6.71) | 21.33 (8.5) |
| D=0 (Low Dosage) | Jones | No | 5 | 12 |
| | Hayes | Yes | 7 | 17 |
| | Porter | Yes | 15 | 19 |
| | Clark | Yes | 16 | 33 |
| | Grant | No | 20 | 32 |
| | | *Average (SD)* | 12.6 (6.35) | 22.6 (9.4) |

**Table 3.** Impact of Additional Funding from SBB via WSF on Elementary School Student Outcomes

| | Pre-Pandemic Effect | | Effect Across All Years | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *A. Test Scores: Grades 3 and 4* | | | | |
| Math (NAEP Standardized Scale) | 0.15** | 0.14* | 0.13* | 0.12* |
| | (0.05) | (0.06) | (0.05) | (0.06) |
| ELA (NAEP Standardized Scale) | 0.12* | 0.12* | 0.11* | 0.12* |
| | (0.05) | (0.06) | (0.05) | (0.06) |
| N. Observations | 427 | 427 | 549 | 549 |
| *B. Attendance: Grades K to 4* | | | | |
| Total Absence Rate (%) | 0.025 | 0.042 | 0.049 | 0.056 |
| | (0.25) | (0.28) | (0.31) | (0.35) |
| Unexcused Abs. Rate (%) | -0.073 | -0.13 | -0.037 | -0.049 |
| | (0.23) | (0.28) | (0.28) | (0.34) |
| N. Observations | 488 | 488 | 671 | 671 |
| *C. Discipline: Grades K to 4* | | | | |
| In-School Suspension Proportion | -0.0044 | -0.0023 | -0.0025 | 0.00022 |
| | (0.0046) | (0.0039) | (0.0031) | (0.0024) |
| Out-of-School Suspension Proportion | -0.0084 | -0.0079 | -0.0046 | -0.0057 |
| | (0.0082) | (0.0072) | (0.0074) | (0.0063) |
| N. Observations | 488 | 488 | 671 | 671 |
| Controls for Student Demographics | No | Yes | No | Yes |
| N. Elementary Schools | 61 | 61 | 61 | 61 |

Notes: Bootstrapped standard errors, clustered at the school level, are reported in parentheses. Statistical significance levels: + $p< 0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$. Observation counts are at the school-by-year level. Results with controls adjust for student race and ethnicity and the logarithm of enrollment.

**Table 4.** Heterogeneity Analyses Among Subgroups Before the Pandemic

| | Math | ELA | Abs. Rate (%) | Unex. Abs (%) | ISS Prop. | OSS Prop. |
|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Ever Economically Disadvantaged (FRPL):* | | | | | | |
| A. ED=0 | -0.14 | -0.040 | -0.078 | -0.041 | 0.0055 | 0.0011 |
| | (0.14) | (0.12) | (0.98) | (0.83) | (0.0045) | (0.016) |
| N. Observations | 371 | 371 | 416 | 416 | 416 | 416 |
| B. ED=1 | 0.16** | 0.12** | -0.41 | -0.25 | -0.0057 | -0.0095 |
| | (0.060) | (0.044) | (0.26) | (0.23) | (0.0063) | (0.0097) |
| N. Observations | 427 | 427 | 488 | 488 | 488 | 488 |
| *Ever Economically Disadvantaged (DC):* | | | | | | |
| A. ED=0 | 0.044 | 0.033 | 0.024 | -0.18 | 0.00023 | -0.0046 |
| | (0.069) | (0.060) | (0.26) | (0.21) | (0.0014) | (0.0060) |
| N. Observations | 420 | 420 | 488 | 488 | 488 | 488 |
| B. ED=1 | 0.13+ | 0.12* | -0.31 | -0.27 | -0.0039 | -0.0076 |
| | (0.069) | (0.053) | (0.28) | (0.25) | (0.0049) | (0.012) |
| N. Observations | 427 | 427 | 488 | 488 | 488 | 488 |
| *Ever English Language Learner:* | | | | | | |
| A. ELL=0 | 0.076 | 0.059 | 0.17 | 0.075 | 0.0012 | -0.0062 |
| | (0.068) | (0.061) | (0.33) | (0.33) | (0.0038) | (0.0090) |
| N. Observations | 427 | 427 | 488 | 488 | 488 | 488 |
| B. ELL=1 | 0.21* | 0.27** | -0.029 | -0.16 | -0.0053 | -0.022+ |
| | (0.10) | (0.097) | (0.34) | (0.35) | (0.0064) | (0.012) |
| N. Observations | 392 | 392 | 440 | 440 | 440 | 440 |
| *Ever Special Education Student:* | | | | | | |
| A. SPED=0 | 0.15* | 0.14* | 0.13 | 0.11 | -0.0030 | -0.0074 |
| | (0.073) | (0.064) | (0.28) | (0.28) | (0.0036) | (0.0080) |
| N. Observations | 427 | 427 | 488 | 488 | 488 | 488 |
| B. SPED=1 | 0.048 | 0.045 | 0.11 | 0.0049 | -0.0023 | 0.0036 |
| | (0.079) | (0.085) | (0.38) | (0.34) | (0.0077) | (0.016) |
| N. Observations | 427 | 427 | 488 | 488 | 488 | 488 |

Notes: Bootstrapped standard errors, clustered at the school level, are reported in parentheses. Statistical significance levels: + $p$< 0.10; * $p$<0.05; ** $p$<0.01; *** $p$<0.001. Each row denoted by A or B reflects a different regression conditional on the group indicated. For example, $ED = 0$ refers to non-economically disadvantaged students, and $ED = 1$ refers to economically disadvantaged students. Observation counts are at the school-by-year level. Results with controls adjust for student race and ethnicity and the logarithm of enrollment.

**Figures**

**Figure 1. Dosage Group Characterization**

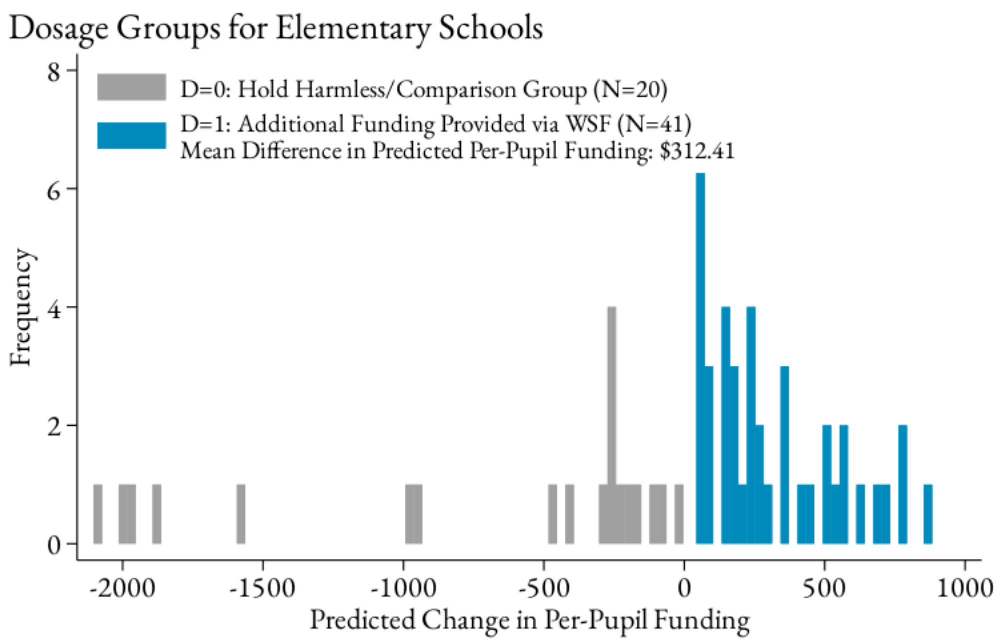
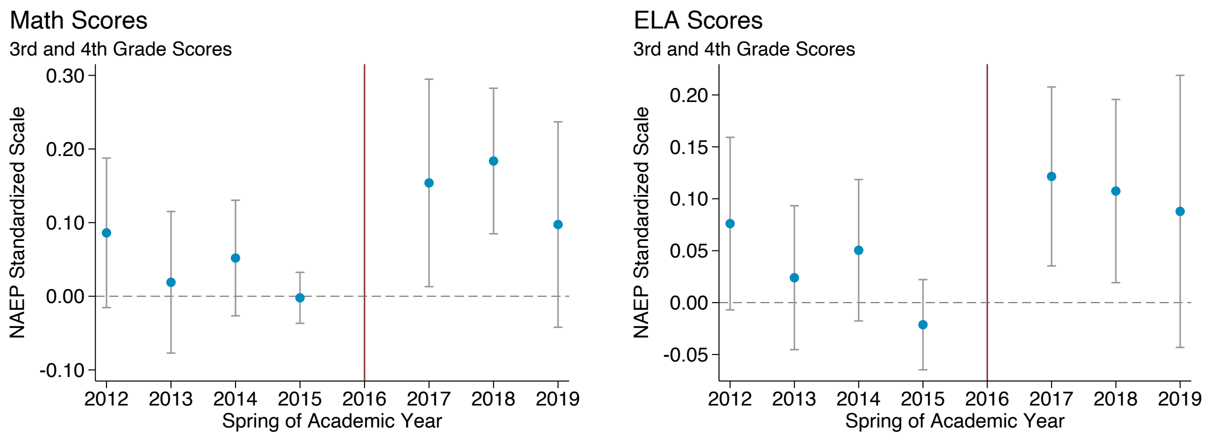


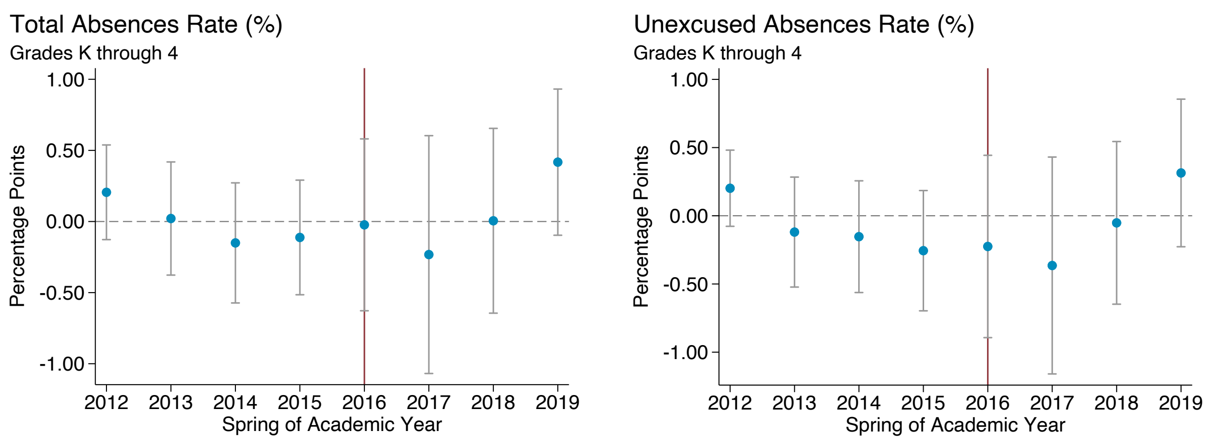
**Figure 2. Event Studies**

**Panel A. Test Score Outcomes**



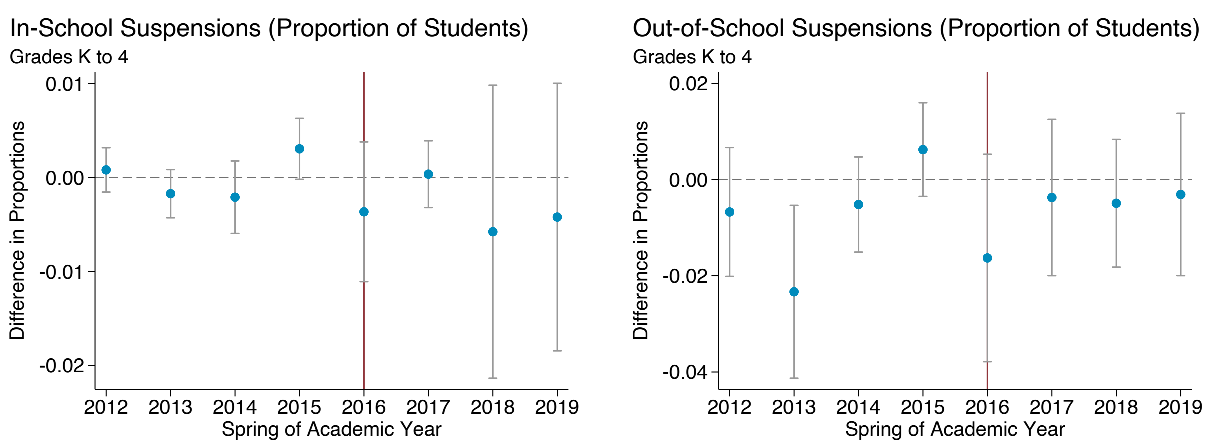

(a) (b)

**Panel B. Attendance Outcomes**




(a) (b)

**Panel C. Discipline Outcomes**




(a) (b)

*Notes*: Event studies based on SDID estimation. Error bars are 95 percent confidence intervals. Details regarding the estimation approach to produce the figure appear in Online Appendix A.

# ONLINE APPENDIX

## Appendix A. Event Studies in the SDID Framework

In an event study framework saturated with binary indicator variables for leads and lags of treatment time, one of the indicator variables is excluded and serves as the base reference year. Typically, this base year is the year before treatment begins. As noted by Clark et al. (in press), the SDID approach does not require excluding a base year. The control and treated group baselines are based on the following equations:

$$\bar{y}_{\text{baseline}}^{Control\ (D=0)} = \sum_{t=1}^{T_{\text{pre-treatment}}} \hat{\mu}_t^{SDID} \times \bar{y}_t^{Control}$$

$$\bar{y}_{\text{baseline}}^{Treatment\ (D=1)} = \sum_{t=1}^{T_{\text{pre-treatment}}} \hat{\mu}_t^{SDID} \times \bar{y}_t^{Treatment},$$

where $\hat{\mu}_t^{SDID}$ is the optimal pre-treatment time weight in year $t$ and $\bar{y}_t^{Control}$ and $\bar{y}_t^{Treatment}$ are the average values of outcome $y_t$ in year $t$ for the control and treatment groups, respectively. Having computed a baseline average, the event study treatment effect in each year $t$ is defined as

$$\left(\bar{y}_t^{Treatment} - \bar{y}_t^{Control}\right) - \left(\bar{y}_{\text{baseline}}^{Treatment} - \bar{y}_{\text{baseline}}^{Control}\right),$$

where $\bar{y}_t^{Treatment}$ is simply the average of the treatment schools in year $t$ and $\bar{y}_t^{Control}$ is the synthetic control average of comparison schools.

## Appendix B. Generalized Difference-in-Difference model

In addition to estimating our difference-in-difference model using the SDID approach, we also estimate a generalized difference-in-difference model via OLS, which is also referred to as a comparative interrupted time series (CITS) model. We operationalize our CITS design by estimating the following equation:

$$y_{st} = \alpha + \theta_s + \mu_t + \theta_s t + \tau(D_s \times Post_t) + \varepsilon_{st},$$

where $y_{st}$ denotes an outcome for school $s$ in year $t$; $\alpha$ is a constant; $\theta_s$ is a school fixed effect, which accounts for time-invariant school level factors; $\mu_t$ is a year fixed effect to account for secular shocks that are common across the low and high-dosage schools in the district; $D_s$ is a binary indicator variable that takes value one among schools in the high-dosage group; $Post_t$ is a binary indicator variable that takes value one beginning in the 2015-16 academic year; and $\varepsilon_{st}$ is assumed to be a mean-zero random error term. We assume that errors are serially correlated within schools over time.

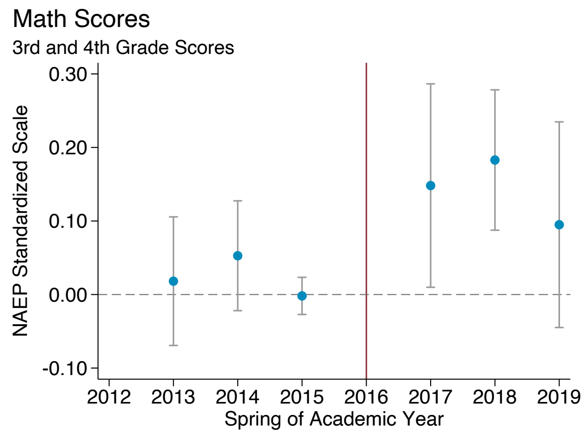
The key difference between the CITS model and Equation (1) is the inclusion of $\theta_s t$, a linear time trend to account for any pre-treatment differential trends. The parameter of interest is still $\tau$—the causal estimate of SBB for the high-dosage group after the introduction of SBB.

## Appendix C. Robustness Check Results

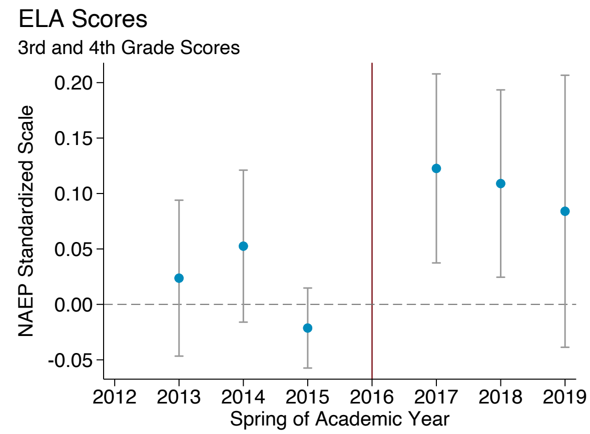
### Test Score Event Study after Dropping 2011-12 as a Pre-Treatment Observation

**Figure C1**. Event Study after Dropping 2011-12

**Panel A. Test Score Outcomes**




(a)




(b)

## Compositional Changes in Student Demographics

**Table C1. Robustness Check: Demographic Compositional Changes Before COVID-19**

| | Grades K-4 | Grades 3 to 4 |
|---|---|---|
| Enrollment | -6.86 | 6.41 |
| | (17.6) | (8.38) |
| Ever Economically Disadvantaged (FRPL) | -0.016+ | 0.00016 |
| | (0.0096) | (0.010) |
| Ever Economically Disadvantaged (DC) | 0.0016 | 0.013 |
| | (0.0094) | (0.016) |
| Ever English Language Learner (ELL) | -0.021* | -0.0048 |
| | (0.010) | (0.011) |
| Ever Special Education (SPED) | -0.0030 | -0.0025 |
| | (0.0067) | (0.013) |
| Female | -0.0031 | -0.0018 |
| | (0.0079) | (0.012) |
| Asian | 0.0073* | 0.0029 |
| | (0.0031) | (0.0037) |
| Black | -0.0050 | -0.0039 |
| | (0.011) | (0.014) |
| Hispanic | 0.0033 | 0.0090 |
| | (0.0084) | (0.011) |
| American Indian | 0.00027 | -0.00034 |
| | (0.00048) | (0.00069) |
| Pacific Islander | -0.00015 | -0.0014* |
| | (0.00046) | (0.00068) |
| White | -0.0026 | -0.0050 |
| | (0.010) | (0.013) |
| N. Obs. | 488 | 488 |

Notes: Bootstrapped standard errors, clustered at the school level are reported in parentheses. Statistical significance levels: + $p< 0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$. Each row reflects a separate regression where the outcome is either school enrollment or the proportion of students in the indicated demographic group.

**Generalized Difference-in-Differences/Comparative Interrupted Time Series Results**

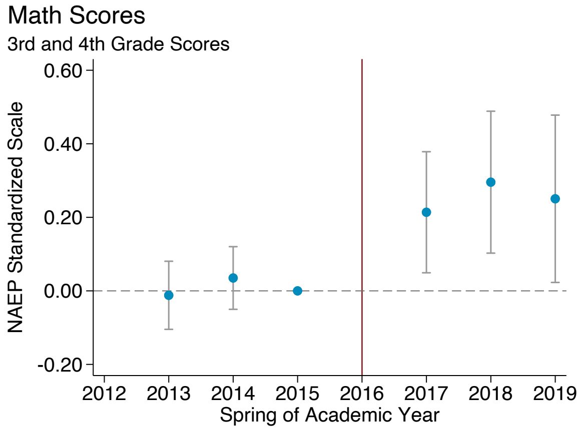
**Table C2. Impact of Additional Funding from SBB via WSF on Elementary School Student Outcomes**

| | Pre-Pandemic Effect | | Effect Across All Years | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *A. Test Scores: Grades 3 and 4* | | | | |
| Math (NAEP Standardized Scale) | 0.22** | 0.21** | 0.15+ | 0.14+ |
| | (0.080) | (0.081) | (0.086) | (0.081) |
| ELA (NAEP Standardized Scale) | 0.20** | 0.20** | 0.15* | 0.14* |
| | (0.070) | (0.070) | (0.064) | (0.065) |
| N. Observations | 427 | 427 | 549 | 549 |
| *B. Attendance: Grades K to 4* | | | | |
| Total Absence Rate (%) | -0.025 | -0.0050 | 0.070 | 0.075 |
| | (0.38) | (0.37) | (0.35) | (0.35) |
| Unexcused Abs. Rate (%) | -0.075 | -0.076 | 0.039 | 0.032 |
| | (0.36) | (0.36) | (0.30) | (0.30) |
| N. Observations | 488 | 488 | 671 | 671 |
| *C. Discipline: Grades K to 4* | | | | |
| In-School Suspension Proportion | -0.0033 | -0.0036 | -0.0067 | -0.0067 |
| | (0.0043) | (0.0040) | (0.0048) | (0.0047) |
| Out-of-School Suspension Proportion | -0.020 | -0.019 | -0.013 | -0.012 |
| | (0.012) | (0.012) | (0.013) | (0.013) |
| N. Observations | 488 | 488 | 671 | 671 |
| Controls for Student Demographics | No | Yes | No | Yes |
| N. Elementary Schools | 61 | 61 | 61 | 61 |

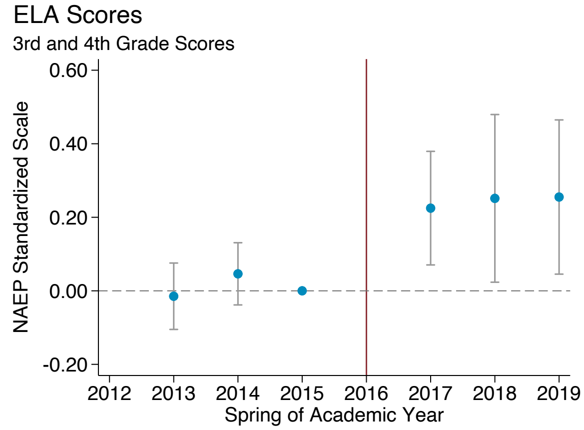
Notes: Robust standard errors, clustered at the school level, are reported in parentheses. Statistical significance levels: + $p< 0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$. Observation counts are at the school-by-year level. Results with controls adjust for student race and ethnicity and the logarithm of enrollment.

**Figure C2**. Event Studies based on generalized difference-in-differences model

**Panel A. Test Score Outcomes**

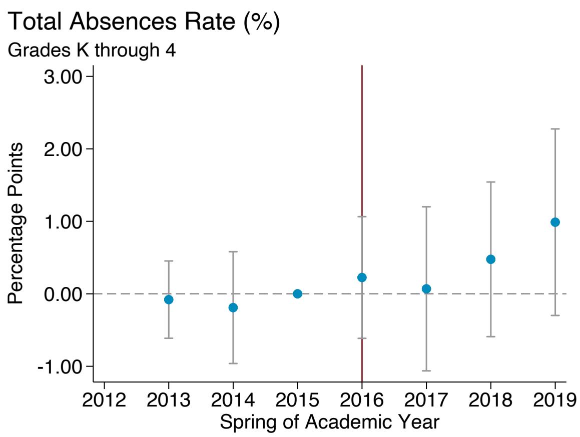



(a)

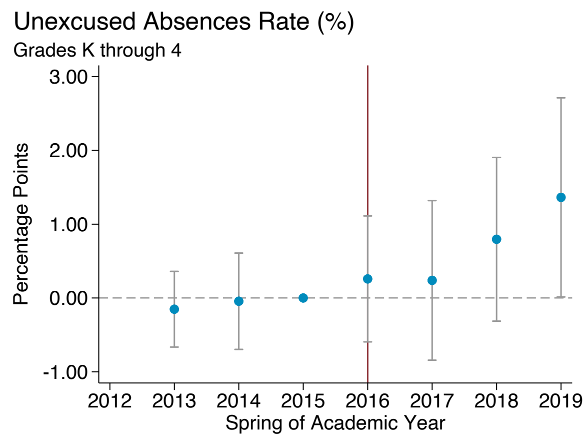



(b)

**Panel B. Attendance Outcomes**

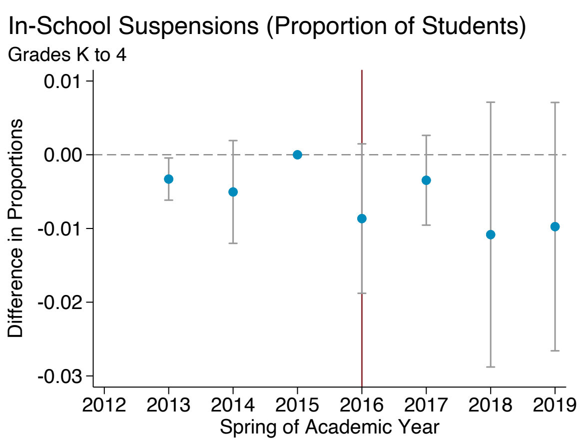



(a)

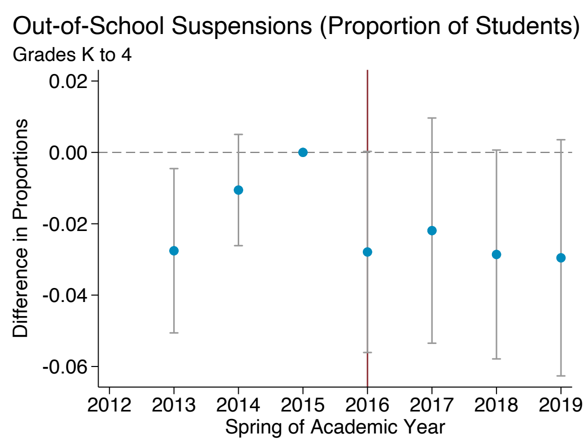



(b)

**Panel C. Discipline Outcomes**




(a)




(b)

Notes: Event studies based on CITS model with leads and lags. Error bars are 95 percent confidence intervals.

## Test Score Missingness

**Table C3.** Robustness Check: Test Score Missingness/Suppression before COVID-19

| | SDID Estimate |
|---|---|
| All Students | 0.0061 |
| | (0.0090) |
| Ever Economically Disadvantaged (FRPL) | -0.00036 |
| | (0.010) |
| Ever Economically Disadvantaged (DC) | -0.0030 |
| | (0.012) |
| Ever English Language Learner (ELL) | -0.0073 |
| | (0.024) |
| Ever Special Education (SPED) | 0.0016 |
| | (0.018) |
| Female | 0.0076 |
| | (0.012) |
| Asian | 0.052 |
| | (0.041) |
| Black | -0.011 |
| | (0.015) |
| Hispanic | 0.017 |
| | (0.020) |
| White | 0.0034 |
| | (0.017) |

Notes: Bootstrapped standard errors, clustered at the school level, are reported in parentheses. Statistical significance levels: + $p < 0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$. Each row reflects a separate regression where the outcome is the proportion of missing test scores when test score data are matched to school-level demographic files. American Indians and Pacific Islanders are not shown due to an insufficient number of observations.

## Test Score Scaling

**Table C4. Impact of Additional Funding from SBB via WSF on Elementary School Student Outcomes (ALTERNATE SCALING)**

| | Pre-Pandemic Effect | | Effect Across All Years | |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| *A. Test Scores: Grades 3 and 4* | | | | |
| Math (Standardized: State Distribution) | 0.14* | 0.13+ | 0.11* | 0.11+ |
| | (0.057) | (0.069) | (0.054) | (0.067) |
| ELA (Standardized: State Distribution) | 0.10* | 0.098 | 0.087+ | 0.11+ |
| | (0.049) | (0.061) | (0.051) | (0.056) |
| N. Observations | 427 | 427 | 549 | 549 |
| *C. Test Scores: Grades 3 and 4* | | | | |
| Math (Standardized: District Distribution) | 0.14* | 0.13+ | 0.11* | 0.11+ |
| | (0.056) | (0.067) | (0.054) | (0.065) |
| ELA (Standardized: District Distribution) | 0.096* | 0.096+ | 0.083+ | 0.10+ |
| | (0.046) | (0.057) | (0.048) | (0.052) |
| N. Observations | 427 | 427 | 549 | 549 |
| Controls for Student Demographics | No | Yes | No | Yes |
| N. Elementary Schools | 61 | 61 | 61 | 61 |

Notes: Bootstrapped standard errors, clustered at the school level, are reported in parentheses. Statistical significance levels: + $p < 0.10$; * $p<0.05$; ** $p<0.01$; *** $p<0.001$. Observation counts are at the school-by-year level. Results with controls adjust for student race and ethnicity and the logarithm of enrollment.