Annenberg Institute
BROWN UNIVERSITY

EdWorkingPaper No. 26-1452

The Effect of School-Based Health Centers on Adolescent Mental Health and Behavior

Anjali Pai
Brown University

Adolescent mental health has experienced significant declines in the past decade, yet take-up of mental health services has remained low among adolescents. This paper examines whether localized access to mental health services has meaningful impacts on adolescent mental health and behavior. I study the effect of school-based health centers — full-service clinics located in K-12 schools that offer physical, mental, and reproductive health services at low to no cost — on suspensions and dropouts, two outcomes that may be linked to untreated mental health issues. Using a staggered difference-in-differences analysis that leverages the timing of health center openings in California and a propensity-score matched control group, I find that the opening of a new school-based health center decreases suspension rates by around 1.1 percentage points, an 18% decrease relative to matched schools. The effects are concentrated in middle schools and high schools which experience decreases of 4.9 and 0.8 percentage points respectively. This decrease is primarily driven by a decrease in suspensions from "disruptive behavior", rather than weapon possession, violence, or drug use. In addition to the extensive margin decrease in share of students suspended, SBHCaccess also decreases number of suspensions per suspended student by 10-26% depending on the grade level. There is no effect on dropout rates, indicating that the decline in suspensions is unlikely to be caused by crowd-out of delinquent behavior by an increase in dropping out. These results suggest large potential impacts of school-based health centers on treating behavioral issues that cause suspendable behavior, especially in middle and high schools, and motivates further study of SBHCs as a tool for improving access to mental health services for adolescents in low-income communities.

VERSION: April 2026

Suggested citation: Pai, Anjali. (2026). The Effect of School-Based Health Centers on Adolescent Mental Health and Behavior. (EdWorkingPaper: 26-1452). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/aga1-ys09

The Effect of School-Based Health Centers on Adolescent Mental Health and Behavior

Anjali Pai *

April 15, 2026

ABSTRACT

Adolescent mental health has experienced significant declines in the past decade, yet take-up of mental health services has remained low among adolescents. This paper examines whether localized access to mental health services has meaningful impacts on adolescent mental health and behavior. I study the effect of school-based health centers — full-service clinics located in K-12 schools that offer physical, mental, and reproductive health services at low to no cost — on suspensions and dropouts, two outcomes that may be linked to untreated mental health issues. Using a staggered difference-in-differences analysis that leverages the timing of health center openings in California and a propensity-score matched control group, I find that the opening of a new school-based health center decreases suspension rates by around 1.1 percentage points, an 18% decrease relative to matched schools. The effects are concentrated in middle schools and high schools which experience decreases of 4.9 and 0.8 percentage points respectively. This decrease is primarily driven by a decrease in suspensions from "disruptive behavior", rather than weapon possession, violence, or drug use. In addition to the extensive margin decrease in share of students suspended, SBHC-access also decreases number of suspensions per suspended student by 10-26% depending on the grade level. There is no effect on dropout rates, indicating that the decline in suspensions is unlikely to be caused by crowd-out of delinquent behavior by an increase in dropping out. These results suggest large potential impacts of school-based health centers on treating behavioral issues that cause suspendable behavior, especially in middle and high schools, and motivates further study of SBHCs as a tool for improving access to mental health services for adolescents in low-income communities.

*Annenberg Institute, Brown University. anjali_pai@brown.edu. Thank you to the California School-Based Health Alliance (CSBHA) and the California Department of Education (CDE) for providing data and contextual information for this paper. This work would not have been possible without the support of Lisa Eisenberg, Amy Ranger, Amy Blackshaw and others at the CSBHA; Maryjane Puffer, Marsha Ellis, and Alex Zepeda at the Los Angeles Trust; Samira Soleimanpour at UCSF; and Jonathan Isler at the CDE. Finally, I am thankful for the feedback of Julian Betts, Prashant Bharadwaj, Itzik Fadlon, and Sally Sadoff, and seminar participants at UC San Diego and Abt Global.

1 Introduction

Adolescent mental health has experienced significant declines in the past decade. The Center for Disease Control and Prevention (CDC) reports that the fraction of adolescents aged 12-17 across the U.S. reporting "persistent feelings of sadness or hopelessness" increased from 26% in 2009 to 36.7% in 2019 (CDC, 2022). By 2021, one year into the COVID-19 pandemic, this share had increased to 44%, marking a 20% increase in just two years (CDC, 2022). Untreated mental health issues can directly impact an individual's adolescent and later-life outcomes. A large body of research in psychology suggests that for adolescents, untreated mental health issues may manifest in disruptive behavior and inattention (Garland et al. (2010), McLeod et al. (2012)). Recent quasi-experimental work suggests that behavioral disorders such as Attention Deficit Hyperactivity Disorder (ADHD), which affects around 11.4% of children, may have negative effects on test scores and education attainment (Currie and Stabile, 2006). Moreover, 50% of mental health illnesses identified in adults are documented as having begun before the age of 14 (Kessler et al., 2005), suggesting that untreated mental health issues in adolescence may have negative spillovers in adulthood.

Despite high rates of reported depression among adolescents, take-up of mental health services for this group has remained relatively low. The 2021 National Survey on Drug Use and Health, administered annually by the Substance Abuse and Mental Health Services Administration (SAMHSA), reports that in a sample of 5 million adolescents aged 12 to 17 who reported having a "major depressive episode" in the past year, only 40% reported receiving treatment for depression (SAMHSA, 2023). The gap between reported need for mental health services and service utilization may be the result of a combination of supply-side and demand-side barriers. On the supply side, the gap between available mental health services and documented mental health needs has been widening since 2016, with the share of need met decreasing from 56% in 2016 to around 27% in 2023 (BHWF, 2023).

On the demand side, there are three primary documented barriers to take-up of mental health services: physical distance, financial cost, and societal stigma (NCBH, 2019). This paper evaluates a model of healthcare provision that is theoretically well designed to overcome all three of these demand-side barriers: school-based health centers. School-based health centers (SBHCs) are student-focused health clinics that are located in or near a K-12 school, provide services at low or no cost to students, and are integrated into the school with which they are associated. These three features have the potential to address distance, cost, and stigma barriers respectively, positioning SBHCs to improve take-up of mental health services for adolescents. While the first SBHCs targeted gaps in physical health in low-income communities, many centers added mental health services in the early 2000s as awareness of mental health issues increased (Flaherty and Osher, 2003). By 2017, nearly 65% of SBHCs nationwide reported employing behavioral health specialists in addition to primary care providers.[1] While school-based health centers have been a feature of many low-income

[1]These statistics come from the 2016-2017 National Census of School-Based Health Centers (Love et al., 2019).

schools for decades, there is limited evidence on their effectiveness. Causal evidence on the mental health impacts of SBHCs has the potential to inform ongoing state and federal discussions about effective allocations of funding to treat adolescent mental health.

This paper studies the impact of SBHC-access on suspensions and dropouts, two outcomes that may result from behavioral issues or poor mental health (So et al. (2024), Dupéré et al. (2018), Cohen et al. (2023)). These outcomes are the best-available proxies for mental health in this context, given that mental health assessments (eg. PHQ-9 for depression and GAD-7 for anxiety) and student-reported mental health are not typically collected for all students at a consistent frequency. Using a novel panel dataset that links data on SBHC openings in California between 2011 - 2019 to school-level outcomes, I present the first quasi-experimental evidence that access to an SBHC with mental health services reduces the rate of suspensions that are caused by behavioral issues.

A key challenge in identifying the impact of SBHCs is selection into opening a center. Due to high overhead and operational costs, the ability to open a center often relies on strong partnerships between a school, school district, and community organizations. As a result, the average school with an SBHC may be meaningfully different from a school without one. To address this challenge, I use propensity-score matching to select a set of control schools with the highest "expected likelihood" of being similar on observable and unobservable characteristics to schools that open an SBHC. I then use a staggered-event difference-in-differences model that isolates the effect of access to an SBHC by comparing outcomes for schools that open an SBHC and their matched comparison schools, before and after the new center opens. To correct for potential negative weighting bias in staggered-event difference-in-differences, all main specifications use the adjusted estimator from Callaway and Sant'Anna (2021).

I find that on average, access to an SBHC reduces suspension rates for a treated school by around 1.1 percentage points. This is a large magnitude effect, amounting to an 18% decrease from the control group baseline rate. These effects are primarily concentrated at middle schools, where the decrease in suspension rates may be as large as 4.8 percentage points, and high schools, where suspension rates decrease by around 0.9 percentage points. In contrast, there are no effects of SBHCs on overall suspension rates at elementary schools. Exploring the mechanisms behind these treatment effects reveals that the decrease in suspension rates is driven by a decrease in "defiance suspensions", suspensions that are caused by disruptive or defiant student behavior. Comparatively, there is no significant change in suspensions resulting from weapon possession, violence, or drug use.[2] This suggests that the decrease in suspensions may be indicative of a decrease in disruptive behavior, which is a common symptom of many underlying psychological disorders in adolescents (Garland et al., 2010). Heterogeneity analyses suggest that the impacts on suspension rates may be larger for male students, students from non-White racial backgrounds, and students with school-

[2]The California Department of Education defines six total categories of offenses that may lead to a suspension: (1) Violent Incident (Injury); (2) Violent Incident (No Injury); (3) Weapons Possession; (4) Illicit Drug-Related; (5) Other; (6) Defiance-only. Categories (1)-(4) consist of federal offenses and Category 5 consists of offenses under state law that are not against federal law. Appendix Table E.1 lists the offenses in each category.

documented disabilities.

I also find a notable decrease in the number of repeat suspensions per suspended student, both for the overall sample and for middle schools. If repeat suspensions are an indication of persistent behavioral issues, this suggests that SBHCs could be treating root causes of disruptive behavior, rather than serving as a temporary "alternative" to suspension for disruptive students. Interestingly, there is a decrease in repeat suspensions at the elementary school level despite the lack of change in overall suspensions. This suggests that while SBHC-access may not affect the overall likelihood of suspension for elementary-aged students, it may still have a positive intensive margin effect for students with a high propensity for suspension. There is no consistent effect of SBHC openings on dropout rates and tight 95% confidence intervals rule out changes larger than 0.5 percentage points. This indicates that the decrease in suspension rates is unlikely to be explained by an increase in dropout rates "crowding out" suspensions. These treatment effects are robust to use of a standard twoway fixed-effect model, as well as to theoretically-appropriate alternate control groups.

Finally, to validate the use of suspension rates and dropout rates as proxies for mental health, I link data on these outcomes for a subset of schools to data from the California Healthy Kids Survey (CHKS), a biannual survey on school climate, risky behavior, and mental health. Controlling for cross-year and cross-school differences, I find that higher rates of student-reported depression and lower rates of "feeling connected" with peers and teachers are correlated with higher suspension rates, but not necessarily with higher dropout rates. This further supports the theory that changes in suspension rates may be capturing improvements in adolescent mental health, and aligns with the zero effect I find for dropout rates.

This paper contributes to a limited existing literature on the impacts of school-based health centers, which is predominantly descriptive and non-causal (Geierstanger et al. (2004), Katz (2020)). In addition to providing the first quasi-experimental evidence that SBHC-access decreases suspension rates, the methods in this paper offer a new approach to overcoming the challenges in identifying causal impacts of SBHCs more broadly. While previous studies point to positive relationships between school-based health centers and attendance, academic performance, physical health, and graduation rates, these results come from either cross-sectional comparisons between schools with and without SBHCs (Kisker and Brown (1996), Santelli et al. (1996), Paschall and Bersamin (2018)), within-school comparisons between students who utilize SBHC services and students who do not utilize these services (Kerns et al. (2011), Walker et al. (2010), McCord et al. (1993)), or single-school program evaluations (Gall et al. (2000), Warren and Fancsali (2000)).

One exception is Lovenheim et al. (2016), which uses data from a national survey of SBHCs and a staggered-event difference-in-differences approach to study the effect of SBHCs on teenage fertility and high-school dropout rates. They find that the first opening of an SBHC in a county leads to a 1.3% decrease in the teenage fertility rate and has no identifiable effect on high school dropout rates. These results suggest large positive effects of SBHCs on reproductive health and null effects for dropout rates that align with my findings. The current paper builds on the empirical design used by Lovenheim et al. (2016) by adding a propensity-score matched control group, which allows for analysis at the school level. In analyzing behavioral outcomes, this paper also expands our understanding of the range of impacts SBHCs might have on student well-being. This paper also complements recent work from Komisarow and Hemelt (2022) that finds that access to a "school-based telemedicine center" in North Carolina decreases the likelihood of chronic absenteeism by 2.5 percentage points and decreases the likelihood of having at least one violent or weapons-related infraction by 40-47% of the baseline mean. While telemedicine is meaningfully different from the in-person services provided by SBHCs, Komisarow and Hemelt's results support the theory that comprehensive school-based health services can have a large impact on mental-health linked outcomes.

Finally, my work contributes to a smaller literature on school-based approaches to mental health treatment. Causal work on this topic has primarily focused on the impact of elementary school counselors on students' behavioral outcomes. Carrell and Carrell (2006) and Carrell and Hoekstra (2014) find that increasing the ratio of counselors to students in elementary schools reduces disciplinary incidents. Similarly, Reback (2010b) finds that state reforms that improve the ratio of counselors to students in elementary schools reduce teacher-reported incidents of delinquent behavior. Finally, Reback (2010a) shows that increased funding for elementary-school counselors has a significant impact on decreasing disciplinary infractions. An important difference between elementary school counselors and the mental health professionals staffed in school-based health centers is that school counselors do not typically provide on-site therapy or formal mental health treatment, and will instead refer students to outside services if mental health needs are identified. The current paper broadens this literature by studying a program that treats mental health issues on-site and is accessible to a broader age-range of students beyond elementary school. The decrease in suspensions identified in this paper is strongest for middle schools and high schools, suggesting that school-based health centers may be well-targeted for middle-school and high-school aged youth, who have been understudied in the literature.

The remainder of this paper proceeds as follows: Section 2 provides background on school-based health centers in California; Section 3 provides an overview of the data and sample construction process; Section 4 discusses the empirical strategy and identifying assumptions; Section 5 shows results from the primary specifications; Section 6 discusses the robustness of the main results; Section 7 shows heterogeneity of impacts by demographic characteristics; Section 8 discusses policy implications; and finally, Section 9 explores avenues for future research and concludes.

2 Background

This paper focuses on school-based health care in California, a state that serves over 5.8 million K-12 students.[3] The reported rate of mental health issues for adolescents in California has closely

[3]Source: *About School-Based Health Centers*, California School Based Health Alliance

followed national trends in the past decade, with nearly 45% of youth aged 12-17 reporting that they struggled with mental health issues in 2021 (Wright et al., 2021). As of 2025 California had 291 active school-based health centers (CSBHA, 2025). The California School-Based Health Alliance (CSBHA) defines school-based health centers as "student-focused health centers or clinics that are located on or near a school campus, are organized through school, community, and health provider relationships, and provide age-appropriate, clinical health care services onsite by qualified health professionals" (CSBHA, 2022b). These centers commonly serve as an alternative to school nurses, who traditionally assess students for health problems, deliver basic health services such as immunizations and insulin, and provide health and nutrition education, but are unable to treat more serious health problems, provide psychological counseling or therapy, or prescribe medication (CSHCA, 2010). School-based health centers are typically staffed by at least one physician or nurse practitioner with the ability to write prescriptions, along with a combination of residents, medical assistants, and nurses who help provide physical, mental, and reproductive health services on site. In 2025, 80% of California SBHCs reported offering mental health services in addition to primary care (CSBHA, 2025).

The services provided by SBHCs are designed to be low cost, age-appropriate, and offered on the school site. This model of healthcare provision is well-equipped to overcome the three most common barriers to take-up of mental health services: physical distance, financial cost, and societal stigmas around mental illness (NCBH, 2019). The "in-school" location and low cost of treatment directly address the distance and financial cost barriers. These features are intended to enable all students to access the services regardless of ability to pay; however, maintaining this model can lead to high operation costs for SBHCs. As a result, most SBHCs fund their operations through a combination of state and local grants, and community partnerships (Katz, 2020). In California, one of the few states that provides no centralized state funding for the opening and maintenance of SBHCs, 65% of centers are operated by Federally-Qualified Health Centers (FQHCs), which are local federally-funded non-profit healthcare organizations. FQHCs receive favorable Medicaid reimbursement rates, allowing them to offer services to low-income individuals at low costs (CSBHA, 2023). The remaining centers are funded by a mix of local education agencies (27%), local hospital or universities (3%), local public health departments (3%), or other community-based organizations (2%) (CSBHA, 2022a).

The provision of age-appropriate and student-focused services that are accessible to all students has the potential to decrease the stigma associated with utilizing mental health services within a school. Moreover, in California, school-based health centers prioritize treating and screening as many students as possible. This is reflected in the three-tier model of mental health services followed by most SBHCs. SBHCs that offer mental health treatment will offer Tier 1 ("Universal Prevention") services at baseline, and offer Tier 2 ("Targeted Early Intervention") and Tier 3 ("Intensive Intervention") depending on available staff and funding.[4] For example, in the Madera

[4] Appendix Figure A.2 outlines the types of services included in each tier.

South School-Based Health Center in Madera County, Tier 1 services include a program that trains students to identify mental health concerns in their peers and provide “peer counseling”, while more intensive services include one-on-one counseling with licensed clinical social workers and referrals to external behavioral-health practitioners for more intensive care.[5] The emphasis on treating as many students as possible makes these centers well-poised to “normalize” the concept of mental health care in the schools they serve.

The adolescent-focused staffing and practices in SBHCs may also give them an advantage over community health clinics in treating youth mental health issues. A 2003 retrospective cohort study from Juszczak et al. (2003) shows that adolescents with access to SBHCs had higher visit rates than students who only had access to community health centers. Moreover, for those students who did not have access to an SBHC and used only community health centers, 97% of visits were for medical services. Comparatively, for students who chose to use an SBHC, at least 34% of visits were for mental health services. While these descriptive statistics do not account for selection bias, they suggest that the mental health services provided by SBHCs may be better-suited for adolescents than those provided by community health centers. While I am unable to measure take-up of services in this paper, discussions with administrators at school-based health centers in California indicate that demand for mental health services in SBHCs is high, suggesting that students are utilizing them when given the option.

3 Data

3.1 Data Sources

The analysis in this paper utilizes three primary data sources: (1) data on school-based health centers from the California School-Based Health Alliance,[6] a non-profit organization that provides support and resources for school-based health centers operating in California; (2) annual suspension rates, dropout rates, and student demographics at the school level from the California Department of Education (CDE); and (3) annual student-reported mental health, wellness, and behavior from the California Healthy Kids Survey (CHKS). The data on SBHC openings is restricted to centers that were operational in August 2022 when the data were compiled, and includes 286 SBHCs spanning 34 counties and 120 school districts in California. For each SBHC, the data includes the opening date, center name, name of the associated school, center address (split into street address, city, county, and zip code), a checklist of services provided, and a set of center-reported characteristics.[7]

[5] This case study come from a report compiled by Lisa Eisenberg, formerly of the California School Based Health Alliance.

[6] The California School-Based Health Alliance is a state affiliate of the National School Based Health Alliance.

[7] The center characteristics include the “type” of SBHC (on-site, off-site, or mobile van), the primary sponsoring organization, a list of schools served, and a list of populations served. The variable describing the list of schools served by the SBHC does not provide a reliable measure of SBHC service area since it is generated from an open-ended text response and is relatively sparse in the data.

The data on suspension rates, dropout rates, and school characteristics come from the CDE's public database. The CDE publishes annual school-level suspension rates for all public elementary, middle, and high schools, and dropout rates for high schools.[8] Data on suspension rates is available for the 2011-12 academic year through the 2021-22 academic year. Dropout rates are available for a more limited set of academic years (2010-11 to 2016-17). For both outcomes, the data is disaggregated by race and gender. Finally, to validate the relationship between student mental health, suspension rates, and dropout rates, I incorporate data from the California Healthy Kids Survey (CHKS), an anonymous student survey with well-validated psychometric properties that is administered in a sample of elementary, middle and high schools across California at an annual or bi-annual frequency (Hanson and Kim (2007), Mahecha and Hanson (2020)).[9]

3.2 Constructing the Analytic Dataset

The California Department of Education uniquely identifies schools in its data using a "County-District-School" (CDS) code. In order to match school-based health centers to schools, each SBHC must be assigned to a "principal school", defined as the main school served by that SBHC. There are three primary types of SBHCs in California: on-site, off-site (which includes telehealth-only centers), and mobile vans. For on-site SBHCs, the "principal school" is the school in which the SBHC is physically located. For off-site and mobile vans, the "principal school" is the school in closest geographic proximity to the SBHC. In the universe of SBHC-openings in California, 73% are on-site, 13% are off-site or "telehealth-only", and 14% are mobile vans. SBHCs are matched to schools using iterative fuzzy string matching on the complete address and school name, which are reported in both datasets but do not follow a single standardized format.[10] At the end of the process, all 286 SBHCs are matched to a principal school. Within the matched sample, the average similarity scores for address similarity are 0.91 for on-site SBHCs, 0.71 for off-site SBHCs, and 0.73 for mobile vans. Appendix C.1 provides further details about the assignment process.

Prior to merging the SBHC data with outcomes data, each SBHC is assigned an "academic year" of opening, defined as the closest academic year during which the SBHC was open for at least one month. For example, an SBHC that opened in April of 2011 would be assigned an academic opening year of 2010-2011, but an SBHC that opened in June would be assigned the academic year 2011-2012. Since this rule treats SBHCs that opened one month before the end of the school year the same as SBHCs that opened at the start of the school year, in the main analysis I weight

[8] In schools with a combined middle school and high school structure dropout rates may be available for grades 7-12.

[9] These data were acquired through a partnership with the California Department of Education and the support of Jonathan Isler at the CDE. The available survey data from the CHKS extends back to 1998, however when used in combination with suspension and dropouts data the analysis is restricted to 2012-2019.

[10] In the SBHC data, the "school name" comes from a survey question where SBHCs were asked to report the name of the "primary school served". Since school names may be repeated in different districts and the data does not contain information on school district, the "school name" field on its own is insufficient for matching. The address is split into four fields: street address, zip code, city, and county. In the CDE data, school name and complete school address are standardized fields available for all California public schools.

estimates in the year of the opening by the fraction of the academic year for which the SBHC was open. The sample of SBHCs is restricted to only on-site SBHCs that opened between the 2012-13 and 2018-19 academic years, inclusive. The lower bound on opening years is one year after the earliest year of data on suspension and dropout rates (the 2011-2012 academic year).[11] The upper bound excludes schools that opened a center during or after the COVID-19 pandemic, since the choice to open an SBHC after 2019 may be a *direct response* to increasing mental health concerns during the pandemic and therefore may be concurrent with other policies that target adolescent mental health.

There are several reasons to restrict to on-site SBHCs. First, if school-based health centers improve take up of mental health services in part by reducing *physical distance* barriers, then those effects should be strongest for SBHCs that are directly located in a school building. On-site SBHCs may also be more closely integrated with the broader school community, improving the potential to decrease stigma. Second, a comparison of on- and off-site SBHCs presented in Appendix Table A.1 reveals that 83% of on-site SBHCs offer mental health services, compared only 48% of non-on-site SBHCs. On-site SBHCs are also more homogeneous in their community partnerships. Over 80% of on-site SBHCs are sponsored either by a Community Health Center (CHC) or a school system while off-site SBHCs are sponsored by a more diverse set of organizations. This suggests that the motivations for opening on-site SBHCs may be more similar than those motivating off-site SBHCs. Finally, Appendix Figure A.1 reveals that on-site SBHCs are more likely to report only serving their "principal school" and less likely to report serving the entire district or multiple districts. This suggests that restricting the sample of treated students to those attending the "principal school" may be more reasonable for on-site SBHCs than for off-site SBHCs.

Finally, all schools located in Los Angeles Unified School District (LAUSD) are omitted from the analytic sample due to a district-wide ban on "willful defiance" suspensions that was implemented in 2013 (Rott, 2013). Since this change in suspension policy is coincidental with the study window, it may be difficult to disentangle changes in suspension rates over this period in LAUSD that are caused by the suspension policy from those caused by SBHC openings. LAUSD is also the singular school district in California to implement such a policy in this time period, raising concerns that SBHCs opening in LAUSD after 2013 may have different goals or motivations than SBHCs opening in other districts. Robustness checks presented in Appendix B.2 show that the effect on overall suspension rates does not meaningfully change when LAUSD is included in the sample. The estimated decline in defiance suspensions is actually smaller than in my primary specifications, which may be the result of the inclusion of control schools from LAUSD that effectively banned defiance suspensions over the study period.

Figure 1 shows the distribution of on-site SBHC openings from 1967 - 2023, with the analytic sample marked by the yellow-shaded bars.[12] Importantly, there is significant year-to-year variation

[11] SBHCs opening in 2011 cannot be included in the sample since they would have zero years of outcomes prior to the opening.

[12] For dropout rates, which are only available from 2010-11 through 2016-17, the analysis is restricted to SBHC

and no visible time trend in the number of SBHC openings during the analysis period.

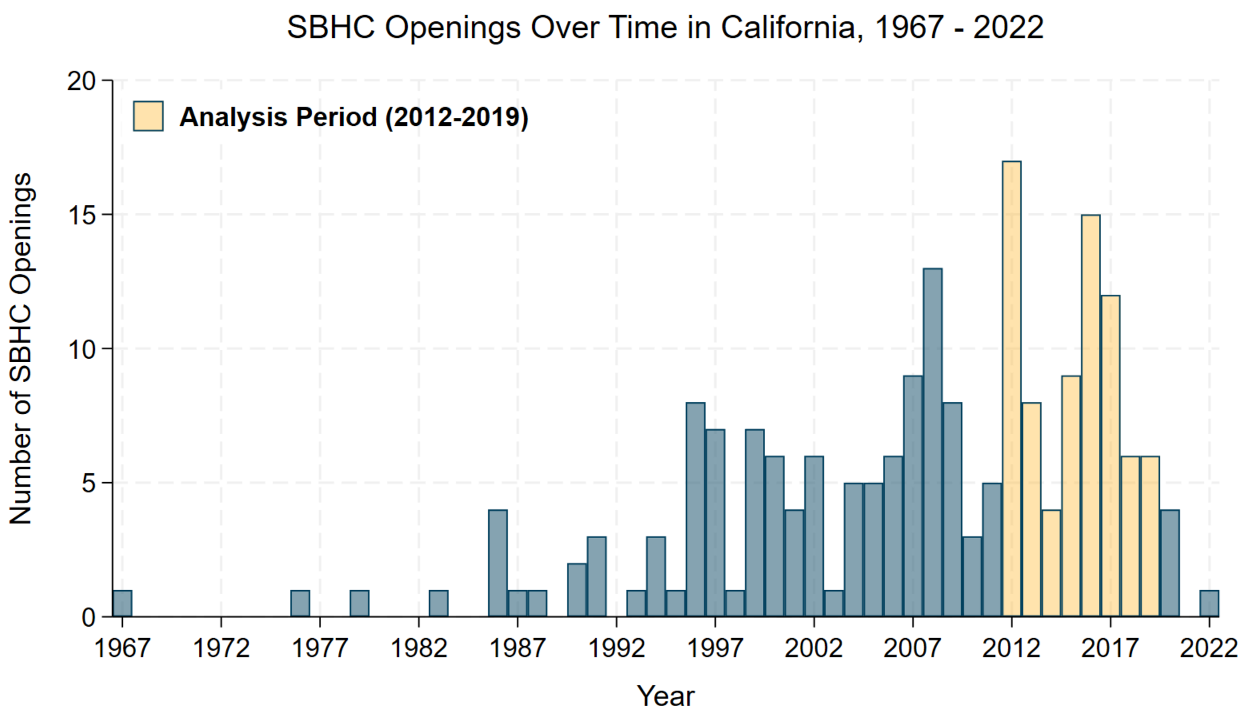

Figure 1: This graph shows the distribution of SBHC openings between 1967-2023 for the set of SBHCs that were active in California as of 2022. The x-axis shows the range of opening years while the y-axis shows the total number of SBHCs that opened in that year. The yellow shaded bars mark the set of years covered by the California Department of Education's suspension and dropout data.

3.3 Defining Outcome Measures

The primary outcomes examined in this paper are suspension rates and dropout rates. For both outcomes, counts are available for all students, as well as disaggregated by gender and race. Suspension counts are additionally available disaggregated into six categories of suspendable offenses defined by the CDE: violent incidents (with injury), violent incidents (with no injury), weapons possession, illicit drug possession or sale, defiance-only incidents, and other offenses.[13] Dropout counts are available primarily for grades 9-12. For certain schools that serve a larger span of grades, separate dropout counts are available for grades 7 and 8. The dropout count for grades 9-12 is treated as the primary count for each school. If a count for grades 9-12 is unavailable, the combined rate for grades 7 and 8 is imputed. To calculate school-level suspension and dropout rates, I divide each count by the relevant school-level cumulative enrollment count. Since cumulative enrollment counts are also available disaggregated by demographics, subgroup-specific rates are calculated using the enrollment count for the respective subgroup.

openings between 2012 and 2016.

[13]The specific offenses included in each of these categories are outlined in Appendix Section E.

A secondary set of outcomes used in this paper comes from the California Healthy Kids Survey (CHKS). CHKS is a part of the California School Climate, Health, and Learning Surveys (CalSCHLS) system, which was designed to provide schools with "quality local data which can be used to improve student academic performance and social-emotional, behavioral, and physical health of all youth" (CalSCHLS, 2026). The restricted CHKS data used in this study is anonymous and contains student-level responses to every question from the core CHKS module, as well as from any supplementary modules the student completed. The data also provides information on students' demographic characteristics, school, district, and grade. I restrict the data to students in grades 7 and 9 to omit grades that are not consistently sampled across all districts and schools.[14]

The primary use of the CHKS data in this paper is to validate suspension rates and dropout rates as reasonable measures of mental health. To do so, I construct four psychometrically validated indices to measure mental health and school climate using a method proposed by researchers at WestEd, the agency that developed the California Healthy Kids Survey. The indices capture the following metrics: (1) Caring Staff-Student Relationships; (2) School Connectedness; (3) Delinquency; (4) Substance Use at School.[15] The first two indices are designed to capture "positive mental health", and include sentiments such as "At school, there is a teacher or adult who really cares about me" and "I am happy to be at this school". The latter two indices capture student reports of their own delinquent behavior and substance use, and are more likely to capture behaviors that may be linked to "negative mental health". Appendix Section E.2 describes the data cleaning and index construction process for the CHKS outcomes, as well as the specific questions included in each index. Finally, to more directly measure student-reported mental health, I use the following questions: (1) During the past 12 months, did you ever seriously consider attempting suicide?; During the past 12 months, did you ever feel so sad or hopeless almost every day for two weeks or more that you stopped doing some usual activities?[16] For each question the relevant outcome is the share of students in a school who responded "yes".

4 Empirical Design

To identify the effect of opening a school-based health center on student outcomes, I use variation in the timing of SBHC openings for a staggered-event difference-in-differences model. The simplest version of this design would compare schools that open an SBHC to schools that do not open an

[14] Although districts are encouraged to administer the CHKS to students in grades 5, 7, 9, and 11, they are only required to administer it to grades 7 and 9.

[15] Each index is a weighted average of the responses to a set of questions. I use the exact questions and weights suggested in Mahecha and Hanson (2020), the paper that proposes these measures and validates their psychometric properties.

[16] While the two questions used to measure student-reported mental health have not explicitly been tested for psychometric validity, they are variations of questions that have been used in other trusted health surveys such as the Youth Behavioral Risk Factor Survey, suggesting that they may be acceptable survey measures for mental health issues. The data for the first question is available going back to 2010, while the data for the second question is only available after 2014, which is when it was added to the core module.

SBHC in the years before and after the opening; however, the validity of this design relies on the assumption that behavioral outcomes in treated and control schools would have evolved in parallel in absence of the school-based health center. This assumption might be violated if school or district characteristics are correlated with both the decision to open an SBHC and student behavior. Since the decision to open an SBHC is often driven by school, community, and district partnerships, schools that open a center may be meaningfully different from schools that do not.[17]

To address non-randomness in the decision to open a center, control schools are selected using a propensity score matching procedure that matches schools that open an SBHC in year y to up to three schools in the same gradespan (elementary school, middle school, and high school) that never open a center but have a similar predicted likelihood of opening an SBHC in year y. The propensity score is based on two school-level factors that should be theoretically strong predictors for the opening of an SBHC: the fraction of low socioeconomic status students and school size. Conversations with SBHC practitioners suggest that two of the primary goals of SBHCs are to serve as many students as possible and fill the gap in healthcare provision in communities that lack sufficient medical services. As a result, we might expect centers to open in lower-income schools and in schools that serve a larger population of students. Larger schools may also have larger buildings that are better equipped to house an on-site clinic. Following standard practice in the education literature, I proxy for the fraction of low socioeconomic-status students with the fraction of students eligible for Free-and-Reduced-Price meals and for school size with total school enrollment.

A small body of literature suggests that propensity-score matching models are sensitive to the choice of predictors, model specification, and the choice of control groups (Smith and Todd, 2005). This sensitivity can be addressed by matching within the same "local labor market" and using consistently measured dependent variables in the treated and control schools (Heckman et al., 1997).[18] To identify control schools in a similar "local labor market" to schools with SBHCs, potential matches are selected from districts that are "amenable to having SBHC", as measured by the opening of at least one SBHC in the district after 2006 (five years before the first opening in the sample) but before the year of matching. While the optimal local labor market in this setting would be a school-district or Local Education Agency (LEA) there are practical concerns with matching within those boundaries. First, limiting to within-district matches is likely to favor large school-districts with more potential controls. This poses a threat to external validity outside of large school districts. Second, if locations of SBHCs are not randomly selected, schools that are untreated in the same district and gradespan may have different school-specific trends in unobservable variables. Appendix D shows that using a control group selected through propensity score matching within

[17] Appendix Table A.4 compares average suspension and dropout rates measured in the years *before an SBHC opens* between schools that eventually open an SBHC and schools that never open an SBHC. Notably, pre-opening suspension rates are significantly higher in schools that open an SBHC, suggesting that these schools may be meaningfully different than the average California school without an SBHC.

[18] Smith and Todd (2005) notes that while geographic restrictions are important in matching, difference-in-differences models allow for fixed differences in outcomes between treated units and matched control units, as these would be netted out. This suggests that the combination of propensity score matching with a difference-in-difference model weakens the need for prefect aligned "local labor markets" across the treated and control groups.

district does not satisfy the parallel pre-trends test. Using districts with similar "openness" to the SBHC-model relaxes geographic constraints that may be impractical in this setting, while maintaining the benefits of matching within a group of schools that may be faced with more similar conditions or share similar attitudes toward school-based health as treated schools. Appendix C discusses the process of selecting the propensity score predictors, the functional form of the predictive model, and the technical implementation of propensity-score matching. Robustness checks in Appendix B.3 show that key results hold when the matching pool is expanded to all untreated schools in any district.

The final matching procedure generates a sample of 47 SBHCs with 138 matched control schools.[19] Of the matched SBHCs, 40 report offering mental health services. To help isolate the mental health impacts of these centers, the final analytic sample limits to the 40 SBHCs that report offering mental health and their 117 matched control schools.[20] Table 1 compares average values of three baseline characteristics that are most likely to be related to SBHC openings, measured one period before an SBHC opening for the final matched sample. The table shows the differences between sample means (control mean minus treated mean) with p-values from a t-test checking that the difference is not equal to zero in parentheses. The comparison group mean is noted in brackets next to each difference. The first column shows these differences for all matched schools, while the next three columns restrict to sub-samples of elementary, middle, and high schools respectively. Across all columns there are no statistically significant differences in the fraction of FRPM students and total enrollment (the two variables used to construct propensity scores). The third row shows no statistically significant imbalance in the fraction of underrepresented minority (URM) students in the full matched sample; however, there is a small but non-negligible imbalance within the matched high school sample, with a magnitude of around 16% of the control mean. To account for residual imbalances, my preferred specification controls for all three school-level characteristics. Notably, the addition of controls does not meaningfully change the magnitude of the main results.

Table 2 summarizes the suspension and dropout outcomes used in this paper in the year prior to the opening of the relevant SBHC. There are no statistically significant differences in suspension rates at the 5% level in the period before an SBHC opens. Although means for repeat suspension rates are statistically different in the pre-period, the magnitude of the differences is relatively small (between 13-18% of the control mean). This suggests that, at least for the primary outcomes in this paper, the propensity score matching process has yielded comparable treated and control groups.

[19] Of the matched SBHC, 45 matched with three unique comparison schools each, one matched with two unique schools, and one matched with a single unique comparison school. Since the matching process used by the *psmatch2* package in Stata selects control schools such that the weighted combination of controls will achieve balance with the treated school, the small subset of treated schools with fewer than three controls may have been matched with the same controls as other treated schools in the sample.

[20] Appendix Table A.12 shows results for the main specifications run on the small subset of treated schools that *do not* report offering mental health services. Notably, there is no overall impact on suspensions, and the direction of impacts on defiance suspensions, non-defiance suspensions and repeat suspensions is the opposite from schools with mental health services. This suggests that SBHCs at these schools may be meaningfully different from SBHCs at schools that report offering mental health services.

Table 1: Difference in Sample Means ($Control - Treatment$)

	All Matched	Elementary	Middle	High
	Diff [Contr μ]	Diff [Contr μ]	Diff [Contr μ]	Diff [Contr μ]
Fraction FRPM	0.041 [0.722]	-0.011 [0.802]	-0.024 [0.659]	0.102 [0.684]
	(0.384)	(0.896)	(0.805)	(0.135)
Total Enrollment	-71.184 [981.991]	64.857 [647.857]	-68.952 [843.048]	-155.724 [1295.907]
	(0.607)	(0.272)	(0.529)	(0.553)
Fraction Minority	0.03 [0.78]	-0.02 [0.82]	-0.12 [0.69]	0.13* [0.78]
	(0.414)	(0.744)	(0.207)	(0.025)
N	157	56	28	73

Coefficients represent the difference between the control schools' sample mean and the treatment schools' sample mean.
Control group mean is printed in brackets.
p-values in parentheses

Table 2: Treatment versus Control Sample Means of Outcome Variables One Period Before an SBHC Opening

	Control	Treated	p-value
Suspension Rate	0.05	0.06	0.091
	[117]	[40]	
Defiance Suspension Rate	0.01	0.02	0.099
	[117]	[40]	
Non-Defiance Suspension Rate	0.03	0.04	0.258
	[117]	[40]	
Repeat Suspension Rate	1.44	1.61	0.014
	[107]	[36]	
Repeat Defiance Suspension Rate	1.27	1.51	0.002
	[83]	[30]	
Repeat Non-Defiance Suspension Rate	1.53	1.70	0.104
	[106]	[36]	
Dropout Rate	0.01	0.01	0.290
	[56]	[20]	

p-values are from a t-test that the treated and un-treated school means are equal
Number of observations in brackets under means

For all analyses, I estimate both a two-period difference-in-differences and an event-study using the Callaway and Sant'Anna (2021) adjustment for staggered-event settings. Equation 1 presents the underlying two-way fixed effects event study specification.

$$Y_{st} = \alpha + \gamma_s + \delta_0 Treated_s + \sum_{\substack{\tau=-3 \\ \tau \neq -1}}^{3} D_t^{\tau} + \omega_{\tau} \sum_{\substack{\tau=-3 \\ \tau \neq -1}}^{3} \delta_{\tau}(Treated_s \times D_t^{\tau}) + \mu \mathbb{X}_{st} + \varepsilon_{st} \tag{1}$$

Y_{st} is the dependent variable of interest, γ_s is a set of school fixed effects that control for fixed differences between schools, $Treated_s$ is a dummy equal to 1 if school s is a treated school, and D_t^{τ} is a dummy equal to 1 if the observation is τ years after (or before if τ is negative) the opening year for its matched pair. ω_{τ} is equal to the fraction of the "opening year" for which each SBHC is open if $\tau = 0$, equal to 0 for $\tau < 0$ and equal to 1 for $\tau > 0$. The purpose of ω_{τ} is to prevent misestimation of the coefficient for event year $\tau = 0$ that may occur due to differences in SBHC opening timings within an "academic year".[21] γ_s is a set of fixed-effects for each school. $\mathbb{X}_{st}$ is a vector of school-level characteristics that includes the fraction of FRPM students, the total enrollment, and the fraction of URM students for school s in year t.[22] Finally, standard errors are clustered at the school level and observations for comparison schools are assigned propensity score matching weights to account for the presence of multiple comparison schools for each treated school.[23] Equation 2 presents the two-period difference-in-differences version of this specification.

$$Y_{st} = \alpha + \gamma_s + \nu_t + \beta\,(Treated_s \times Post_t) + \mu \mathbb{X}_{st} + \varepsilon_{st} \tag{2}$$

where Y_{st}, $Treated_s$, γ_s, and $\mathbb{X}_{st}$ are as defined as above. ν_t is a set of year fixed effects, controlling for differences between time periods respectively. $Post_t$ is a dummy equal to 1 if year t is after the opening year of the SBHC (including the opening year itself). Weights and standard errors are the same as in the event study.

With a propensity-score matched control group, estimate validity relies on the assumption that conditional on having similar predicted likelihoods of opening an SBHC, outcomes for treated schools and matched control schools would evolve similarly in the absence of treatment. This assumption will be met if, given two schools with similar predicted likelihoods of opening an SBHC, the actual opening of an SBHC in school s in year y is plausibly random. There are two contextual reasons why we may expect conditional parallel trends to be satisfied. First, the standard timeline for constructing an SBHC can take around 2-3 years and may vary by district, leading to randomness in the length of time between a school's decision to open an SBHC and the actual opening. Second,

[21] In practice, this is implemented by weighting all observations in the year of an SBHC opening by the $(Opening\ Month)/12$. All other observations are assigned a weight of 1.

[22] The Callaway and Sant'Anna estimator is implemented using the *csdid* package in Stata, which uses baseline values of these covariates to effectively "match" and weight observations appropriately. This avoids concerns about using time-varying characteristics as controls.

[23] The use of school-level clustering follows standard difference-in-differences guidance to cluster standard errors at the level at which the policy is implemented. The propensity score matching weights are equivalent to $1/K$ for each comparison school where K is the total number of comparison schools that are matched to the corresponding treated school. For example, for a treated school that is matched to three comparison schools, each of those matched schools will have a weight of 0.333.

anecdotal evidence from school-based health administrators suggests that the most common reason for opening these centers relates to physical health concerns rather than mental health concerns; therefore, even if the timing of an SBHC opening is not perfectly exogenous to *all* student-level outcomes, there is reason to believe it may be exogenous with relation to the mental health and behavioral outcomes examined in this paper. The event study results in Section 5 suggest that the test for parallel trends is satisfied in the pre-event period.

5 Results

5.1 Mental Health Correlations

The connection between mental health issues and suspension-inducing behavior or dropout decisions has been understudied in existing literature. To motivate the use of suspension rates and dropout rates as outcomes, I use data from the California Healthy Kids Survey to document their relationship with student-reported mental health and student perceptions of school climate. I use a set of four psychometrically-validated indices measuring school climate (defined in Mahecha and Hanson (2020)) and two questions measuring student-reported mental health.[24] The four school-climate indices measure: delinquency; substance use; caring staff-student relationships; and school connectedness. The indices for delinquency and substance abuse capture student behaviors that we might expect to be directly correlated with suspensions and dropouts. The remaining two indices aim to capture positive sentiments of student-belonging, which we would expect to be correlated with better mental health based on existing research (Arslan, 2021).

Table 3 presents regressions where the dependent variable is the z-scored average school-level suspension rate, normalized within the relevant analysis sample. The independent variable is the school-level average value of a given CHKS index. For each CHKS measure, the second column shows the preferred specification that includes year and school fixed-effects. All estimates presented from these regressions should be interpreted as correlational relationships that control for fixed cross-year and cross-school differences. Examining student-reported behaviors first, Columns 1-4 of Table A.2 show that suspension rates are positively correlated with higher levels of delinquency and substance use. Columns 5-8 reveal that suspension rates are negatively correlated with higher levels of caring staff-student relationships and school connectedness, both of which are viewed as indicators of positive school climate. This suggests that schools with higher suspension rates may have worse perceived school climate or a larger fraction of students who do not feel that they are supported in their school.

Table 4 shows the same regressions, where the dependent variables are the fraction of students who report that they have considered attempting suicide in the past 12 months (Columns 1 and 2) and the fraction of students who report that they have experienced depression in the past 12 months

[24] Appendix E.2 lists the questions included in each index.

Table 3: Correlations Between School Climate and Z-Scored Suspension Rates

	Z-Scored Suspension Rates							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Delinquency	1.682***	0.378***						
	(0.127)	(0.074)						
Substance Use			0.667***	0.127***				
			(0.091)	(0.047)				
Worse Caring Staff-Student Relationships					1.112***	0.161***		
					(0.047)	(0.051)		
Worse School Connectedness							1.100***	0.259***
							(0.037)	(0.040)
Constant	-0.810***	0.146**	-0.366***	0.272***	-3.011***	-0.075	-2.269***	-0.215**
	(0.099)	(0.067)	(0.114)	(0.069)	(0.148)	(0.160)	(0.093)	(0.104)
Observations	10878	10436	10881	10438	10882	10441	10882	10441
Sample Mean	0.0	0.012	-0.0	0.012	-0.0	0.012	0.0	0.012
Fixed Effects	Year	Year/School	Year	Year/School	Year	Year/School	Year	Year/School

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

(columns 3 and 4). These results indicate that higher suspension rates are positively correlated with both measures of poor mental health. The magnitude of the correlations is lower in part due to a smaller sample of years for which these metrics are available, and response sparseness in years for which data is available.

Table 4: Correlation Between Mental Health and Z-Scored Suspension Rates

	Z-Scored Suspension Rates			
	(1)	(2)	(3)	(4)
Fraction of Students Considered Suicide	0.590**	0.208		
	(0.237)	(0.363)		
Fraction of Students Experienced Depression			0.872***	0.287**
			(0.128)	(0.142)
Constant	0.363***	0.417***	-0.049	0.097**
	(0.080)	(0.093)	(0.047)	(0.043)
Observations	4809	3759	8819	8348
Sample Mean	-0.0	0.021	0.0	0.010
Fixed Effects	Year	Year/School	Year	Year/School

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

While these estimates are only suggestive of potential relationships, they offer valuable insight into channels through which SBHCs might impact suspension rates. First, more common reports of feeling less comfortable and less interconnected in school are positively correlated with higher school-level suspension rates. The SBHC model's emphasis on broad provision of services, integration with a school, and confidentiality has the potential to generate a "safe space" that could improve students' feelings of belonging and comfort. Second, higher rates of reported depression and suicidal thoughts are positively correlated with higher suspension rates. This suggests that there may be a link between the behaviors that lead to suspensions and feelings commonly associated with anxiety or depression. If school-based health centers are able to increase the share of students with mental health issues who receive treatment, this may lead to a decrease in suspensions.

Tables 5 and 6 show regressions of dropout rates on the same set of school climate measures and mental health measures respectively. A first observation from Table 5 is that the magnitudes of the correlations between school climate and dropout rates are meaningfully smaller than those between school climate and suspension rates. A second observation is that the magnitude and direction of estimates are volatile across specifications. In all specifications with only year fixed-effects, relationships are statistically significant and positive; however, the addition of school fixed effects meaningfully attenuates these coefficients toward zero (and in the case of caring staff-student relationship and school connectedness, even changes the sign of the relationship). Table 6 reveals a similar pattern for feelings of depression. The inconsistency of signs across specifications makes these relationships difficult to interpret with confidence. While this may be due to the small sample size of the dropout rates data, it might also suggest that any impact that SBHCs have on adolescent mental health may not cleanly translate to changes in dropout rates.

Table 5: Correlations Between School Climate and Z-Scored Dropout Rates

	Z-Scored Dropout Rates							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Delinquency	0.870***	0.151**						
	(0.097)	(0.074)						
Substance Use			0.354***	0.010				
			(0.064)	(0.032)				
Worse Caring Staff-Student Relationships					0.335***	-0.157***		
					(0.072)	(0.053)		
Worse School Connectedness							0.451***	-0.127**
							(0.039)	(0.059)
Constant	-0.465***	0.106	-0.213**	0.213***	-0.831***	0.713***	-0.928***	0.543***
	(0.102)	(0.102)	(0.101)	(0.080)	(0.226)	(0.188)	(0.111)	(0.188)
Observations	8169	7761	8172	7765	8173	7767	8175	7769
Sample Mean	-0.0	-0.009	0.0	-0.009	0.0	-0.008	-0.0	-0.008
Fixed Effects	Year	Year/School	Year	Year/School	Year	Year/School	Year	Year/School

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

5.2 Suspension Rates

Table 7 estimates the average effect of an SBHC over a period of 4 post-opening years, including the year of the opening.[25] Columns 1 and 2 show results for the full sample, while columns 3-8 show results for separate subsamples of elementary, middle, and high schools. Within each estimation sample, the first column shows the baseline specification, while the second column shows

[25] The average effect is constructed by the *csdid* package in Stata, which calculates the average effect of treatment on the treated (ATT) for each "cohort" of treated schools and averages ATTs across all cohorts. This estimate should be interpreted as the average treatment effect across all schools that opened an SBHC in any year between 2012-2019.

Table 6: Correlations Between Mental Health and Z-Scored Dropout Rates

	Z-Scored Dropout Rates			
	(1)	(2)	(3)	(4)
	Year FE	Year/School FE	Year FE	Year/School FE
Fraction of Students Considered Suicide	-0.149	0.157		
	(0.284)	(0.182)		
Fraction of Students Experienced Depression			1.311***	-0.463
			(0.289)	(0.428)
Constant	0.247***	0.168***	-0.347***	0.193
	(0.079)	(0.050)	(0.078)	(0.123)
Observations	2865	2642	5271	4752
Sample Mean	0.0	-0.019	0.0	-0.015

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

the preferred specification which adds school-level controls for the fraction of students qualifying for free-and-reduced-price meals, fraction of minority students, and the total enrollment. Column 2 suggests that relative to control schools, overall suspension rates decrease by around 1.1 percentage points in schools with SBHCs in the years following the opening. Although this result is only marginally significant (at the 10% level) it is a notably large magnitude decrease that amounts to an 18% change from the comparison school pre-event mean. Columns 4, 6, and 8 suggest that the decrease in suspension rates is largest for middle schools, which see a 4.8 percentage point (or 66%) decrease. There is a smaller, marginally significant decrease for high schools of 0.9 percentage points and no identifiable effect for elementary schools.

Table 7: Suspension Rates: Difference-in-Differences Estimates by Sample

	All Grades		Elementary		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Demographics	Baseline	Demographics	Baseline	Demographics	Baseline	Demographics
Treated X Post	-0.0109**	-0.0106*	-0.0035	-0.0006	-0.0468***	-0.0476***	-0.0067	-0.0089*
	(0.0053)	(0.0054)	(0.0044)	(0.0051)	(0.0135)	(0.0162)	(0.0051)	(0.0049)
Pre-Period Control Mean	0.0541	0.0541	0.0253	0.0253	0.0720	0.0720	0.0682	0.0682
Observations	938	938	354	353	168	168	416	416

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

To examine change over time in the treatment effect, Figure 2 plots the coefficients from event-study versions of the specifications in columns 2, 4, 6, and 8 of table 7. Table 8 shows the corresponding event-study estimates. As expected based on the difference-in-differences results, the full sample and high school samples show no treatment effect in the pre-period and a slight declining trend in the years following the opening. While the individual event study estimates are underpowered for statistical significance, the trend suggests that the treatment effect increases in magnitude over time for schools opening an SBHC. It is worth noting that for the middle-school subsample, there is an increase in suspension rates in the year before the SBHC opens. This suggests that the true "baseline" level of suspension rates for these schools may be lower than the baseline used to calculate the difference-in-differences estimates. A simple back of the envelope calculation indicates that a reasonable lower bound for the treatment effect given the pre-period levels is 1 percentage point, which is the same magnitude as the treatment effect identified for the full sample.

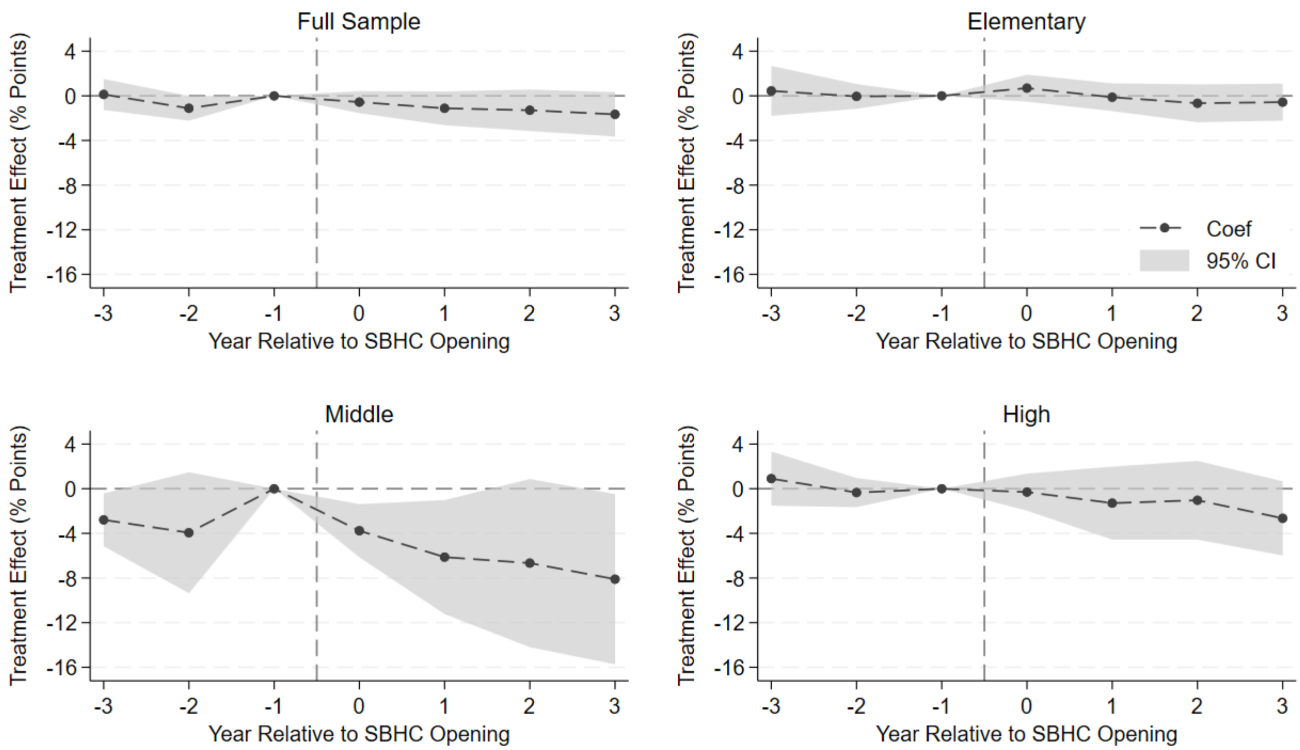

Figure 2: This figure plots the event time coefficients from event studies where the outcome is the school-level suspension rate. Each panel shows estimates for a different subsample of schools. All event studies control for school fixed effects and a vector of school characteristics that includes fraction of Free and Reduced Price Meal (FRPM) students, fraction of minority students, and total school enrollment. All lags prior to event time -3 and all leads after event time 3 are dropped from the estimation sample. Standard errors are clustered at the school level.

To begin to understand *why* suspension rates decrease in response to the opening of a new SBHC, table 9 shows difference-in-differences estimates for two sub-categories of suspensions: defiance-only and non-defiance. Here, I focus on the full sample with splits for middle school and high school only, acknowledging that suspension rates in elementary schools seem to be unaffected by the presence of an SBHC. The California Department of Education defines a defiance-only suspension as "any suspension associated with a student in which the only offense committed by a student is disruption." Non-defiance suspensions are a constructed category that includes all other suspendable offenses that are *not* defined as defiance-only.[26] This category includes violent incidents with an injury, violent incidents without an injury, weapons possession, illicit drug possession or sale, and "other

[26] The outcome for non-defiance suspensions is calculated by subtracting the CDE-reported count of "defiance-only" suspensions from their reported count of total school-level suspensions

Table 8: Suspension Rates: Event Study Estimates by Sample

	All Grades		Elementary		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Demographics	Baseline	Demographics	Baseline	Demographics	Baseline	Demographics
Treated x (τ = -3)	-0.0001	0.0013	0.0005	0.0044	-0.0293	-0.0279**	0.0022	0.0090
	(0.0073)	(0.0071)	(0.0094)	(0.0114)	(0.0196)	(0.0121)	(0.0134)	(0.0124)
Treated x (τ = -2)	-0.0105**	-0.0112**	-0.0023	-0.0005	-0.0335*	-0.0394	-0.0072	-0.0035
	(0.0052)	(0.0056)	(0.0061)	(0.0057)	(0.0187)	(0.0276)	(0.0061)	(0.0067)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0058	-0.0056	0.0037	0.0069	-0.0343***	-0.0376***	-0.0026	-0.0030
	(0.0048)	(0.0050)	(0.0051)	(0.0062)	(0.0094)	(0.0121)	(0.0072)	(0.0085)
Treated x (τ = 1)	-0.011	-0.011	-0.003	-0.001	-0.059***	-0.061**	-0.006	-0.013
	(0.007)	(0.008)	(0.005)	(0.006)	(0.022)	(0.026)	(0.011)	(0.017)
Treated x (τ = 2)	-0.013	-0.013	-0.009	-0.007	-0.061**	-0.067*	-0.007	-0.010
	(0.009)	(0.009)	(0.007)	(0.009)	(0.027)	(0.038)	(0.016)	(0.018)
Treated x (τ = 3)	-0.016*	-0.017	-0.007	-0.006	-0.068**	-0.081**	-0.020	-0.027
	(0.009)	(0.010)	(0.007)	(0.008)	(0.027)	(0.039)	(0.018)	(0.017)
χ^2	3.233	3.163	8.941	9.722	19.874	37.999	2.196	3.406
p-value	0.520	0.531	0.063	0.045	0.001	0.000	0.700	0.492
Pre-Period Control Mean	0.054	0.054	0.025	0.025	0.072	0.072	0.068	0.068
Observations	938	938	354	353	168	168	416	416

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

offenses".[27] Evidence from psychology suggests that one common way for mental health issues to manifest in adolescents is through "disruptive behavior" (Garland et al., 2010); therefore, the fraction of suspensions due to disruptive behavior may be a proxy for the share of suspensions that are caused by mental health issues.[28]

Table 9: Suspension Rates: Difference-in-Differences Heterogeneity by Suspension Category

	All Grades		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)
	Non-Defiance	Defiance	Non-Defiance	Defiance	Non-Defiance	Defiance
Treated X Post	0.0002	-0.0108*	0.0015	-0.0493***	0.0049	-0.0137***
	(0.0030)	(0.0060)	(0.0079)	(0.0179)	(0.0036)	(0.0031)
Pre-Period Control Mean	0.0344	0.0198	0.0432	0.0289	0.0427	0.0255
Observations	938	938	168	168	416	416

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

Table 9 shows that across all three samples, the decrease in suspension rates is driven by a decrease in defiance suspensions. This is especially striking given that non-defiance suspensions

[27] The specific offenses included in each of these categories are outlined in Appendix Section E.1.

[28] While violence and drug-use are behaviors that have also been linked to mental health issues, these are more extreme behaviors that are likely to result from severe mental health issues affecting a smaller fraction of students. Disruptive behavior can be viewed as a proxy for mental health issues that may be less severe but still detrimental to students if left untreated. Anecdotal evidence also suggests that given the staffing limitations faced by many school-based health centers they may be better equipped to directly treat less-severe mental health issues.

occur nearly twice as often at baseline, as evidenced by the relative size of the pre-period suspension rates at untreated schools. In the full sample and the middle school subsample the magnitude of the decrease in defiance suspensions is similar to the overall decrease in suspension rates. In the high school sample, however, there is a 1.4 percentage point decrease in defiance suspensions that is precisely estimated at the 1% level. Relative to the baseline rate of defiance suspensions in high schools in this sample, this amounts to a 53% decrease. This indicates that the pooled effect of SBHC-access on all high school suspensions may be an *underestimate* of the true effect on the subcategory of suspensions that is most impacted. Figure 3 plots the corresponding event studies for the difference-in-differences specifications in table 9.[29] The trends are as we would expect based on the difference-in-differences estimates. It is worth observing that the baseline difference between treated and comparison middle schools is closer to zero for defiance suspensions than it was for all suspensions. The corresponding panel for non-defiance suspensions in middle schools suggests that the pre-period "increase" in suspensions visible for middle schools in figure 2 is likely driven by an increase in non-defiance suspensions. This increases confidence that the decrease in suspensions at middle schools may in fact be as large as 4 percentage points on average. Moreover the event study estimates for defiance suspensions suggest that the decrease may be as large as 8 percentage points two years following the opening and 10 percentage points three years following the opening.[30]

There are two potential channels through which the opening of an SBHC could lead to a decrease in defiance-only suspensions. The first is that the opening of an SBHC increases the share of students with behavioral issues who receive treatment by improving access to mental health services. As a result, the frequency of disruptive behaviors that would lead to defiance suspensions decreases. We can call this first scenario, the "mental health channel". The second potential channel is that the opening of an SBHC leads teachers and principals to refer students with behavioral issues to a mental health professional at the SBHC as an *alternate policy* to suspensions. We can refer to this as the "displacement channel" whereby SBHCs decrease suspensions by serving as an alternative response to disruptive behavior.

These mechanisms are not mutually exclusive and under certain conditions may achieve the same impact. Consider, for example, a student who has untreated ADHD that results in inattention and disruptive behaviors. Prior to the opening of an SBHC, after repeated incidents of acting out in class, a school may have no option but to suspend this student for defiance. Under the "mental health channel", after an SBHC opens this student seeks out psychological services and receives the relevant treatments (either in the form of medication or counseling) to improve their focus and prevent disruptive behaviors that would have led to suspension. Under the "displacement channel", the student's teacher chooses to refer them to see the SBHC's psychologist after multiple disruptive incidents, rather than immediately issuing a suspension. If the student is only referred to mental

[29]Event study estimates are shown in appendix table A.5.

[30]It is worth noting that the estimation sample size decreases for later years. Therefore estimates for two years and three years following opening may be more susceptible to spurious results than estimates in event-years zero and one.

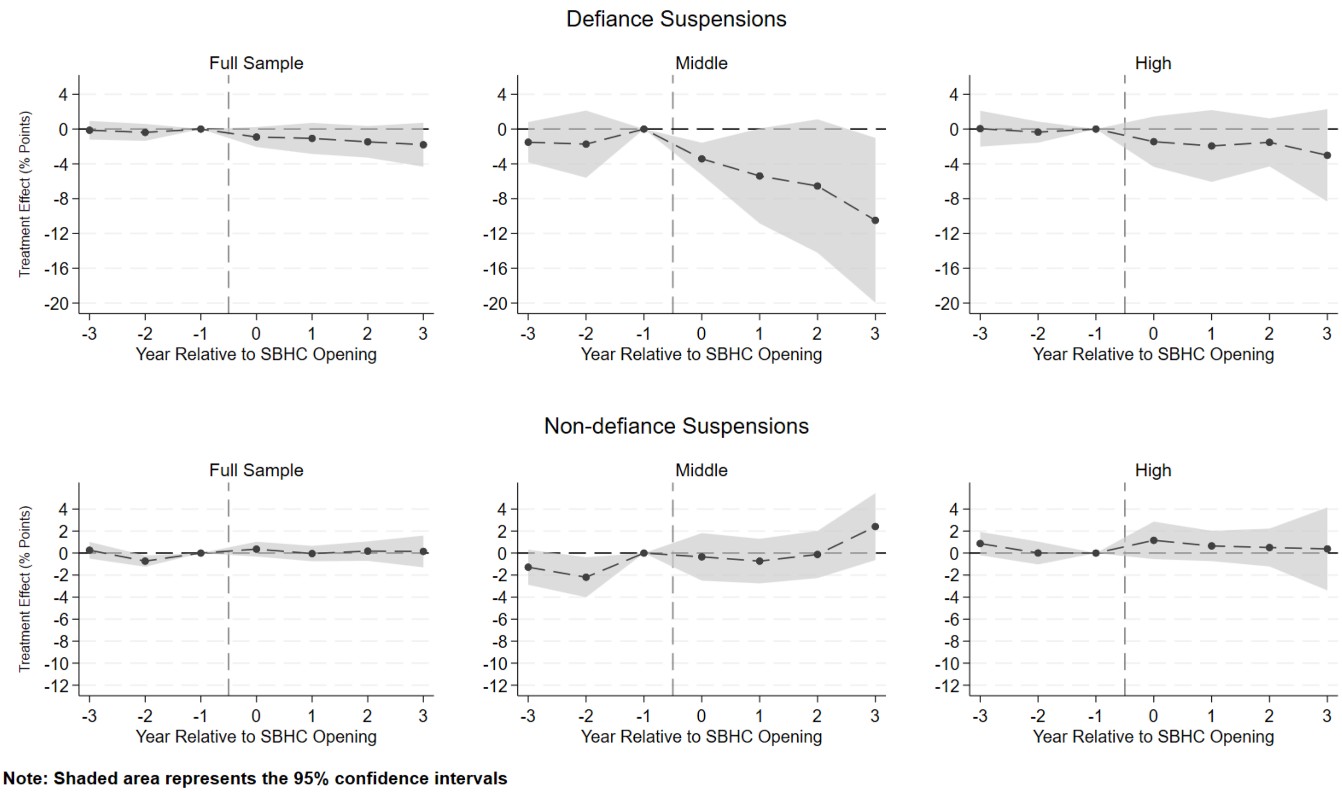

Figure 3: This figure plots the *Event Time* coefficients from separate event studies by sample where the outcomes are: the suspension rate for defiance suspensions **(top row)** and the suspension rate for non-defiance suspensions **(bottom row)**. All event studies control for school fixed effects and a vector of school characteristics that includes fraction of Free and Reduced Price Meal (FRPM) students, fraction of minority students, and total school enrollment. All lags prior to event time -3 and all leads after event time 3 are dropped from the estimation sample. Standard errors are clustered at the school level.

health services as an alternative to suspension but never actually receives treatment or changes their behavior, then their decreased likelihood of suspension is not caused by an improvement in mental health; however, if that initial referral ultimately leads the student to receive treatment for the issues that lead to their disruptive behavior, then the "displacement channel" has an equivalent outcome to the "mental health channel".

One approach to disentangling these two channels is to examine the effect of SBHCs on the rate of repeat suspensions, defined as the number of suspensions per suspended student. If SBHCs serve primarily as an alternate policy to suspending disruptive student and have no impact on treating students' mental health, we would expect disruptive behavior to recur frequently. Given that SBHCs have limited staff and resources, if a student's disruptive behavior never changes despite multiple visits to an SBHC, the school may eventually be forced to suspend the student; therefore, a student who was predisposed for repeat suspensions may see no impacts under the "displacement" channel. If, on the other hand, SBHCs are reducing improving mental health for certain students, this should both reduce the extensive margin probability of suspension for some subset of students, and decrease the likelihood of repeated suspensions for some subset of students who have been suspended at least once.

Table 10 shows difference-in-difference estimates where the outcome is the average number of suspensions per suspended student.[31] As in previous tables, columns 1 and 2 show the full sample of SBHCs, while columns 3-8 show subsample results for elementary, middle, and high schools. The even numbered columns show the preferred specification with demographic controls. Table 10 suggests a precisely estimated decrease in number of suspensions per suspended student for the full sample. This decrease of 0.14 suspensions represents a 10% decrease from the comparison school mean. Unsurprisingly given previous results, there is a notably larger decrease in repeat suspensions for middle schools of 0.42 suspensions (a 26% decrease). Interestingly, in column 2 there is also a statistically significant decrease of 0.28 suspensions per suspended student in elementary schools despite the lack of effect on the overall suspension rate. This may suggest that even if SBHC access does not impact the overall number of suspensions at elementary schools, it could have an impact on the intensive margin for students with the most persistent issues.

Table 10: Number of Repeat Suspensions: Difference-in-Differences by Sample

	All Grades		Elementary		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Demographics	Baseline	Demographics	Baseline	Demographics	Baseline	Demographics
Treated X Post	-0.1588***	-0.1425**	-0.2753***	-0.2876***	-0.3663***	-0.4217***	-0.0379	-0.0652
	(0.0556)	(0.0621)	(0.0930)	(0.0769)	(0.0900)	(0.0696)	(0.0661)	(0.0657)
Pre-Period Control Mean	1.4601	1.4601	1.4403	1.4403	1.6343	1.6343	1.3948	1.3948
Observations	852	852	303	301	166	166	383	382

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

Table 11 examines whether the decrease in repeat suspensions is driven by a specific type of

[31] Appendix table A.8 shows the corresponding event study estimates.

suspension.[32] For each subsample, the first column show the number of non-defiance suspensions per student while the second column shows the number of defiance suspensions per student. Unlike with the overall suspension rate, the decrease in repeat suspensions is not driven by a specific type of suspension. For the full sample, the decrease in repeat defiance suspensions is slightly larger than the decrease in repeat non-defiance suspensions; however for the middle school subsample, the decrease in repeat defiance suspensions is smaller than the decrease in repeat non-defiance suspensions. This result for middle schools does not necessarily contradict the results for the overall suspension rate. If the overall number of students who were suspended for "defiance-behaviors" decreased in schools with SBHCs, then the post-opening rate of defiance suspensions at these schools would be lower than the rate of non-defiance suspensions. This naturally leaves fewer students with any *defiance suspensions* and offers less room for change on the intensive margin.

Table 11: Number of Repeat Suspensions: Heterogeneity in Difference-in-Differences Estimates by Suspension Type

	All Grades		Elementary		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Defiance	Defiance	Non-Defiance	Defiance	Non-Defiance	Defiance	Non-Defiance	Defiance
Treated X Post	-0.1654*	-0.2255***	-0.2479*	-0.2435	-0.6971***	-0.3393***	-0.0662	-0.1001
	(0.0975)	(0.0798)	(0.1350)	(0.3291)	(0.1823)	(0.1091)	(0.0811)	(0.0716)
Pre-Period Control Mean	1.5779	1.3135	1.5208	1.3494	1.8422	1.3743	1.5003	1.2841
Observations	847	645	298	173	165	139	381	313

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

[32] Appendix table A.9 shows the corresponding event study estimates.

5.3 Dropout Rates

Table 12 shows the results from a two-period difference-in-differences specification where the dependent variable is the school-level dropout rate. Recall that the correlations presented in Section 5.1 suggest that the relationship between mental health and dropout rates is ambiguous; theoretically, however, severe mental health issues have the potential to lead to extreme actions such as dropping out of school. Since the data on dropouts primarily covers high schools and some combined middle-and-high schools, table 12 presents estimates for the full sample followed by splits for only the middle school and high school subsamples. For each sample shown in table 12, the first column shows the baseline specification while the second column shows the preferred specification with demographic controls. Across all columns with the preferred specification there is no statistically significant impact of the opening of an SBHC on dropout rates.[33] Appendix table A.10 shows no treatment effects in the years prior to an SBHC opening, suggesting that the parallel trends assumption is satisfied.[34]

Table 12: Dropout Rates: Difference-in-Differences by Sample

	All Grades		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Demographics	Baseline	Demographics	Baseline	Demographics
Treated X Post	-0.0017*	-0.0001	0.0009	0.0011	-0.0022	-0.0009
	(0.0009)	(0.0020)	(0.0015)	(0.0012)	(0.0014)	(0.0017)
Pre-Period Control Mean	0.0104	0.0104	0.0045	0.0045	0.0139	0.0139
Observations	348	348	110	110	205	200

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

An effect-size of zero is in line with the results from Lovenheim et al. (2016). Importantly, the 95% confidence intervals on the difference-in-differences estimates rule out increases or decreases in the dropout rate of greater than 0.5 of a percentage point. Given that the estimated decrease in suspension rates is around 1.1 percentage points, these bounds suggests that if there was a true *increase in dropout rates*, it would be unlikely to fully explain the decrease in suspension rates. This alleviates concerns that the decrease in suspension rates could be caused by students shifting from misbehaving to dropping out of school.

[33] Since the available data on dropout rates is more limited than the available data on suspension rates these analyses have smaller sample sizes, reducing estimation power.

[34] It is worth noting that for the full sample, appendix table A.10 shows a statistically significant decrease of around 1.1 percentage points in year 3 that is statistically significant at the 1% level. Compared to the control school baseline dropout rate of 1% this is a nearly 110% decrease; however, this coefficient is estimated on a relatively small sample of schools and should be interpreted with caution. This is further confirmed by the observation that for the middle school sub-sample, there are insufficient observations to estimate the coefficients for two and three years following the SBHC opening.

6 Robustness

6.1 Staggered Implementation Adjustments

The primary specifications in this paper use the difference-in-differences estimator proposed by Callaway and Sant'Anna (2021) to address concerns about bias from implicit negative weighting that may occur in a standard ordinary least squares (OLS) regression in the presence of staggered treatment years and heterogeneous treatment effects. Appendix B.1 shows that simple two-way fixed effects models yield comparable results to the main specifications shown in Section 5.

To further verify that negative weighting is unlikely to be a threat to my research design, I use a procedure proposed in de Chaisemartin and D'Haultfœuille (2020) to calculate the weights on each decomposed difference-in-difference estimate.[35] Following Chaisemartin and D'Haultfœuille, the standard deviation of the weights can be used to estimate bounds on the level of treatment effect heterogeneity that would be necessary for the true treatment effect to be zero or for all subgroup average treatment effects to have *different signs* from the estimated average treatment effect for the full sample. Applying this weight-calculation exercise to my preferred specification for the effect of SBHCs on suspension rates reveals that 100% of weights are positive. The calculated bounds suggest that treatment effect heterogeneity would need to be 0.02 percentage points for the true treatment effect to be 0. This level of treatment effect heterogeneity would be approximately equivalent to 36% of the baseline suspension rate in the study sample. Based on these bounds and the presence of no negative weights, it is reasonable to conclude that the true effect of SBHC openings on suspension rates is most likely a decrease.

6.2 Alternate Control Groups

The pool of control schools used in the main specifications of this paper was selected to ensure the highest level of comparability with schools that select into opening an SBHC. One restriction made in that matching process that may raise concerns about external validity is the selection of control schools from the pool of untreated schools in districts that have at least one open SBHC. Appendix Section B.3 shows that the results from Section 5 are robust in direction and significance to selecting control schools from the pool of all never-treated schools, although the magnitudes of the effect are slightly smaller with this alternate control group. Notably, this alternate control group is still matched on school-level characteristics. This suggests that the results in this paper may be local to schools that are similar at baseline to the types of schools that choose to open an SBHC. Appendix D explores alternate control groups that yield different results and expands on reasons these groups may not be contextually valid controls.

[35] In practice, this is implemented using the *twowayfeweights* package in Stata.

7 Heterogeneity Analysis

Given the large decrease in suspension rates identified in Section 5.2, it is valuable to examine whether this effect differs across student demographics. Table 13 shows difference-in-differences estimates for suspension rates across a set of population subsamples: male and female students (columns 1-2); six categories of race (columns 3-8); and background characteristics such as English Language Learners (ELL), disabled, and socioeconomically disadvantaged (columns 9-11).[36]

Looking first at gender, columns 1 and 2 show that overall, the decrease in suspension rates is nearly twice as large for male students than for female students. Middle schools show the same pattern, while in high schools, the change in suspension rates for female students is slightly higher than that of male students. It is important to note, however, that across all samples the baseline suspension rate for male students is nearly twice as large as that for female students. For example, the 1.3 percentage point decrease in suspension rates for male students amounts to an 18% decrease from the control group baseline. Comparatively, a 0.8 percentage point decrease for female students, while smaller in magnitude, is a 22% decrease from the baseline rate for female students.

Looking to effects by racial identity, columns 3-5 suggest that effects may be larger for Black and Hispanic students than for White students. In the middle school sample in particular, there is additional suggestive evidence of large decreases for Asian and Filipino students; however, it is worth noting that due to the smaller sample of treated middle schools, these heterogeneity results could be spurious correlations rather than true effects. Finally, looking at background characteristics, column 10 suggests that effects may be large for students with disabilities. In the middle school sample, there are additional large effects for English language learners and students defined by the CDE as "socioeconomically disadvantaged".

These results align with the theory of change behind the SBHC model. In particular, SBHCs aim to bridge gaps in access and affordability for disadvantaged students. Given historic correlations between racial minority status and lower socioeconomic status, it is unsurprising that the impacts of this intervention should be concentrated among non-White students. Moreover, if mental health services are the primary driver of the decrease in suspensions, it is reasonable that effects would be strong for students who are identified as "disabled", a category that includes behavioral issues such as ADHD. Finally, a larger magnitude decrease for male students than female students is in line with previous research on interventions that impact delinquency. Komisarow and Hemelt (2022), for example, finds that the effects of their telemedicine program on chronic absenteeism and delinquency are largest for male students. It is worth noting, however, that access to SBHCs seems to have a positive impact for both male and female students, suggesting that this unique intervention may be universally beneficial.

[36] Each of these subsamples is based on a category of disaggregation provided by the CDE in their publicly available suspension data. The outcome in each subsample should be interpreted as the fraction of students in that subsample who were suspended in a given year.

Table 13: Suspension Rate Heterogeneity

	Gender		Race							Background	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
	Male	Female	White	Black	Hispanic	Asian	Filipino	Multiple Races	ELL	Disability	SE Disadv.
All Grades	-0.0133*	-0.0077**	-0.0067	-0.0215**	-0.0110*	-0.0038	-0.0006	0.0049	-0.0118	-0.0180**	-0.0056
	(0.0080)	(0.0039)	(0.0049)	(0.0108)	(0.0064)	(0.0049)	(0.0067)	(0.0099)	(0.0072)	(0.0084)	(0.0050)
Pre-Period Control Mean	0.0732	0.0336	0.0497	0.1112	0.0543	0.0198	0.0213	0.0645	0.0580	0.0895	0.0594
Observations	938	938	726	640	938	566	389	447	929	931	938
Elementary	-0.0007	-0.0002	0.0030	-0.0080	0.0012	-0.0153**	0.0018	-0.0479***	0.0013	-0.0011	0.0000
	(0.0078)	(0.0043)	(0.0115)	(0.0101)	(0.0049)	(0.0073)	(0.0012)	(0.0183)	(0.0038)	(0.0117)	(0.0057)
Pre-Period Control Mean	0.0388	0.0115	0.0361	0.0810	0.0217	0.0123	0.0193	0.0313	0.0209	0.0434	0.0248
Observations	353	353	235	183	353	172	37	110	353	353	353
Middle	-0.0637***	-0.0321**	0.0114	-0.1133***	-0.0475**	-0.0655***	-0.0655***	0.0246	-0.0865***	-0.0458*	-0.0510***
	(0.0196)	(0.0132)	(0.0226)	(0.0268)	(0.0187)	(0.0141)	(0.0244)	(0.0236)	(0.0196)	(0.0262)	(0.0167)
Pre-Period Control Mean	0.0985	0.0447	0.0575	0.1646	0.0768	0.0220	0.0128	0.0592	0.0799	0.1296	0.0823
Observations	168	168	150	125	168	101	75	97	168	166	168
High	-0.0074	-0.0110**	-0.0086	-0.0126	-0.0090	0.0061	0.0059	-0.0103	0.0010	-0.0185	-0.0003
	(0.0071)	(0.0051)	(0.0057)	(0.0135)	(0.0080)	(0.0049)	(0.0111)	(0.0106)	(0.0101)	(0.0129)	(0.0064)
Pre-Period Control Mean	0.0883	0.0456	0.0542	0.1070	0.0694	0.0219	0.0270	0.0808	0.0773	0.1077	0.0757
Observations	416	416	332	323	416	280	214	206	407	411	416

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

8 Discussion

The analyses presented in this paper show that access to a school-based health center leads to a significant and large drop in suspension rates, especially for middle school students. Regardless of whether this decrease is directly caused by improvements in mental health, these results suggest that the provision of comprehensive in-school health services should be considered as a potential policy tool for decreasing students' exposure to discipline. The observation that the decrease in suspensions is driven by a drop in "defiance" suspensions suggest that SBHCs may be treating mental health issues that lead to disruptive behavior. It is important, however, to consider the alternative channels through which the opening of a school-based health center might lead to a drop in suspension rates.

One possible alternate channel for these impacts is that schools are using SBHCs as an alternative "policy" to disciplinary actions. For example, we might see a decrease in suspension rates after the opening of an SBHC if teachers are now advised by the school to send disruptive students to the SBHC *instead of* sending them to the principal's office to receive a suspension. The decrease in repeat suspensions provides some evidence that this channel is unlikely to explain the entire decrease in suspension. Specifically, the repeat suspension rate is calculated as the total number of suspensions divided by the number of students who were ever suspended; therefore, implicitly, it is a statistic for the subsample of students who have been suspended at least once. A decrease in this rate suggests that SBHC-access has impacts for students who are suspended at least once, which suggests that SBHCs are not entirely displacing suspension as a disciplinary tool. There are also contextual reasons that the "displacement channel" is unlikely to fully explain the decline in suspension rates. Discussions with SBHC administrators indicate that SBHCs are often under-staffed relative to student demand for mental health services. The existing strain on these services should disincentivize school principals from sending students to SBHCs unless the student is expected to actually receive and benefit from treatment. Moreover, for many students, an initial visit with a mental health professional, (regardless of what channels led to that visit), may help them identify their own mental health treatment needs or overcome stigmas they hold regarding therapy. Due to a relatively small sample of schools and data limitations, this paper is unable to clearly conclude that the decrease in suspensions is linked to improved mental health. Future work should focus on more clearly disentangling these potential mechanisms through conversations with SBHC leaders and school administrators and student-level data on SBHC utilization.

It is also important to note that the results presented in this paper are local to a set of schools that have historically selected into opening school-based health centers. This locality stems from the propensity-matching approach to selecting comparison schools. As a result, we should expect the estimated decrease in suspension rates to translate most directly to other schools that are similar to the type of school that has traditionally chosen to open an SBHC. Anecdotally, these may be schools that serve a larger fraction of low-income students or have a larger size. Policymakers and educators should consider school characteristics when assessing whether SBHCs would be an efficient intervention for addressing delinquency or mental health. This paper is unable to conclude that school-based health centers would have the same magnitude impact on schools that look meaningfully different from the treated and control samples. Exploring the impact of SBHCs on schools with different demographic profiles requires a larger sample of SBHCs than are available in this study, and is a valuable goal for future research.

Finally, this paper focuses on a set of SBHCs that explicitly report offering mental health services and select outcomes that are most likely to be impacted through mental health channels. However, policymakers should note that SBHCs typically offer a wide range of health services that may improve students' physical health and broader well-being. Any considerations of SBHCs as a tool for treating mental health should evaluate the full benefit of these centers as well as the full costs of offering such comprehensive healthcare. Moreover, the school-level data used in this paper is unable to assess any student-level spillovers or indirect benefits of SBHC-access that may come from changes in long-run achievement or labor market outcomes. Due to the limited availability of reliable mental health data this paper also cannot assess the direct impact of SBHCs on clinically-validated mental health outcomes.

While it is theoretically possible to calculate longer-run impacts from short-run effects, this calculation requires a clear and precisely estimated causal pathway from decreased suspension rates to long-run earnings and labor market participation. Unfortunately, existing research on these causal pathways is limited and difficult to translate to the current study setting due to contextual differences. The research does, however, suggest that higher suspension rates may have long-run impacts on student performance, dropout rates, and future incarceration. Descriptive work from Fabelo et al. (2011) shows that being suspended is positively correlated with repeating a grade and dropping out of school. Recent quasi-experimental work from Bacher-Hicks et al. (2019) finds that students who are randomly assigned to attend a middle school with a one standard deviation higher suspension rate than their previous school are 1.7 percentage points more likely to ever drop out of school and 2.5 percentage points more likely to ever be incarcerated. These results suggest that by decreasing suspension rates, school-based health centers may also have longer-run impacts on students' high school completion and employment. Finally, work from Carrell et al. (2018) shows that exposure to a disruptive peer in elementary school has long-run effects on reducing earnings by around 3%. If the decrease in suspension rates identified in this paper is indicative of a true decrease in disruptive behavior, there may be positive spillovers for students who are not directly using SBHC services from decreased classroom disruptions. Policy discussions around school-based health centers would benefit from student-level research on the impact of SBHCs on mental health and academic spillovers, as well as additional evidence on the causal impact of decreased discipline and improved mental health on academic and labor market outcomes. This evidence, in combination with the results from the current paper, will be instrumental in evaluating whether the benefits of SBHCs outweigh the costs.

9 Conclusion

Worsening trends in adolescent mental health have been a focal point of recent policy discussions and funding investments. This paper aims to contribute to that policy discussion by evaluating the impact of access to a school-based health center on suspension and dropout rates, two behavioral outcomes that are likely to be directly impacted by untreated mental health issues. In addition to being the first paper to examine the effect of school-based health center access on suspensions and the first to examine the impact on dropout rates using school-level data, this paper also provides novel evidence regarding the correlation between these outcomes and student-reported mental health and school climate. To address selection in the decision to open an SBHC, I use a propensity-score matching approach to identify a theoretically reasonable and empirically valid comparison group.

The opening of an on-site SBHC that offers mental health care leads to a significant decrease in suspension rates, that may be as large as 4.9 percentage points in middle schools and 1.4 percentage points in high schools. Even the lower bound estimate of 1.1 percentage points amounts to an 18% decrease from the baseline average for comparison schools. I provide various pieces of evidence that this decrease in suspension rates is driven by an improvement in mental health. First, decomposing suspension rates by the category of offense reveals large decreases in suspensions due to "disruptive behavior", but minimal effects for suspensions due to violence or drug use. Second, access to an SBHC leads to a decrease in repeat suspensions, which are more likely to be driven by "persistent" behavioral issues. This indicates that SBHCs may be treating the root cause of suspensions rather than simply serving as a temporary alternative policy. Finally, correlations from the California Healthy Kids Survey suggest that high suspension rates are negatively correlated with feelings of "belonging" and "staff support" and positively correlated with feelings of depression. This combination of results suggests that school-based health centers may shift behaviors that often result from untreated mental health issues.

While I find no overall effect of SBHC-access on dropout rates, tight confidence intervals on these estimates allow me to rule out increases and decreases in dropout rates of more than 0.5 of a percentage points. This helps rule out the possibility that SBHCs lead students who would otherwise have been suspended, to drop out of school instead. It is also helpful to note that a zero-effect for dropout rates is in line with the results of previous research on school-based health centers by Lovenheim et al. (2016). That SBHCs have no effect on dropout rates is not unreasonable given that the long-run repercussions to dropping out are larger than the repercussions of a single suspension. Moreover, while the decision to drop out could be linked to poor mental health, alternative factors such as family issues and academic performance may be stronger drivers. The survey correlations presented in this paper indicate a more ambiguous correlation between mental health and dropout rates that further supports this idea. Finally, it is worth noting that this study is unable to rule out longer-run effects on dropout rates that show up more than three years after the SBHC opening.

The results in this paper should be cautiously interpreted relative to the assumptions made in selecting the control group. Specifically, we can conclude that suspension rates decrease in the years following the opening of an SBHC for treated schools relative to untreated schools with similar demographic compositions. Appendix D shows that these results do not necessarily hold for all possible alternate control groups. While pre-trend tests and contextual evidence suggest that the control group used in this paper is the strongest comparison, it is impossible to entirely rule out the validity of alternate control groups. As such, the impact of SBHCs on suspension rates may not replicate in schools that are meaningfully different from the treated schools in this paper. By showing the potential variation in results across different comparison groups, this paper highlights a key challenge to causal identification using available data on SBHCs. Rigorous evaluation of these centers would benefit from more exogenous source of variation such as a policy that randomizes the timing at which new schools receive grants to fund new SBHC construction.

This paper contributes to a small but growing literature on the impacts of school-based health centers by providing novel evidence that these centers may have large impacts on student behavior. My results offer some natural paths for future research. First, developing a clear recommendation on the use of SBHCs to address adolescent mental health requires evidence on the direct mental-health impacts of these centers. Second, given the large overhead costs to opening and operating SBHCs, subsequent studies should evaluate whether less comprehensive alternatives to on-site SBHCs, such as in-school mental health professionals or mobile SBHC clinics, could have the same behavioral impacts. Finally, to better assess how much variation there is the impact of SBHCs, one research priority should be to collect and standardize data on individual SBHC operations, services, and utilization. This data would improve our understanding of the specific features of SBHCs that benefit students, and provide valuable insight into how the effects of these centers can be replicated across different settings.

References

Arslan, G. (2021). School belongingness, well-being, and mental health among adolescents: Exploring the role of loneliness. *Australian Journal of Psychology*, 73(1):70–80.

Bacher-Hicks, A., Billings, S. B., and Deming, D. J. (2019). The school to prison pipeline: Long-run impacts of school suspensions on adult crime. Technical report, National Bureau of Economic Research.

BHWF (2023). Primary care health professional shortage areas (hpsas).

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of econometrics*, 225(2):200–230.

CalSCHLS (2026). The surveys: California healthy kids survey (chks). https://calschls.org/about/the-surveys/. Retrieved March 19, 2026.

Carrell, S. E. and Carrell, S. A. (2006). Do lower student to counselor ratios reduce school disciplinary problems? *The BE Journal of Economic Analysis & Policy*, 5(1):0000101515153806451463.

Carrell, S. E. and Hoekstra, M. (2014). Are school counselors an effective education input? *Economics Letters*, 125(1):66–69.

Carrell, S. E., Hoekstra, M., and Kuka, E. (2018). The long-run effects of disruptive peers. *American Economic Review*, 108(11):3377–3415.

CDC (2022). New cdc data illuminate youth mental health threats during the covid-19 pandemic.

CDE (2023). Income eligibility scales for school year 2023–24.

Cohen, D. R., Lewis, C., Eddy, C. L., Henry, L., Hodgson, C., L. Huang, F., Reinke, W. M., and Herman, K. C. (2023). In-school and out-of-school suspension: Behavioral and psychological outcomes in a predominately black sample of middle school students. *School psychology review*, 52(1):1–14.

CSBHA (2022a). From vision to reality. Technical report, California School-Based Health Alliance, Oakland, CA.

CSBHA (2022b). Navigating the promise of sbhcs: A guide for health care leaders.

CSBHA (2023). Funding school-based health & wellness centers: California school-based health alliance.

CSBHA (2025). School-based health centers wellness centers in california: A growing trend. Technical report, California School-Based Health Alliance.

CSHCA (2010). California school nursing roles and responsibilities.

Currie, J. and Stabile, M. (2006). Child mental health and human capital accumulation: the case of adhd. *Journal of health economics*, 25(6):1094–1118.

de Chaisemartin, C. and D'Haultfœuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, 110(9):2964–96.

Deshpande, M. and Li, Y. (2019). Who is screened out? application costs and the targeting of disability programs. *American Economic Journal: Economic Policy*, 11(4):213–48.

Dupéré, V., Dion, E., Nault-Brière, F., Archambault, I., Leventhal, T., and Lesage, A. (2018). Revisiting the link between depression symptoms and high school dropout: Timing of exposure matters. *Journal of Adolescent Health*, 62(2):205–211.

Fabelo, T., Thompson, M. D., Plotkin, M., Carmichael, D., Marchbanks, M. P., and Booth, E. A. (2011). Breaking schools' rules: A statewide study of how school discipline relates to students' success and juvenile justice involvement. *New York: Council of State Governments Justice Center*.

Fadlon, I. and Nielsen, T. H. (2021). Family labor supply responses to severe health shocks: Evidence from danish administrative records. *American Economic Journal: Applied Economics*, 13(3):1–30.

Flaherty, L. T. and Osher, D. (2003). History of school-based mental health services in the united states. In *Handbook of school mental health advancing practice and research*, pages 11–22. Springer.

Gall, G., Pagano, M. E., Desmond, M. S., Perrin, J. M., and Murphy, J. M. (2000). Utility of psychosocial screening at a school-based health center. *Journal of School Health*, 70(7):292–298.

Garland, A. F., Brookman-Frazee, L., Hurlburt, M. S., Accurso, E. C., Zoffness, R. J., Haine-Schlagel, R., and Ganger, W. (2010). Mental health care for children with disruptive behavior problems: A view inside therapists' offices. *Psychiatric Services*, 61(8):788–795.

Geierstanger, S. P., Amaral, G., Mansour, M., and Walters, S. R. (2004). School-based health centers and academic performance: Research, challenges, and recommendations. *Journal of School Health*, 74(9):347–352.

Hanson, T. L. and Kim, J.-O. (2007). Measuring resilience and youth development: the psychometric properties of the healthy kids survey.

Heckman, J. J., Ichimura, H., and Todd, P. (1998). Matching as an econometric evaluation estimator. *The review of economic studies*, 65(2):261–294.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, 64(4):605–654.

Juszczak, L., Melinkovich, P., and Kaplan, D. (2003). Use of health and mental health services by adolescents across multiple delivery sites. *Journal of Adolescent Health*, 32(6):108–118.

Katz, E. (2020). Realizing the potential of school-based health centers.

Kerns, S. E., Pullmann, M. D., Walker, S. C., Lyon, A. R., Cosgrove, T., and Bruns, E. J. (2011). Adolescent use of school-based health centers and high school dropout. *Archives of Pediatrics & Adolescent Medicine*, 165(7):617–623.

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., and Walters, E. E. (2005). Lifetime prevalence and age-of-onset distributions of dsm-iv disorders in the national comorbidity survey replication. *Archives of general psychiatry*, 62(6):593–602.

Kisker, E. E. and Brown, R. S. (1996). Do school-based health centers improve adolescents' access to health care, health status, and risk-taking behavior? *Journal of Adolescent Health*, 18(5):335–343.

Komisarow, S. and Hemelt, S. W. (2022). School-based healthcare and absenteeism: Evidence from telemedicine. *Education Finance and Policy*, pages 1–50.

Love, H. E., Schlitt, J., Soleimanpour, S., Panchal, N., and Behr, C. (2019). Twenty years of school-based health care growth and expansion. *Health Affairs*, 38(5):755–764.

Lovenheim, M. F., Reback, R., and Wedenoja, L. (2016). How does access to health care affect teen fertility and high school dropout rates? evidence from school-based health centers. Technical report, National Bureau of Economic Research.

Mahecha, J. and Hanson, T. (2020). Measurement structure of the california school climate, health, and learning surveys.

McCord, M. T., Klein, J. D., Foy, J. M., and Fothergill, K. (1993). School-based clinic use and school performance. *Journal of Adolescent Health*, 14(2):91–98.

McLeod, J. D., Uemura, R., and Rohrman, S. (2012). Adolescent mental health, behavior problems, and academic achievement. *Journal of health and social behavior*, 53(4):482–497.

NCBH (2019). America's mental health.(2018).

Paschall, M. J. and Bersamin, M. (2018). School-based health centers, depression, and suicide risk among adolescents. *American journal of preventive medicine*, 54(1):44–50.

Reback, R. (2010a). Noninstructional spending improves noncognitive outcomes: Discontinuity evidence from a unique elementary school counselor financing system. *Education Finance and Policy*, 5(2):105–137.

Reback, R. (2010b). Schools' mental health services and young children's emotions, behavior, and learning. *Journal of Policy Analysis and Management*, 29(4):698–725.

Rott, N. (2013). La schools throw out suspensions for 'willful defiance'.

SAMHSA (2023). Highlights for the 2021 national survey on drug use and health.

Santelli, J., Kouzis, A., and Newcomer, S. (1996). Student attitudes toward school-based health centers. *Journal of Adolescent Health*, 18(5):349–356.

Smith, J. A. and Todd, P. E. (2005). Does matching overcome lalonde's critique of nonexperimental estimators? *Journal of econometrics*, 125(1-2):305–353.

So, M., Freese, R. L., and Barnes, A. J. (2024). Pushed out and drawn in: Exclusionary discipline, mental health, and protective factors among youth in public schools. *Journal of School Health*, 94(2):128–137.

Walker, S. C., Kerns, S. E., Lyon, A. R., Bruns, E. J., and Cosgrove, T. (2010). Impact of school-based health center use on academic outcomes. *Journal of Adolescent Health*, 46(3):251–257.

Warren, C. and Fancsali, C. (2000). Lessons from the evaluation of new jersey's school-based youth services program. *Improving Results for Children and Families: Linking Collaborative Services with School Reform Efforts*, pages 91–120.

Wright, B., Padilla-Frausto, D. I., Tse, H. W., Crawford-Roberts, A., Kabir, F., and Salem, S. (2021). Nearly 1 in 3 adolescents in california reports serious psychological distress. Technical report, Los Angeles, Calif.: UCLA Center for Health Policy Research.

A Additional Tables & Figures

Table A.1: Summary Statistics: All SBHCs, On-Site, and Off-Site

	All SBHCs	On-Site	Off-Site/Mobile
Opened Between 2011-2019	0.38 [0.49]	0.37 [0.48]	0.41 [0.49]
High School	0.52 [0.50]	0.57 [0.50]	0.37 [0.49]
Middle School	0.13 [0.34]	0.14 [0.35]	0.10 [0.30]
Elementary School	0.33 [0.47]	0.27 [0.44]	0.48 [0.50]
Other/Unidentified	0.02 [0.15]	0.01 [0.12]	0.05 [0.22]
Mental Health	0.73 [0.44]	0.83 [0.38]	0.48 [0.50]
Medical	0.82 [0.39]	0.83 [0.38]	0.80 [0.40]
Reproductive Health	0.62 [0.49]	0.63 [0.48]	0.61 [0.49]
Dental or Vision	0.62 [0.49]	0.60 [0.49]	0.66 [0.48]
Serves Other Students	0.42 [0.49]	0.38 [0.49]	0.53 [0.50]
Serves Other Youth	0.56 [0.50]	0.51 [0.50]	0.68 [0.47]
Serves Community	0.44 [0.50]	0.38 [0.49]	0.62 [0.49]
Serves Families	0.59 [0.49]	0.57 [0.50]	0.65 [0.48]
Serves Staff	0.26 [0.44]	0.29 [0.45]	0.18 [0.38]
CHC Sponsored	0.52 [0.50]	0.52 [0.50]	0.52 [0.50]
Hospital Sponsored	0.04 [0.20]	0.02 [0.14]	0.10 [0.30]
Health Department Sponsored	0.08 [0.27]	0.05 [0.22]	0.14 [0.35]
School System Sponsored	0.26 [0.44]	0.31 [0.47]	0.13 [0.33]
Private Nonprofit Sponsored	0.08 [0.27]	0.07 [0.26]	0.10 [0.30]
Other Sponsored	0.02 [0.14]	0.02 [0.15]	0.01 [0.11]
Observations	286	207	79

Standard deviations in brackets

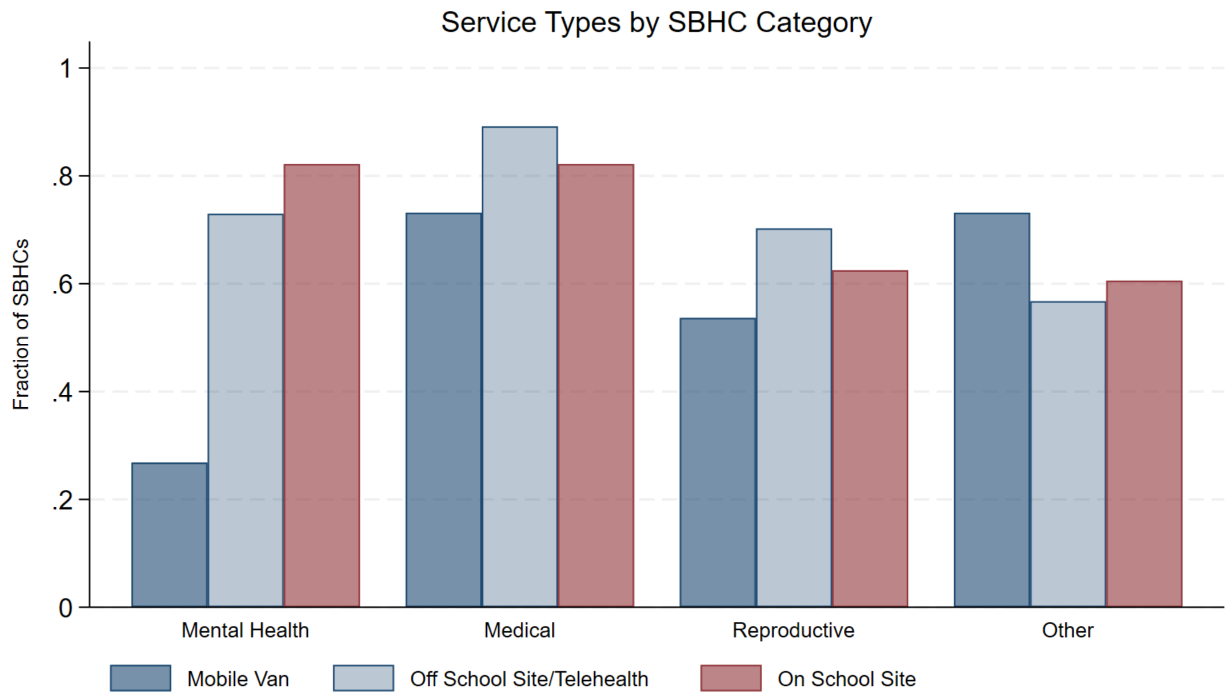

Figure A.1: The figure graphs the fraction of school-based health centers within each center "type", that report serving each of the following areas: the principal school (i.e. the school to which the SBHC is attached), multiple schools, the entire district, or multiple districts. Data on the schools served comes from a self-reported text field where the school-based health center provides a list of all schools it serves. I determine the four categories as follows: SBHCs that only list one school, where the school matches the principal school are classified as serving the "principal school"; SBHCs that list the name of more than one school are classified as serving "multiple schools"; SBHCs that list their principle district are classified as serving the "principal district"; and SBHCs that list multiple districts are classified as serving "multiple districts".

Table A.2: Correlations Between School Climate and Average Suspension Rates

	Mean Suspension Rates							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
More Delinquency	0.094***	0.021***						
	(0.007)	(0.004)						
More Substance Use			0.037***	0.007***				
			(0.005)	(0.003)				
Worse Caring Staff-Student Relationships					0.062***	0.009***		
					(0.003)	(0.003)		
Worse School Connectedness							0.062***	0.015***
							(0.002)	(0.002)
Constant	0.017***	0.070***	0.042***	0.078***	-0.107***	0.058***	-0.065***	0.050***
	(0.006)	(0.004)	(0.006)	(0.004)	(0.008)	(0.009)	(0.005)	(0.006)
Observations	10878	10436	10881	10438	10882	10441	10882	10441
Sample Mean	0.062	0.063	0.062	0.063	0.062	0.063	0.062	
Fixed Effects	Year	Year/School	Year	Year/School	Year	Year/School	Year	Year/School

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

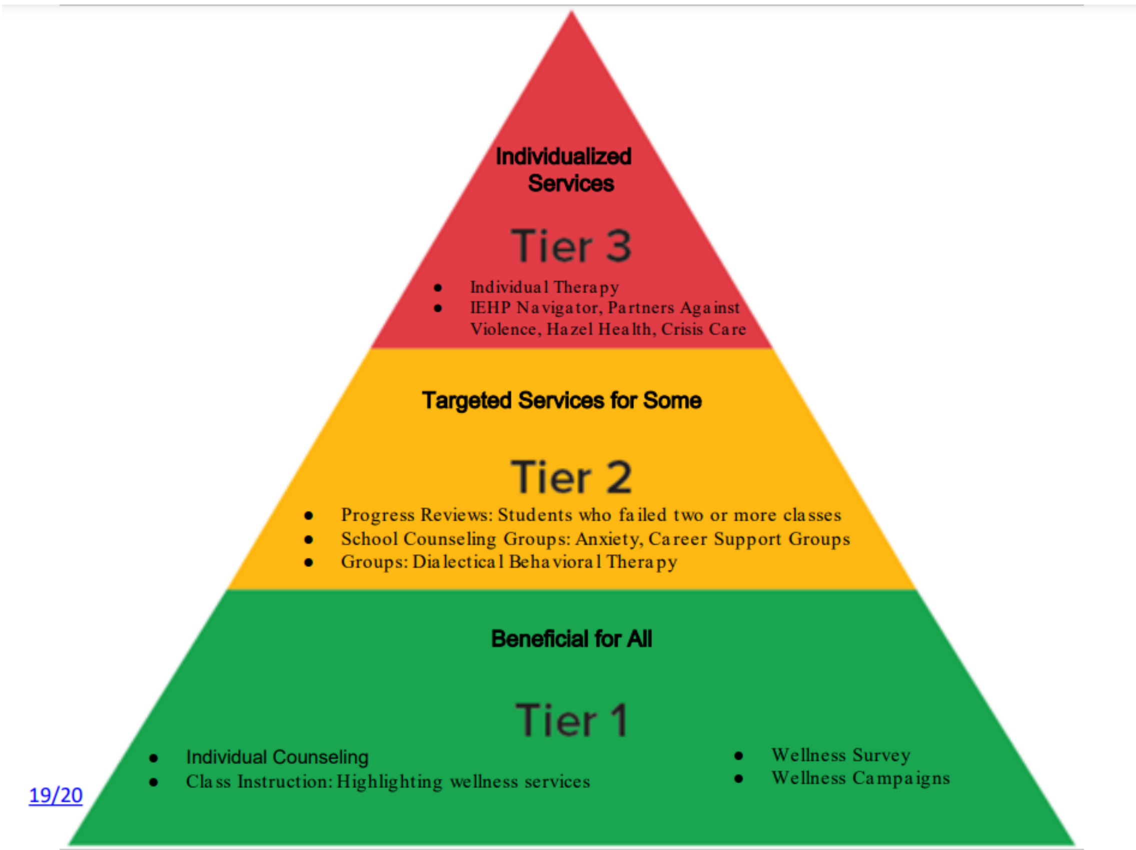

Figure A.2: This figure comes from a presentation delivered by the California School-Based Health Alliance (CBSHA) at their annual conference in 2023. It outlines the tiers of mental-health service provision recommended for new SBHCs by the CBSHA. Tier 1 services are services that all SBHCs in California that report offering "mental health" services will provide. The ability to offer Tier 2 and Tier 3 services will vary from center to center and may depend on funding, staffing, and student demand amongst other factors.

Table A.3: Correlations Between School Climate and Average Dropout Rates

	Mean Dropout Rates							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
More Delinquency (1-5)	0.020***	0.004**						
	(0.002)	(0.002)						
More Substance Use (1-5)			0.008***	0.000				
			(0.001)	(0.001)				
Worse Caring Staff-Student Relationships (1-5)					0.008***	-0.004***		
					(0.002)	(0.001)		
Worse School Connectedness (1-5)							0.011***	-0.003**
							(0.001)	(0.001)
Constant	-0.003	0.010***	0.003	0.013***	-0.012**	0.025***	-0.014***	0.021***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.005)	(0.004)	(0.003)	(0.004)
Observations	8169	7761	8172	7765	8173	7767	8175	7769
Sample Mean	0.008	0.008	0.008	0.008	0.008	0.008	0.008	
Fixed Effects	Year	Year/School	Year	Year/School	Year	Year/School	Year	Year/School

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

Table A.4: Average suspension and dropout rates in the period before an SBHC opens for "ever-treated" and "never-treated" schools

	Control	Treated	p-value
Suspension Rate	0.06	0.08	0.079
Defiance Suspension Rate	0.02	0.03	0.163
Non-Defiance Suspension Rate	0.04	0.05	0.140
Dropout Rate	0.01	0.01	0.306

p-values are from a t-test that the treated and un-treated school means are equal

Table A.5: Suspension Rate Event Studies by Defiance vs Non-Defiance Categories

	Full Sample		Elementary		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Defiance	Defiance	Non-Defiance	Defiance	Non-Defiance	Defiance	Non-Defiance	Defiance
Treated x (τ = -3)	0.0018	-0.0017	-0.0017	0.0068	-0.0116	-0.0119	0.0103**	-0.0047
	(0.0037)	(0.0050)	(0.0073)	(0.0048)	(0.0081)	(0.0099)	(0.0049)	(0.0097)
Treated x (τ = -2)	-0.0056*	-0.0020	-0.0018	0.0042	-0.0269**	-0.0132	0.0062	0.0006
	(0.0032)	(0.0046)	(0.0045)	(0.0030)	(0.0105)	(0.0172)	(0.0059)	(0.0060)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	0.0007	-0.0058	0.0064	-0.0012	-0.0133	-0.0285***	0.0058	-0.0071
	(0.0035)	(0.0053)	(0.0045)	(0.0020)	(0.0102)	(0.0092)	(0.0072)	(0.0115)
Treated x (τ = 1)	0.000	-0.007	0.000	0.001	-0.008	-0.054*	0.006	-0.011
	(0.003)	(0.008)	(0.005)	(0.002)	(0.010)	(0.028)	(0.006)	(0.016)
Treated x (τ = 2)	0.003	-0.010	-0.001	-0.002	-0.002	-0.065*	0.008	-0.009
	(0.004)	(0.008)	(0.006)	(0.002)	(0.011)	(0.039)	(0.007)	(0.012)
Treated x (τ = 3)	0.003	-0.012	0.001	-0.001	0.024	-0.105**	0.003	-0.016
	(0.007)	(0.011)	(0.006)	(0.003)	(0.016)	(0.048)	(0.012)	(0.020)
χ^2	0.692	2.795	6.143	7.747	9.502	24.748	3.039	0.778
p-value	0.952	0.593	0.189	0.101	0.050	0.000	0.551	0.941
Pre-Period Control Mean	0.035	0.021	0.023	0.010	0.044	0.027	0.039	0.026
Observations	1109	1109	410	410	184	184	515	515

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table A.6: Middle School Suspension Rates by All Available Categories of Suspension-Type

	(1) Any Offense	(2) Defiance	(3) Non-Defiance (All)	(4) Violence	(5) Weapon Poss.	(6) Illicit Drug
Treated X Post	-0.0271	-0.0217	-0.0056	-0.0430***	0.0014	0.0009
	(0.0163)	(0.0136)	(0.0102)	(0.0151)	(0.0022)	(0.0031)
Fraction FRPM	0.1508	0.0703	0.0816	0.1318	0.0181***	0.0131
	(0.0931)	(0.0628)	(0.0785)	(0.0930)	(0.0064)	(0.0313)
Fraction Minority	-0.0594	-0.0812	0.0187	0.0534	-0.0110	-0.0068
	(0.1520)	(0.1154)	(0.1039)	(0.1375)	(0.0168)	(0.0407)
School Size	0.000**	0.000**	-0.000	0.000	0.000**	0.000**
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	-0.033	-0.038	0.007	-0.047	-0.006	-0.022
	(0.056)	(0.070)	(0.042)	(0.037)	(0.009)	(0.021)
School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Period Control Mean	0.0720	0.0289	0.0432	0.0545	0.0043	0.0096
R^2	0.001	0.000	0.032	0.111	0.000	0.009
Observations	168	168	168	168	168	168

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

Table A.7: High School Suspension Rates by All Available Categories of Suspension-Type

	(1) Any Offense	(2) Defiance	(3) Non-Defiance (All)	(4) Violence	(5) Weapon Poss.	(6) Illicit Drug
Treated X Post	-0.0016	-0.0060	0.0045	0.0011	0.0005	-0.0009
	(0.0087)	(0.0108)	(0.0045)	(0.0051)	(0.0007)	(0.0029)
Fraction FRPM	0.0482	0.0589	-0.0108	0.0189	0.0017	-0.0170
	(0.0481)	(0.0434)	(0.0223)	(0.0376)	(0.0024)	(0.0201)
Fraction Minority	-0.0861	-0.1314	0.0464	-0.0116	-0.0026	0.0086
	(0.0935)	(0.1393)	(0.0605)	(0.0561)	(0.0055)	(0.0347)
School Size	0.000	-0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Constant	0.200***	0.108	0.091**	0.100**	0.005	0.094***
	(0.070)	(0.103)	(0.042)	(0.046)	(0.004)	(0.024)
School Characteristics	Yes	Yes	Yes	Yes	Yes	Yes
Pre-Period Control Mean	0.0682	0.0255	0.0427	0.0419	0.0022	0.0181
R^2	0.048	0.190	0.005	0.089	0.092	0.033
Observations	416	416	416	416	416	416

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

Table A.8: Repeat Suspension Rates Event Study Estimates by Sample

	All Grades		Elementary		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Baseline	Demographics	Baseline	Demographics	Baseline	Demographics	Baseline	Demographics
Treated x (τ = -3)	0.0491	0.0473	0.4755	0.1266	-0.2294	-0.1089	-0.0933	-0.1069
	(0.1368)	(0.1453)	(0.4820)	(0.1484)	(0.1948)	(0.1996)	(0.0842)	(0.0867)
Treated x (τ = -2)	0.0419	0.0007	0.3327	0.0129	-0.2798*	-0.1243	-0.0479	-0.0883
	(0.1074)	(0.1045)	(0.2971)	(0.1568)	(0.1675)	(0.1756)	(0.0951)	(0.0852)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.1240**	-0.1138*	-0.1858	-0.1616	-0.2250**	-0.2745**	-0.0725	-0.1096
	(0.0517)	(0.0583)	(0.1293)	(0.1392)	(0.1137)	(0.1272)	(0.0590)	(0.0753)
Treated x (τ = 1)	-0.169**	-0.154	-0.359***	-0.486***	-0.485**	-0.553***	0.043	0.014
	(0.085)	(0.094)	(0.120)	(0.145)	(0.203)	(0.201)	(0.116)	(0.124)
Treated x (τ = 2)	-0.285***	-0.250**	-0.416***	-0.316*	-0.654***	-0.704***	-0.079	-0.123
	(0.093)	(0.107)	(0.152)	(0.187)	(0.187)	(0.160)	(0.112)	(0.122)
Treated x (τ = 3)	-0.279**	-0.320***	-0.307	-0.309*	-0.755***	-0.811***	-0.124	-0.167
	(0.118)	(0.123)	(0.192)	(0.159)	(0.181)	(0.109)	(0.086)	(0.127)
χ^2	11.669	9.084	11.898	14.972	18.031	71.031	6.760	9.624
p-value	0.020	0.059	0.018	0.005	0.001	0.000	0.149	0.047
Pre-Period Control Mean	1.460	1.460	1.440	1.440	1.634	1.634	1.395	1.395
Observations	852	852	303	301	166	166	383	382

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table A.9: Repeat Suspension Rates Event Study Estimates by Suspension Category

	All Grades		Elementary		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Non-Defiance	Defiance	Non-Defiance	Defiance	Non-Defiance	Defiance	Non-Defiance	Defiance
Treated x (τ = -3)	-0.1014	-0.1060	0.2006	0.0498	-0.0096	-0.2035	-0.1135	-0.0840
	(0.1087)	(0.0881)	(0.1731)	(0.2603)	(0.4199)	(0.1565)	(0.1255)	(0.1151)
Treated x (τ = -2)	-0.0309	-0.0550	0.1620	-0.5770	-0.2216	-0.0907	-0.1208	0.0349
	(0.1374)	(0.0924)	(0.2212)	(0.4237)	(0.2605)	(0.1287)	(0.1123)	(0.1007)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.1402	-0.2090*	-0.0598	-0.2713	-0.5075**	-0.1742	-0.1450	0.0072
	(0.0903)	(0.1092)	(0.1356)	(0.5010)	(0.2200)	(0.2444)	(0.1154)	(0.0948)
Treated x (τ = 1)	-0.157	-0.186*	-0.482*	0.063	-0.855**	-0.459***	0.069	-0.115
	(0.146)	(0.106)	(0.262)	(0.439)	(0.368)	(0.133)	(0.182)	(0.134)
Treated x (τ = 2)	-0.297*	-0.400***	-0.321	-0.246	-1.066***	-0.491**	-0.124	-0.173
	(0.175)	(0.136)	(0.276)	(0.653)	(0.373)	(0.224)	(0.184)	(0.184)
Treated x (τ = 3)	-0.380	-0.354**	-0.206	-0.144	-1.340***		-0.279	-0.215
	(0.237)	(0.175)	(0.231)	(0.286)	(0.336)		(0.296)	(0.259)
χ^2	4.455	9.666	4.083	1.995	17.717	12.220	22.357	2.164
p-value	0.348	0.046	0.395	0.737	0.001	0.007	0.000	0.706
Pre-Period Control Mean	1.578	1.314	1.521	1.349	1.842	1.374	1.500	1.284
Observations	847	645	298	173	165	139	381	313

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table A.10: Dropout Rates: Event Study Estimates by Sample

	All Grades		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)
	Baseline	Demographics	Baseline	Demographics	Baseline	Demographics
Treated x (τ = -3)	0.0011	0.0014	0.0053	0.0052	0.0003	0.0013
	(0.0019)	(0.0019)	(0.0042)	(0.0038)	(0.0027)	(0.0035)
Treated x (τ = -2)	0.0017	0.0010	-0.0003	-0.0000	0.0055	0.0072
	(0.0030)	(0.0040)	(0.0012)	(0.0016)	(0.0051)	(0.0054)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0013	0.0006	0.0021	0.0024	-0.0019	-0.0010
	(0.0011)	(0.0028)	(0.0021)	(0.0019)	(0.0017)	(0.0026)
Treated x (τ = 1)	-0.000	0.002	0.000	-0.000	0.002	0.010
	(0.001)	(0.005)	(0.002)	(0.002)	(0.002)	(0.010)
Treated x (τ = 2)	-0.000	0.010			0.004	0.010
	(0.002)	(0.007)			(0.002)	(0.008)
Treated x (τ = 3)	-0.017***	0.011***			-0.014***	0.005*
	(0.003)	(0.003)			(0.004)	(0.003)
χ^2	43.580	15.729	1.867	2.605	142.763	15.119
p-value	0.000	0.003	0.393	0.272	0.000	0.004
Pre-Period Control Mean	0.010	0.010	0.004	0.004	0.014	0.014
Observations	348	348	110	110	205	200

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table A.11: Suspension Rates: Event Study Estimated by Gender

	All Grades		Elementary		Middle		High	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Male	Female	Male	Female	Male	Female	Male	Female
Treated x (τ = -3)	0.0020	0.0004	0.0094	-0.0011	-0.0332*	-0.0210**	0.0097	0.0079
	(0.0101)	(0.0049)	(0.0200)	(0.0055)	(0.0199)	(0.0101)	(0.0161)	(0.0088)
Treated x (τ = -2)	-0.0132*	-0.0088	0.0013	-0.0024	-0.0435	-0.0336	-0.0060	-0.0010
	(0.0078)	(0.0054)	(0.0112)	(0.0023)	(0.0329)	(0.0247)	(0.0115)	(0.0069)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0067	-0.0038	0.0093	0.0046	-0.0489***	-0.0257**	-0.0018	-0.0023
	(0.0082)	(0.0039)	(0.0105)	(0.0054)	(0.0158)	(0.0101)	(0.0137)	(0.0055)
Treated x (τ = 1)	-0.012	-0.010	0.000	-0.002	-0.086***	-0.036	-0.007	-0.022
	(0.010)	(0.007)	(0.010)	(0.005)	(0.033)	(0.022)	(0.017)	(0.020)
Treated x (τ = 2)	-0.017	-0.009*	-0.009	-0.004	-0.093*	-0.042	-0.009	-0.014
	(0.014)	(0.005)	(0.015)	(0.005)	(0.048)	(0.028)	(0.025)	(0.012)
Treated x (τ = 3)	-0.022	-0.011*	-0.007	-0.004	-0.100**	-0.064*	-0.033	-0.024*
	(0.016)	(0.006)	(0.014)	(0.005)	(0.046)	(0.033)	(0.025)	(0.013)
χ^2	2.469	4.224	8.554	3.558	33.259	13.796	3.360	3.830
p-value	0.650	0.377	0.073	0.469	0.000	0.008	0.499	0.430
Pre-Period Control Mean	0.073	0.034	0.039	0.011	0.098	0.045	0.088	0.046
Observations	938	938	353	353	168	168	416	416

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table A.12: Suspension Rates: Difference-in-Differences for All Schools with No Mental Health Services

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	Baseline	Demographics	Non-Def.	Def.	Repeat	Repeat Non-Def.	Repeat Def	Male	Female
Treated X Post	0.0076	-0.0034	-0.0204***	0.0170**	0.9792***	2.1196***	-0.0229	-0.0133	0.0072
	(0.0088)	(0.0096)	(0.0078)	(0.0080)	(0.2260)	(0.3611)	(0.3315)	(0.0141)	(0.0059)
Pre-Period Control Mean	0.0601	0.0601	0.0354	0.0247	1.7116	1.8887	1.5781	0.0782	0.0398
Observations	171	166	166	166	122	120	96	166	166

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

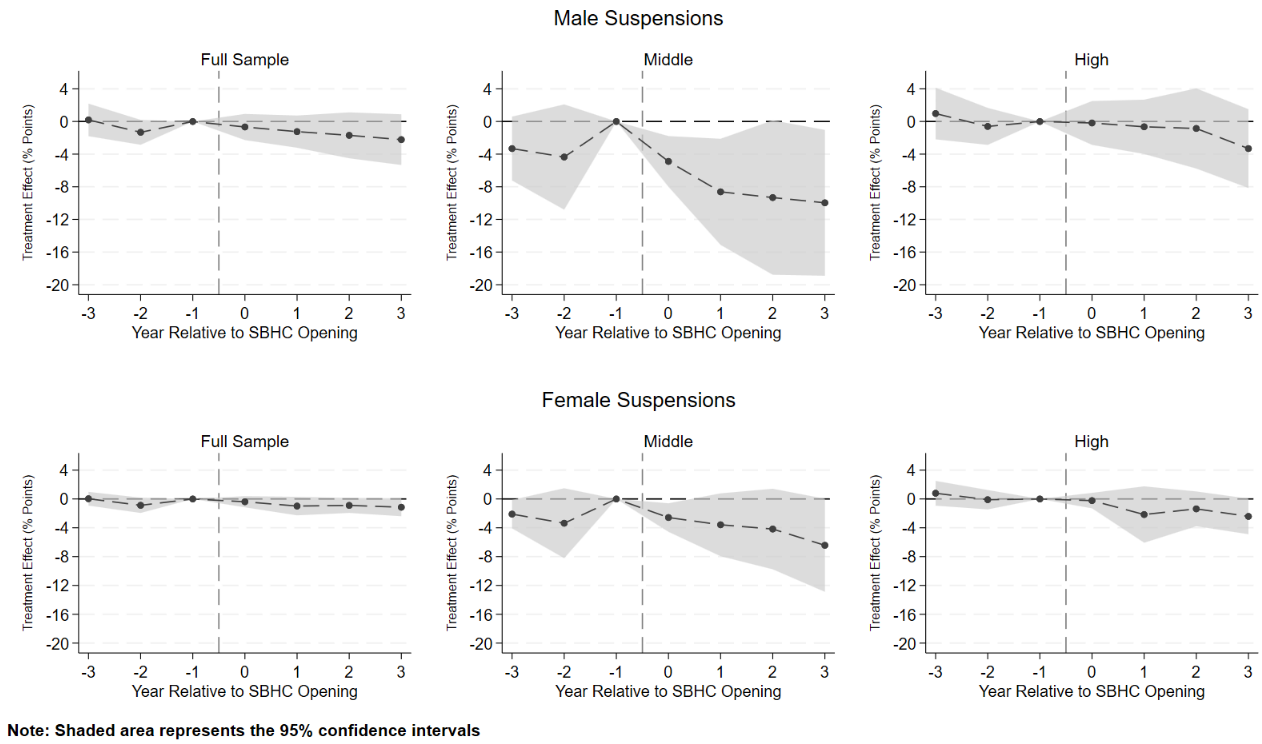

Figure A.3: This figure plots the *Event Time* coefficients from separate event studies for the outcomes of male suspension rate **(top row)** and female suspension rates **(bottom row)**. All event studies control for school fixed effects and a vector of school characteristics that includes fraction of Free and Reduced Price Meal (FRPM) students, fraction of underrepresented minority students, and total school enrollment. All lags prior to event time -3 and all leads after event time 3 are dropped from the estimation sample. Standard errors are clustered at the school level.

B Robustness Checks

B.1 Two Way Fixed Effects Regressions

Table B.13: Suspension Rates: Two Way Fixed Effects Difference-in-Differences Regressions

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def
All Grades	-0.0157***	-0.0158***	0.0038	-0.0196***	-0.2099**	-0.2497**	-0.2864***
	(0.0058)	(0.0056)	(0.0029)	(0.0065)	(0.0818)	(0.1043)	(0.0804)
Pre-Period Control Mean	0.0541	0.0541	0.0344	0.0198	1.4601	1.5779	1.3072
R^2	0.737	0.738	0.783	0.499	0.478	0.423	0.343
Observations	938	938	938	938	864	859	677
Elementary	-0.0080	-0.0084	-0.0011	-0.0074**	-0.3563*	-0.2411	-0.3366*
	(0.0050)	(0.0051)	(0.0035)	(0.0032)	(0.1860)	(0.1565)	(0.1927)
Pre-Period Control Mean	0.0253	0.0253	0.0184	0.0070	1.4403	1.5208	1.3007
R^2	0.118	0.026	0.044	0.003	0.060	0.001	0.014
Observations	354	354	354	354	314	311	207
Middle	-0.0422**	-0.0410**	0.0057	-0.0468***	-0.3168**	-0.5258***	-0.3706***
	(0.0178)	(0.0198)	(0.0102)	(0.0166)	(0.1148)	(0.1822)	(0.1254)
Pre-Period Control Mean	0.0720	0.0720	0.0432	0.0289	1.6343	1.8422	1.3743
R^2	0.159	0.003	0.018	0.001	0.001	0.007	0.000
Observations	168	168	168	168	166	165	144
High	-0.0120	-0.0107	0.0047	-0.0154*	-0.0525	-0.0889	-0.1776*
	(0.0082)	(0.0071)	(0.0032)	(0.0086)	(0.0858)	(0.1302)	(0.0933)
Pre-Period Control Mean	0.0682	0.0682	0.0427	0.0255	1.3948	1.5003	1.2793
R^2	0.045	0.040	0.005	0.090	0.063	0.019	0.008
Observations	416	416	416	416	384	383	326

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

Table B.14: Full Sample Suspension Rates: Two Way Fixed Effects Event Studies

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def
Treated x (τ = -3)	-0.0049	-0.0045	-0.0003	-0.0041	0.0463	-0.0947	-0.0144
	(0.0097)	(0.0095)	(0.0044)	(0.0079)	(0.1487)	(0.1121)	(0.1085)
Treated x (τ = -2)	-0.0098*	-0.0097*	-0.0065*	-0.0031	0.0252	0.0349	-0.1617
	(0.0057)	(0.0056)	(0.0034)	(0.0044)	(0.1184)	(0.1494)	(0.1218)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0011	-0.0015	0.0053	-0.0068	-0.1004	-0.1242	-0.1866
	(0.0062)	(0.0063)	(0.0042)	(0.0065)	(0.0804)	(0.1016)	(0.1640)
Treated x (τ = 1)	-0.012	-0.013	-0.002	-0.010	-0.151	-0.175	-0.225*
	(0.008)	(0.008)	(0.004)	(0.008)	(0.099)	(0.152)	(0.127)
Treated x (τ = 2)	-0.017*	-0.018*	-0.003	-0.014	-0.289***	-0.332**	-0.345***
	(0.010)	(0.010)	(0.006)	(0.009)	(0.098)	(0.161)	(0.129)
Treated x (τ = 3)	-0.014*	-0.014*	0.002	-0.016*	-0.206	-0.237	-0.342**
	(0.008)	(0.008)	(0.007)	(0.009)	(0.125)	(0.176)	(0.168)
Baseline Treatment Effect	-0.138***	-0.145***	-0.091***	-0.054	0.579	-0.110	-0.545
	(0.005)	(0.030)	(0.016)	(0.038)	(0.479)	(0.879)	(0.495)
School Characteristics		X	X	X	X	X	X
χ^2	1.405	1.416	1.579	0.927	2.501	1.362	2.321
p-value	0.235	0.231	0.183	0.450	0.045	0.250	0.060
Pre-Period Control Mean	0.054	0.054	0.034	0.020	1.460	1.578	1.307
R^2	0.749	0.752	0.789	0.530	0.482	0.431	0.378
Observations	938	938	938	938	864	859	677

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table B.15: Elementary Suspension Rates: Two Way Fixed Effects Event Studies

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def
Treated x (τ = -3)	0.0085	0.0090	0.0044	0.0046	0.4417	0.1815	-0.1842
	(0.0092)	(0.0097)	(0.0079)	(0.0037)	(0.4797)	(0.1641)	(0.2812)
Treated x (τ = -2)	-0.0009	-0.0003	-0.0017	0.0014	0.3296	0.4760	-0.6323
	(0.0076)	(0.0079)	(0.0053)	(0.0051)	(0.3133)	(0.3196)	(0.3900)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	0.0119**	0.0115**	0.0143**	-0.0028	-0.1720	-0.0826	-0.5017
	(0.0055)	(0.0055)	(0.0058)	(0.0032)	(0.1805)	(0.1706)	(0.4817)
Treated x (τ = 1)	-0.000	-0.001	-0.000	-0.000	-0.265	-0.248	-0.453
	(0.005)	(0.005)	(0.005)	(0.002)	(0.179)	(0.242)	(0.310)
Treated x (τ = 2)	-0.008	-0.009	-0.004	-0.005*	-0.438**	-0.425*	-0.634*
	(0.007)	(0.007)	(0.007)	(0.002)	(0.178)	(0.236)	(0.352)
Treated x (τ = 3)	-0.005	-0.005	-0.002	-0.003	-0.170	-0.109	-0.826**
	(0.006)	(0.006)	(0.006)	(0.003)	(0.237)	(0.243)	(0.316)
Baseline Treatment Effect	-0.025***	0.008	0.012	-0.004	0.448	0.348	-1.458
	(0.004)	(0.030)	(0.025)	(0.025)	(1.473)	(1.691)	(2.129)
School Characteristics		X	X	X	X	X	X
χ^2	2.399	2.483	2.872	1.944	1.756	0.928	1.899
p-value	0.061	0.054	0.031	0.116	0.151	0.455	0.126
Pre-Period Control Mean	0.025	0.025	0.018	0.007	1.440	1.521	1.301
R^2	0.058	0.060	0.077	0.013	0.011	0.000	0.003
Observations	354	354	354	354	314	311	207

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table B.16: Middle School Suspension Rates: Two Way Fixed Effects Event Studies

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def
Treated x (τ = -3)	-0.0172	-0.0098	-0.0037	-0.0061	-0.1224	-0.2818	-0.1167
	(0.0196)	(0.0210)	(0.0106)	(0.0172)	(0.1933)	(0.3595)	(0.2005)
Treated x (τ = -2)	-0.0313	-0.0259	-0.0179*	-0.0079	-0.1511	-0.2625	-0.0690
	(0.0187)	(0.0180)	(0.0092)	(0.0146)	(0.1802)	(0.3337)	(0.1525)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0297	-0.0261	-0.0141*	-0.0119	-0.0851	-0.1351	-0.1511
	(0.0195)	(0.0246)	(0.0075)	(0.0224)	(0.1348)	(0.2407)	(0.1240)
Treated x (τ = 1)	-0.042*	-0.041*	-0.015	-0.026	-0.274	-0.352	-0.375**
	(0.024)	(0.020)	(0.012)	(0.023)	(0.219)	(0.414)	(0.181)
Treated x (τ = 2)	-0.049	-0.042	-0.009	-0.033	-0.297	-0.442	-0.221
	(0.030)	(0.028)	(0.014)	(0.026)	(0.223)	(0.420)	(0.214)
Treated x (τ = 3)	-0.043	-0.028	0.014	-0.042	-0.350**	-0.486	-0.496***
	(0.026)	(0.026)	(0.027)	(0.030)	(0.162)	(0.378)	(0.160)
Baseline Treatment Effect	0.107***	0.047	0.075**	-0.028	0.469	-0.159	0.867*
	(0.014)	(0.064)	(0.029)	(0.061)	(0.506)	(0.849)	(0.468)
School Characteristics		X	X	X	X	X	X
χ^2	1.297	1.418	1.763	0.640	1.407	0.437	3.260
p-value	0.296	0.255	0.165	0.639	0.258	0.780	0.026
Pre-Period Control Mean	0.072	0.072	0.043	0.029	1.634	1.842	1.374
R^2	0.075	0.002	0.013	0.000	0.000	0.004	0.002
Observations	168	168	168	168	166	165	144

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table B.17: High School Suspension Rates: Two Way Fixed Effects Event Studies

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def
Treated x (τ = -3)	-0.0109	-0.0098	-0.0015	-0.0083	-0.1039	-0.1668	0.1160
	(0.0178)	(0.0174)	(0.0063)	(0.0148)	(0.0941)	(0.1587)	(0.1458)
Treated x (τ = -2)	-0.0080	-0.0086	-0.0051	-0.0035	-0.0728	-0.0708	0.0287
	(0.0079)	(0.0080)	(0.0050)	(0.0071)	(0.1151)	(0.1720)	(0.1285)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	0.0002	-0.0002	0.0048	-0.0050	-0.0511	-0.0775	0.0020
	(0.0095)	(0.0097)	(0.0060)	(0.0120)	(0.0878)	(0.1458)	(0.1489)
Treated x (τ = 1)	-0.010	-0.010	-0.001	-0.009	0.023	0.061	-0.030
	(0.013)	(0.012)	(0.006)	(0.014)	(0.132)	(0.174)	(0.155)
Treated x (τ = 2)	-0.011	-0.014	-0.001	-0.012	-0.140	-0.150	-0.184
	(0.018)	(0.019)	(0.010)	(0.016)	(0.121)	(0.190)	(0.121)
Treated x (τ = 3)	-0.010	-0.010	0.002	-0.012	-0.175*	-0.287**	-0.006
	(0.014)	(0.014)	(0.008)	(0.011)	(0.091)	(0.136)	(0.123)
Baseline Treatment Effect	0.067***	0.045***	-0.036***	0.082***	0.149	0.296	0.295
	(0.008)	(0.016)	(0.010)	(0.021)	(0.276)	(0.479)	(0.249)
School Characteristics		X	X	X	X	X	X
χ^2	0.409	0.392	0.641	0.473	3.173	5.478	0.774
p-value	0.801	0.814	0.635	0.756	0.019	0.001	0.546
Pre-Period Control Mean	0.068	0.068	0.043	0.025	1.395	1.500	1.279
R^2	0.039	0.035	0.004	0.088	0.099	0.046	0.081
Observations	416	416	416	416	384	383	326

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

B.2 Matched Sample Including Los Angeles Schools

Table B.18: Suspension Rates: Difference-in-Differences Regressions

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def	(8) Male	(9) Female
All Grades	-0.0087*	-0.0079*	0.0016	-0.0096*	-0.1553***	-0.1687*	-0.2002**	-0.0107	-0.0048
	(0.0046)	(0.0047)	(0.0027)	(0.0050)	(0.0583)	(0.0981)	(0.0845)	(0.0071)	(0.0034)
Pre-Period Control Mean	0.0352	0.0352	0.0236	0.0116	1.3621	1.4456	1.2893	0.0482	0.0211
Observations	1087	1086	1086	1086	901	894	577	1086	1082
Elementary	-0.0044	-0.0036	-0.0034	-0.0002	-0.2658**	-0.1089	-0.4398**	-0.0037	-0.0033
	(0.0042)	(0.0046)	(0.0037)	(0.0018)	(0.1063)	(0.1161)	(0.1982)	(0.0070)	(0.0039)
Pre-Period Control Mean	0.0182	0.0182	0.0130	0.0052	1.4018	1.4890	1.3584	0.0276	0.0065
Observations	377	372	372	372	263	258	122	372	368
Middle	-0.0420***	-0.0336***	0.0044	-0.0382**	-0.3498***	-0.6357**	-0.2206**	-0.0553***	-0.0116
	(0.0136)	(0.0124)	(0.0100)	(0.0183)	(0.0950)	(0.2786)	(0.0948)	(0.0173)	(0.0094)
Pre-Period Control Mean	0.0666	0.0666	0.0424	0.0242	1.5462	1.7406	1.4814	0.0898	0.0429
Observations	168	168	168	168	163	162	119	168	168
High	-0.0030	-0.0009	0.0053	-0.0062***	-0.0155	0.0001	-0.0151	-0.0016	0.0003
	(0.0035)	(0.0039)	(0.0034)	(0.0019)	(0.0587)	(0.0734)	(0.1028)	(0.0055)	(0.0043)
Pre-Period Control Mean	0.0353	0.0353	0.0239	0.0114	1.2682	1.3079	1.2076	0.0469	0.0229
Observations	542	542	542	542	462	461	312	542	542

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

Table B.19: Suspension Rates: Event Study Regressions (All Grades)

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def	(8) Male	(9) Female
Treated x (τ = -3)	0.0044	0.0039	0.0016	0.0022	0.0002	-0.1388	-0.1023	0.0045	0.0027
	(0.0064)	(0.0064)	(0.0035)	(0.0049)	(0.1225)	(0.0988)	(0.0771)	(0.0093)	(0.0041)
Treated x (τ = -2)	-0.0069	-0.0071	-0.0068***	-0.0003	0.0096	-0.0120	-0.0005	-0.0078	-0.0066
	(0.0048)	(0.0048)	(0.0022)	(0.0042)	(0.0993)	(0.1337)	(0.0741)	(0.0067)	(0.0046)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0047	-0.0042	0.0039	-0.0080*	-0.1283**	-0.1689*	-0.2026*	-0.0057	-0.0023
	(0.0042)	(0.0043)	(0.0025)	(0.0045)	(0.0510)	(0.0863)	(0.1069)	(0.0073)	(0.0035)
Treated x (τ = 1)	-0.010*	-0.009	-0.000	-0.009	-0.147*	-0.146	-0.185	-0.011	-0.007
	(0.006)	(0.006)	(0.003)	(0.007)	(0.086)	(0.137)	(0.115)	(0.009)	(0.005)
Treated x (τ = 2)	-0.009	-0.008	0.005	-0.012	-0.224**	-0.216	-0.344**	-0.011	-0.005
	(0.008)	(0.008)	(0.004)	(0.008)	(0.094)	(0.167)	(0.135)	(0.012)	(0.004)
Treated x (τ = 3)	-0.012	-0.012	0.003	-0.015	-0.293**	-0.323	-0.241	-0.018	-0.007
	(0.007)	(0.008)	(0.006)	(0.010)	(0.126)	(0.231)	(0.222)	(0.013)	(0.005)
χ^2	3.810	3.753	5.663	5.599	9.257	4.653	6.986	3.028	3.411
p-value	0.432	0.440	0.226	0.231	0.055	0.325	0.137	0.553	0.492
Pre-Period Control Mean	0.035		0.024	0.012	1.362	1.446	1.289	0.048	0.021
Observations	1087	1086	1086	1086	901	894	577	1086	1082

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table B.20: Suspension Rates: Event Study Regressions (Elementary)

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def	(8) Male	(9) Female
Treated x (τ = -3)	0.0044	0.0011	-0.0012	0.0023	-0.1359	-0.0659	-0.1844	0.0034	-0.0017
	(0.0090)	(0.0090)	(0.0074)	(0.0030)	(0.1232)	(0.1353)	(0.1933)	(0.0165)	(0.0040)
Treated x (τ = -2)	-0.0017	0.0009	-0.0032	0.0041	-0.0668	0.0675	0.0562	0.0026	-0.0035
	(0.0062)	(0.0065)	(0.0045)	(0.0041)	(0.1644)	(0.2145)	(0.1076)	(0.0122)	(0.0032)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	0.0020	0.0029	0.0042	-0.0013	-0.2603*	-0.1405	-0.6907**	0.0060	0.0014
	(0.0049)	(0.0053)	(0.0043)	(0.0023)	(0.1401)	(0.2081)	(0.3095)	(0.0089)	(0.0049)
Treated x (τ = 1)	-0.003	-0.004	-0.005	0.001	-0.379***	-0.240	-0.109	-0.004	-0.005
	(0.005)	(0.005)	(0.005)	(0.003)	(0.137)	(0.179)	(0.331)	(0.007)	(0.005)
Treated x (τ = 2)	-0.006	-0.004	-0.000	-0.004**	-0.210	0.043	-0.277	-0.003	-0.004
	(0.006)	(0.007)	(0.006)	(0.002)	(0.189)	(0.248)	(0.182)	(0.012)	(0.005)
Treated x (τ = 3)	-0.010	-0.009	-0.009	0.000	-0.258	-0.089	-1.518***	-0.012	-0.007
	(0.007)	(0.007)	(0.007)	(0.005)	(0.264)	(0.309)	(0.298)	(0.012)	(0.005)
χ^2	6.441	6.975	6.830	7.282	7.921	3.517	36.056	6.848	3.192
p-value	0.169	0.137	0.145	0.122	0.094	0.475	0.000	0.144	0.526
Pre-Period Control Mean	0.018		0.013	0.005	1.402	1.489	1.358	0.028	0.007
Observations	377	372	372	372	263	258	122	372	368

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table B.21: Suspension Rates: Event Study Regressions (Middle)

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def	(8) Male	(9) Female
Treated x (τ = -3)	-0.0256	-0.0139	-0.0081	-0.0059	-0.2261	-0.4536	-0.3558	-0.0220	-0.0054
	(0.0159)	(0.0171)	(0.0091)	(0.0208)	(0.2093)	(0.3677)	(0.3069)	(0.0258)	(0.0117)
Treated x (τ = -2)	-0.0340*	-0.0216	-0.0134***	-0.0082	-0.2442	-0.5502	-0.1895	-0.0289	-0.0137
	(0.0180)	(0.0228)	(0.0040)	(0.0226)	(0.1838)	(0.3501)	(0.2114)	(0.0288)	(0.0196)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0316***	-0.0232**	0.0014	-0.0247**	-0.2406*	-0.4817*	-0.1716	-0.0397***	-0.0060
	(0.0092)	(0.0091)	(0.0105)	(0.0118)	(0.1442)	(0.2525)	(0.3035)	(0.0148)	(0.0079)
Treated x (τ = 1)	-0.056**	-0.050**	-0.008	-0.042	-0.449*	-0.760	-0.310*	-0.078**	-0.021
	(0.024)	(0.025)	(0.015)	(0.033)	(0.250)	(0.528)	(0.185)	(0.034)	(0.020)
Treated x (τ = 2)	-0.053*	-0.050*	0.008	-0.058	-0.735***	-1.104*	-0.523**	-0.081**	-0.022
	(0.028)	(0.029)	(0.014)	(0.036)	(0.254)	(0.597)	(0.232)	(0.040)	(0.018)
Treated x (τ = 3)	-0.061**	-0.057**	0.041	-0.098**	-0.901***	-1.581**	-0.594***	-0.081**	-0.036**
	(0.027)	(0.025)	(0.028)	(0.047)	(0.133)	(0.627)	(0.199)	(0.034)	(0.017)
χ^2	33.633	13.894	5.069	33.097	47.454	6.484	13.247	14.538	6.981
p-value	0.000	0.008	0.280	0.000	0.000	0.166	0.010	0.006	0.137
Pre-Period Control Mean	0.067		0.042	0.024	1.546	1.741	1.481	0.090	0.043
Observations	168	168	168	168	163	162	119	168	168

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table B.22: Suspension Rates: Event Study Regressions (High)

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def	(8) Male	(9) Female
Treated x (τ = -3)	0.0095	0.0095	0.0055	0.0040	-0.0794	-0.0867	-0.0132	0.0102	0.0084
	(0.0095)	(0.0094)	(0.0040)	(0.0077)	(0.0730)	(0.0972)	(0.0893)	(0.0125)	(0.0062)
Treated x (τ = -2)	-0.0015	0.0003	-0.0025	0.0027	-0.0768	-0.0770	0.0231	0.0004	-0.0003
	(0.0056)	(0.0054)	(0.0026)	(0.0054)	(0.0804)	(0.0997)	(0.1057)	(0.0075)	(0.0065)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0013	0.0004	0.0066*	-0.0062	-0.0251	-0.0284	0.0087	-0.0001	0.0019
	(0.0054)	(0.0055)	(0.0037)	(0.0067)	(0.0427)	(0.0521)	(0.1095)	(0.0106)	(0.0051)
Treated x (τ = 1)	-0.004	-0.002	0.004	-0.006	0.030	0.080	-0.068	-0.000	-0.003
	(0.008)	(0.008)	(0.005)	(0.010)	(0.103)	(0.123)	(0.180)	(0.012)	(0.007)
Treated x (τ = 2)	-0.000	0.003	0.011*	-0.008	-0.003	0.024	-0.073	0.002	0.005
	(0.011)	(0.012)	(0.006)	(0.010)	(0.080)	(0.101)	(0.123)	(0.020)	(0.005)
Treated x (τ = 3)	-0.005	-0.005	0.004	-0.009	0.001	0.041	-0.016	-0.010	-0.000
	(0.012)	(0.015)	(0.008)	(0.012)	(0.067)	(0.072)	(0.128)	(0.027)	(0.006)
χ^2	1.437	1.462	6.598	1.621	0.654	1.797	0.489	0.912	2.416
p-value	0.838	0.833	0.159	0.805	0.957	0.773	0.975	0.923	0.660
Pre-Period Control Mean	0.035		0.024	0.011	1.268	1.308	1.208	0.047	0.023
Observations	542	542	542	542	462	461	312	542	542

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

B.3 Matched Never-Treated Schools

Table B.23: Suspension Rates: Difference-in-Differences Regressions

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def	(8) Male	(9) Female
All Grades	-0.0128**	-0.0125*	0.0003	-0.0129*	-0.1763***	-0.2793***	-0.1334	-0.0174*	-0.0078*
	(0.0062)	(0.0064)	(0.0033)	(0.0069)	(0.0607)	(0.1047)	(0.0836)	(0.0096)	(0.0046)
Pre-Period Control Mean	0.0494	0.0494	0.0342	0.0152	1.4172	1.4816	1.2837	0.0680	0.0298
Observations	774	774	774	774	712	708	554	774	771
Elementary	-0.0028	-0.0037	-0.0037	-0.0001	-0.2838**	-0.1850*	-0.6289	-0.0096	0.0001
	(0.0067)	(0.0059)	(0.0044)	(0.0018)	(0.1203)	(0.0971)	(0.4888)	(0.0098)	(0.0052)
Pre-Period Control Mean	0.0275	0.0273	0.0213	0.0060	1.4559	1.5401	1.2613	0.0427	0.0106
Observations	229	226	226	226	197	195	103	226	224
Middle	-0.0458***	-0.0384***	-0.0040	-0.0346*	-0.3249***	-0.6626**	-0.0832	-0.0494**	-0.0279***
	(0.0139)	(0.0144)	(0.0089)	(0.0205)	(0.0976)	(0.2638)	(0.1129)	(0.0199)	(0.0098)
Pre-Period Control Mean	0.0825	0.0825	0.0549	0.0275	1.5712	1.6604	1.4395	0.1127	0.0509
Observations	167	167	167	167	162	162	145	167	167
High	-0.0053	-0.0090**	0.0033	-0.0122***	0.0384	0.0488	-0.0199	-0.0105*	-0.0065
	(0.0041)	(0.0043)	(0.0028)	(0.0026)	(0.0656)	(0.0799)	(0.0896)	(0.0055)	(0.0056)
Pre-Period Control Mean	0.0466	0.0440	0.0315	0.0125	1.3159	1.3551	1.2303	0.0588	0.0281
Observations	378	374	374	374	339	338	297	374	373

Standard errors in parentheses. Observations are at the school level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

Table B.24: Suspension Rates: Event Study Regressions (All Grades)

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def	(8) Male	(9) Female
Treated x (τ = -3)	0.0033	0.0025	0.0023	0.0002	0.0845	-0.0447	-0.0231	0.0004	0.0045
	(0.0075)	(0.0078)	(0.0046)	(0.0061)	(0.1460)	(0.1140)	(0.0865)	(0.0110)	(0.0057)
Treated x (τ = -2)	-0.0124**	-0.0072	-0.0091***	0.0019	0.1142	0.1089	-0.0299	-0.0042	-0.0120*
	(0.0058)	(0.0059)	(0.0032)	(0.0059)	(0.1338)	(0.1747)	(0.0779)	(0.0089)	(0.0065)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0072	-0.0071	0.0024	-0.0094*	-0.1424**	-0.2473**	-0.1737	-0.0099	-0.0048
	(0.0053)	(0.0056)	(0.0031)	(0.0055)	(0.0593)	(0.1021)	(0.1104)	(0.0096)	(0.0049)
Treated x (τ = 1)	-0.015*	-0.017**	-0.002	-0.015	-0.178	-0.298*	-0.133	-0.021*	-0.014**
	(0.008)	(0.008)	(0.004)	(0.009)	(0.116)	(0.165)	(0.128)	(0.012)	(0.006)
Treated x (τ = 2)	-0.015	-0.015	0.004	-0.018*	-0.293***	-0.432**	-0.210*	-0.022	-0.007
	(0.011)	(0.011)	(0.005)	(0.010)	(0.108)	(0.186)	(0.115)	(0.017)	(0.007)
Treated x (τ = 3)	-0.024**	-0.020	0.006	-0.026	-0.453***	-0.708**	-0.078	-0.033	-0.007
	(0.012)	(0.012)	(0.009)	(0.016)	(0.144)	(0.298)	(0.246)	(0.021)	(0.011)
χ^2	7.204	7.850	5.005	4.719	14.110	7.904	5.335	5.707	5.729
p-value	0.125	0.097	0.287	0.317	0.007	0.095	0.255	0.222	0.220
Pre-Period Control Mean	0.049		0.034	0.015	1.417	1.482	1.284	0.068	0.030
Observations	774	774	774	774	712	708	554	774	771

Standard errors in parentheses. Observations are at the school level. * $p<0.1$, ** $p<0.05$, *** $p<0.01$

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table B.25: Suspension Rates: Event Study Regressions (Elementary))

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def	(8) Male	(9) Female
Treated x (τ = -3)	-0.0027	-0.0109	-0.0093	-0.0015	0.0760	0.4085**	-0.1601	-0.0103	-0.0065*
	(0.0122)	(0.0134)	(0.0105)	(0.0043)	(0.1170)	(0.1899)	(0.2046)	(0.0233)	(0.0038)
Treated x (τ = -2)	-0.0094	-0.0091	-0.0101	0.0010	0.3985	0.4232	0.1285	-0.0143	-0.0043
	(0.0076)	(0.0081)	(0.0067)	(0.0042)	(0.4757)	(0.4135)	(0.2232)	(0.0161)	(0.0035)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	0.0039	0.0029	0.0008	0.0020	-0.3649	-0.2176	-1.2375	-0.0011	0.0029
	(0.0072)	(0.0070)	(0.0039)	(0.0042)	(0.2838)	(0.2415)	(0.8728)	(0.0109)	(0.0063)
Treated x (τ = 1)	-0.002	-0.003	-0.004	0.000	-0.475**	-0.421***	-0.299	-0.006	-0.004
	(0.008)	(0.008)	(0.006)	(0.004)	(0.196)	(0.152)	(0.495)	(0.011)	(0.007)
Treated x (τ = 2)	-0.007	-0.007	-0.005	-0.002	-0.239	-0.139	-0.039	-0.017	0.001
	(0.010)	(0.010)	(0.007)	(0.005)	(0.172)	(0.164)	(0.101)	(0.016)	(0.008)
Treated x (τ = 3)	-0.008	-0.009	-0.007	-0.002	-0.276*	-0.177		-0.017	-0.002
	(0.010)	(0.010)	(0.006)	(0.005)	(0.164)	(0.243)		(0.016)	(0.008)
χ^2	4.556	5.871	1.932	7.284	6.058	7.711	2.366	5.280	2.752
p-value	0.336	0.209	0.748	0.122	0.109	0.103	0.500	0.260	0.600
Pre-Period Control Mean	0.027		0.021	0.006	1.456	1.540	1.261	0.043	0.011
Observations	229	226	226	226	197	195	103	226	224

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table B.26: Suspension Rates: Event Study Regressions (Middle)

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def	(8) Male	(9) Female
Treated x (τ = -3)	-0.0199	-0.0098	-0.0116	0.0017	-0.0344	-0.1978	-0.0860	-0.0107	-0.0087
	(0.0161)	(0.0204)	(0.0094)	(0.0242)	(0.2143)	(0.4037)	(0.2161)	(0.0304)	(0.0157)
Treated x (τ = -2)	-0.0294	-0.0273	-0.0207***	-0.0063	-0.2128	-0.4648	-0.1705	-0.0327	-0.0221
	(0.0202)	(0.0216)	(0.0079)	(0.0198)	(0.1749)	(0.3364)	(0.1366)	(0.0261)	(0.0194)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0382***	-0.0298**	-0.0095	-0.0204	-0.1661	-0.3805	0.0619	-0.0308*	-0.0287**
	(0.0128)	(0.0143)	(0.0088)	(0.0136)	(0.1167)	(0.2346)	(0.2522)	(0.0185)	(0.0136)
Treated x (τ = 1)	-0.056**	-0.047**	-0.008	-0.039	-0.462**	-0.898*	-0.353**	-0.059**	-0.034*
	(0.023)	(0.023)	(0.011)	(0.031)	(0.214)	(0.472)	(0.165)	(0.029)	(0.019)
Treated x (τ = 2)	-0.060*	-0.062**	-0.007	-0.055	-0.680***	-1.177**	-0.388*	-0.088**	-0.037
	(0.031)	(0.030)	(0.015)	(0.035)	(0.245)	(0.526)	(0.216)	(0.038)	(0.023)
Treated x (τ = 3)	-0.072**	-0.069**	0.019	-0.088*	-0.741***	-1.488**	-0.117	-0.076*	-0.063***
	(0.028)	(0.031)	(0.025)	(0.051)	(0.234)	(0.660)	(0.133)	(0.045)	(0.022)
χ^2	16.830	6.899	6.301	6.173	11.642	6.180	8.372	5.555	12.337
p-value	0.002	0.141	0.178	0.187	0.020	0.186	0.039	0.235	0.015
Pre-Period Control Mean	0.082		0.055	0.028	1.571	1.660	1.440	0.113	0.051
Observations	167	167	167	167	162	162	145	167	167

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

Table B.27: Suspension Rates: Event Study Regressions (High)

	(1) Baseline	(2) Demographics	(3) Non-Def.	(4) Def.	(5) Repeat	(6) Repeat Non-Def.	(7) Repeat Def	(8) Male	(9) Female
Treated x (τ = -3)	0.0164	0.0161	0.0113	0.0048	-0.0286	-0.0391	-0.0297	0.0144	0.0175
	(0.0116)	(0.0141)	(0.0076)	(0.0092)	(0.0817)	(0.1230)	(0.1103)	(0.0155)	(0.0124)
Treated x (τ = -2)	-0.0086	-0.0043	-0.0031	-0.0012	-0.0754	-0.0678	-0.0005	-0.0022	-0.0066
	(0.0080)	(0.0076)	(0.0039)	(0.0058)	(0.0783)	(0.1111)	(0.0698)	(0.0085)	(0.0073)
Treated x (τ = -1)	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.	ref.
Treated x (τ = 0)	-0.0009	-0.0053	0.0063	-0.0115	-0.0037	-0.0082	-0.0156	-0.0056	-0.0037
	(0.0070)	(0.0101)	(0.0047)	(0.0127)	(0.0456)	(0.0571)	(0.0993)	(0.0177)	(0.0063)
Treated x (τ = 1)	-0.008	-0.014	0.002	-0.015	0.116	0.140	-0.019	-0.011	-0.014
	(0.011)	(0.016)	(0.004)	(0.016)	(0.117)	(0.131)	(0.138)	(0.020)	(0.012)
Treated x (τ = 2)	-0.005	-0.009	0.010	-0.019	0.026	0.041	-0.052	-0.012	-0.004
	(0.018)	(0.021)	(0.007)	(0.017)	(0.100)	(0.130)	(0.096)	(0.033)	(0.010)
Treated x (τ = 3)	-0.020	-0.032	0.004	-0.035	-0.018	-0.050	0.230***	-0.057	-0.005
	(0.028)	(0.038)	(0.009)	(0.031)	(0.043)	(0.060)	(0.072)	(0.071)	(0.006)
χ^2	4.123	3.712	4.495	2.464	2.016	3.377	0.408	2.890	2.938
p-value	0.390	0.446	0.343	0.651	0.733	0.497	0.939	0.576	0.568
Pre-Period Control Mean	0.047		0.032	0.012	1.316	1.355	1.230	0.059	0.028
Observations	378	374	374	374	339	338	297	374	373

Standard errors in parentheses. Observations are at the school level. * p<0.1, ** p<0.05, *** p<0.01

χ^2 and p-value come from a test that the coefficients on Treatment X Event-Time for all post-event years are jointly equal to 0.

C Data Cleaning, Propensity Score Matching, and Factor Selection

C.1 Data Cleaning

To assign county-district-school codes to each SBHC, I use an iterative fuzzy string matching process that aims to select the closest matching school for each SBHC based on the similarity between addresses, school names, and gradespans. The matching process begins by identifying the set of *all potential matching schools* for each SBHC using the composite addresses of the SBHC and the school. Each potential match is assigned a "similarity score" between zero and one, with a higher value indicating a "better match".[37] The potential matches with the highest similarity score for each school are marked as "top matches". Each potential match is also assigned separate similarity scores based on the school name, gradespan, city, county, and zip code fields.[38]

There are eleven steps of matching based on these similarity scores, with each step imposing less stringent requirements than the prior one. The process is designed to generate matches that are as exact as possible and ends when a single, unique match is identified for each SBHC. Table C.1 outlines each of the eleven-steps, including the fraction of the sample that has an identified match at each step.

Table C.1: Matching criteria for nine iterations of fuzzy string matching

Iteration	Matching Criteria	% of Sample Matched
1	"Top match" with address SS > 0.917 - **If multiple matches:** select "top match" with highest school name SS - **If multiple matches with same address and school name SS:** select the first "top match" in the list.	58.39
2	Exact match on city and school name	11.54
3	Exact match on city and zip code + highest gradespan SS *if school name SS > 0.45*	3.5
4	Exact match on city + partially matching zip code *if school name SS > 0.45*	4.9
5	Exact match on city + highest school name SS *if school name SS > 0.45*	4.2
6	Exact match on city + highest zip code SS	1.75
7	Exact match on city + partially matching zip code	9.44
8	Exact match on zip code	0.35
9	Partially matching zip code	2.8
10	Manual matching for three cases	1.05
11	For SBHCs with no matches, pick potential match with highest address SS	2.1

SS = Similarity Score. In all rows, "Exact match" indicates a similarity score of 1.
"Partially matching zip code", indicates that the zipcode of the school or SBHC is *nested* in the zipcode of the other
(eg. SBHC has a zipcode of "95121" and the school has a zipcode of "95121-1845")

At the end of the matching process each SBHC is matched to a "County-District-School" (CDS) code, which is a unique school-identifier utilized across all CDE datasets. This CDS code allows each SBHC to be attached to a panel of outcomes data on suspension rates and dropout rates for its associated primary school. It also allows school-level characteristics such as racial composition,

[37] Mechanically, similarity scores are generated with the *matchit* package in Stata which decomposes the text into bigrams before calculating a Jaccard similarity index.

[38] The "gradespan" field is a standardized field in the CDE data, but does not exist in the SBHC openings data. To generate a similarity score for this field, I generate a corresponding "gradespan" field in the SBHC data by parsing the "school name" field for keywords such as "elementary", "middle", and "high".

enrollment, and fraction of students on Free-or-Reduced-Price Lunch to be merged onto the SBHC dataset. These same CDS codes allow the data on suspensions and dropouts to be linked to the California Healthy Kids Survey data.

C.2 Propensity Score Matching and Factor Selection

Recent papers on propensity-score matching have argued for a careful selection of factors in constructing the propensity scores. Smith and Todd (2005) notes that one of the concerns with propensity score matching is that the results may be sensitive to choice of predictors and the specified prediction model. Moreover, the use of "bad predictors" can be equally problematic for matching. In order to avoid the use of bad predictors or a poorly specified model, I focus on evaluating three potential predictors that are grounded in contextual knowledge of SBHCs: the socioeconomic status of students attending a school, the racial composition of the school, and the size of the school. The first two factors are motivated by the credo that first inspired the SBHC model in the 1960s. In their incipience, SBHCs were intended to bridge gaps in healthcare access specifically in *low-income* areas and for students from *racial minority backgrounds* (Flaherty and Osher, 2003). Current guidance from the California School-Based Health Alliance suggests that this continues to be a goal of SBHCs in California; specifically, they recommend that SBHCs in California focus on "health care services for children and youth with Medi-Cal coverage" and providing "culturally competent, high-quality, first-contact primary care" with the potential to " reduce health inequities and improve health outcomes for LGBTQ+ youth, low-income youth, and youth of color" (CSBHA, 2022b). While socioeconomic status could be measured through median income, this data is not available at the school level.[39] Instead, I rely on the fraction of students at the school who receive free- or reduced-price lunch, which is a standard proxy for socioeconomic status in the academic literature.[40] The racial composition of a school is measured by the fraction of students from racial minority backgrounds, which includes students who identify as Black, Hispanic or Latinx, Filipino, Pacific Islander, American Indian or Alaskan Native.[41] There is a high correlation between this constructed metric for "racial composition" and the fraction of students receiving free-and reduced price lunch, indicating that this measure is appropriately capturing underlying facets of a school that would increase its likelihood of opening an SBHC.

The third factor, school size, is motivated by background on the process of opening an "on-site" SBHC. Since on-site SBHCs are often located *inside* a physical school building, the ability to construct an SBHC on-site necessitates either a large school or large school campus, both of which are plausibly correlated with a large student body. An additional reason that school size may be a

[39] The most granular geography for Census data on median household income is the census tract level.

[40] In California, students from households with income below 130% of the federal poverty level qualify for free meals, while students from households that fall between 130% and 185% of the federal poverty level qualify for "reduced-price" meals. (CDE, 2023)

[41] This definition is based on definitions of "underrepresented minority" or URM students commonly used at the university level in California.

reasonable predictor of opening is that if districts are concerned with improving healthcare access for as many students as possible, they would be incentivized to place the SBHC in a school with a larger concentration of the district's students. To identify the correct set of predictors to use, I run a set of logit models based on equation 3.

$$Treated_{s,t} = \alpha_{s,t} + \mathbb{X}_{s,t-1} + \varepsilon_{s,t} \tag{3}$$

where $Treated_{s,t}$ is an dummy equal to 1 if school s has an active SBHC in year t. $\mathbb{X}_{s,t-1}$ is a vector of lagged predictors for school s in year $t-1$. In the most saturated specification $\mathbb{X}_{s,t-1}$ contains 1-year lags of the fraction of students qualified for FRPM, the fraction minority students, and the total enrollment. Table C.2 shows the coefficients from a set of logit models. Columns (1) - (3) show logit models for each of the three potential factors separately. The primary takeaway from these first three specifications is that individually, each of these factors is a statistically significant predictor of an increase in the likelihood of a school having an SBHC. The χ^2 statistics indicate that 1-year lagged enrollment is the most predictive, while fraction of URM students is the least predictive.

Work from Heckman et al. (1998) suggests that one effective method of selecting appropriate predictors is by sequentially adding potential predictors to the model and testing for significance. Following this approach, Table C.2 show three logit models beginning with including only lagged fraction of FRPM students as a predictor in Column (1), and adding lagged enrollment in Column (2) and lagged fraction of minority students in Column (3). Columns (1) and (2) reveal that the lags for both fraction FRPM and total enrollment are statistically significant predictors of opening an SBHC in school s at time t. Column (3) shows that the inclusion of the lagged fraction of minority students provides no additional predictive power. This is not surprising given the high correlation between the fraction of minority students the fraction of FRPM students (with a correlation coefficient of 0.798).

Once the correct predictors have been determined, it remains to select the correct function of these predictors. Table C.3 shows potential linear and non-linear functions of lagged fraction of FRPM students and lagged enrollment. Column (1) shows the original linear function of both variables; Column (2) adds a square term for lagged enrollment to the baseline specification; Column(3) adds a square term for lagged FRPM to the baseline specification; finally, Column (4) includes square terms for both predictors. Column 2 reveals that the addition of a square term for lagged enrollment is statistically significant; however the sign of the coefficient is negative and the χ^2 statistic for the specification in Column (2) is lower than the χ^2 statistic for Column (1). This suggests that while the inclusion of a square term for lagged enrollment may be statistically significant, it does not necessarily increase the predictive power of the model.

One further check that is relevant here is which model specification is most predictive *within* each of the three gradespan types. This is worth examining here since the propensity matching process

Table C.2: Predicted Likelihood of Having an SBHC (Pooled)

	(1)	(2)	(3)
1Y FRPM Lag	2.179***		0.815
	(0.291)		(1.249)
1Y Frac URM Lag	-0.618**		
	(0.313)		
1Y Enrollment Lag	0.00106***		0.00116***
	(0.0000750)		(0.000337)
1Y FRPM Lag Q2		0.598***	
		(0.186)	
1Y FRPM Lag Q3		0.590***	
		(0.187)	
1Y FRPM Lag Q4		1.039***	
		(0.175)	
1Y Enr Lag Q2		0.256	
		(0.180)	
1Y Enr Lag Q3		-0.267	
		(0.202)	
1Y Enr Lag Q4		1.062***	
		(0.159)	
$(1YFRPMLag)^2$			0.781
			(1.044)
$(1YEnrollmentLag)^2$			-4.05e-08
			(0.000000110)
Constant	-5.557***	-4.639***	-5.546***
	(0.211)	(0.192)	(0.369)
χ^2	293.1	175.9	292.1
Observations	12570	12571	12571

Standard errors in parentheses. Observations are at the school level.

Covariates are one-year lags relative to a specific cohort event.

* p<0.1, ** p<0.05, *** p<0.01

Table C.3: Predicted Likelihood of Having an SBHC - Model Selection

	(1)	(2)	(3)	(4)
1Y FRPM Lag	1.757***	1.756***	0.824	0.815
	(0.243)	(0.243)	(1.251)	(1.249)
1Y Enrollment Lag	0.00104***	0.00115***	0.00104***	0.00116***
	(0.0000719)	(0.000337)	(0.0000725)	(0.000337)
$(1YEnrollmentLag)^2$		-3.92e-08		-4.05e-08
		(0.000000111)		(0.000000110)
$(1YFRPMLag)^2$			0.774	0.781
			(1.046)	(1.044)
Constant	-5.713***	-5.764***	-5.495***	-5.546***
	(0.193)	(0.245)	(0.326)	(0.369)
Chi Squared	293.6	288.6	296.9	292.1
Observations	12571	12571	12571	12571

Standard errors in parentheses. Observations are at the school level.

Covariates are one-year lags relative to a specific cohort event.

* p<0.1, ** p<0.05, *** p<0.01

restricts to matching schools in the same gradespan. I focus on the specifications in Columns (1) and (2) of Table C.3, as the two specifications where all predictors are statistically significant. Table C.4 runs each of those logit specifications separately for elementary, middle, and high schools. Columns (1), (3), and (5) reveal that the predictors in the linear specification are consistently significant across all three gradespans. For the non-linear model, all predictors are significant for the subsample of elementary schools, but not for middle and high schools. More concerningly, for the sample of middle schools, the addition of the square term for lagged enrollment decreases the χ^2 statistic, indicating that this model may be less predictive.

In order to ensure the use of a model with consistent predictive power, both for the whole sample and each gradespan subsample, the final logit regression follows equation 3 where the vector $\mathbb{X}_{s,t-1}$ contains the lagged fraction of students qualified for FRPM and the lagged total enrollment of a school. The final matching occurs within gradespan and limits the sample of potential control school districts to those that already have an SBHC; therefore the propensity matching implicitly takes into account grade-levels and underlying openeness to having an SBHC in addition to the selected observable predictors.

Table C.4: Predicted Likelihood of Having an SBHC (Logit)

	Elementary		Middle		High School	
	(1)	(2)	(3)	(4)	(5)	(6)
1Y FRPM Lag	3.714***	3.724***	4.227***	4.161***	0.508	4.161***
	(0.719)	(0.714)	(1.017)	(0.972)	(0.311)	(0.972)
1Y Enrollment Lag	0.000615*	0.00994***	-0.00140***	-0.00421***	0.000643***	-0.00421***
	(0.000362)	(0.00237)	(0.000536)	(0.000999)	(0.0000757)	(0.000999)
$(1Y EnrollmentLag)^2$		-0.00000803***		0.00000204***		0.00000204***
		(0.00000202)		(0.000000549)		(0.000000549)
Constant	-7.505***	-9.955***	-5.567***	-4.746***	-3.852***	-4.746***
	(0.653)	(0.991)	(0.866)	(0.869)	(0.229)	(0.869)
χ^2	29.71	38.65	20.37	31.86	78.66	31.86
Observations	7652	7652	1660	1660	2629	1660

Standard errors in parentheses. Observations are at the school level.

Covariates are one-year lags relative to a specific cohort event.

* p<0.1, ** p<0.05, *** p<0.01

D Alternative Control Groups

The primary results in this paper must be cautiously interpreted with respect to the selected control group. Tthere are two design choices in this paper that deserve further justification: (1) the choice of a propensity-matching specification instead of a simple two-way fixed effects model using never or not-yet treated units; and (2) the selection of control schools from the pool of *untreated schools* in districts that have at least one open SBHC.

The use of a propensity-score matching approach to select control schools is motivated by the expectation that there is *selection into treatment*. Specifically, since schools do not randomly receive an SBHC, the schools/districts that choose to open an SBHC may be meaningfully different from those districts that never open an SBHC. Propensity-score matching addresses this concern by matching schools on characteristics that are predictive of the likelihood of opening an SBHC. An alternate method of addressing selection is to use schools that are treated in the *future* as controls for schools that are treated earlier. This approach relies on the assumption that conditional on two schools having the same underlying propensity to *ever* open an SBHC, the *exact timing* of the open is random. This approach would fail to produce well-matched treatment and control groups if the specific timing at which an SBHC is non-random, and specifically if the timing is driven by district or school-level trends that are correlated with the outcomes of interest. The existence of a pre-trend in suspension rates that is in the same direction as the treatment further complicates the use of future treated schools, since any "decreases" in suspension rates in the control group may be capturing the pre-trend for the future-treated school, and therefore would not be a proper counterfactual for the trajectory of the treated school in absence of treatment.

Figure D.1 shows the separate event studies for treated and control schools using a "future-treated" control group. Specifically, for the cohort of schools that open an SBHC in year y the control group includes all schools that open an SBHC in year $z > y + 3$. The imposition of a three year buffer allows for the examination of a three-year post-event window in which all of the "control" schools are pure controls.[42] Figure D.1 shows a visible difference in pre-trends for treated and control schools in this sample, which is in line with the theory that the exact timing of SBHC-openings may not be random.

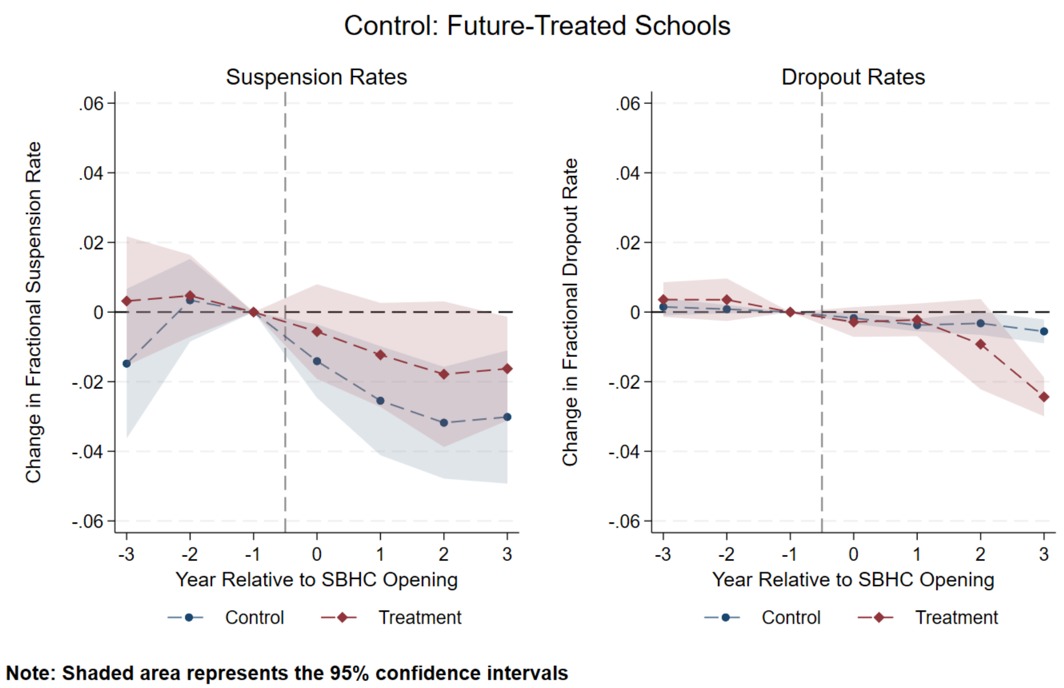

Figure D.1: This figure plots the *Event Time* coefficients from treated and control group event studies for suspension rates **(left)** and dropout rates **(right)** using *future-treated* schools as controls. Both sub-sample event studies control for school fixed effects and a vector of school characteristics that includes fraction of Free and Reduced Price Meal (FRPM) students, fraction of underrepresented minority students, and total school enrollment. All lags prior to event time -3 and all leads after event time 3 are dropped from the estimation sample. Standard errors are clustered at the school level.

Having provided some evidence that propensity-score matching is a more appropriate control group than future-treated schools in this setting, we can consider whether the restrictions placed on the propensity score matching process are appropriate. The primary matching process in this paper uses the one-year lagged fraction of FRPM students and one-year lagged school enrollment as the predictors used to construct the propensity scores; however, it also imposes two additional restrictions that theoretically strengthen the matching: (1) that matches are selected from the same gradespan as the treated school (i.e. a high school with an SBHC can be matched to a *high school*

[42]This model of using future-treated schools that open outside of a certain window is used in several recent papers including Deshpande and Li (2019) and Fadlon and Nielsen (2021). The primary goal of the buffer is to prevent violations of the Stable Unit Treatment Values Assumption (SUTVA) which requires that units do not change their treatment status after the time of the treatment.

without an SBHC but not an *middle school*).; and (2) that match for a school with an SBHC that opens in year y is selected from the pool of *never-treated* schools in districts that have at least one SBHC that opened in year $k <= y$. The first restriction has a natural motivation: both the types of services offered by SBHCs and the outcomes of interest (suspensions and dropout rates) are likely to differ meaningfully across different gradespans. Therefore gradespan mismatches could be a significant source of bias for my difference-in-differences estimates. The second restriction aims to improve the quality of matches by limiting to districts that have similar "openness" to having an SBHC. An alternate way of accomplishing this it to match within district, following recent recommendations from the propensity-score matching literature to match within the same "local labor market". Figure D.2 shows treated and control group event studies on suspension rates for a sample where controls are selected using *within-district* propensity score matching. The figure reveals that control schools matched from the same district do not have a parallel pre-trend in suspension rates to the set of treated schools.

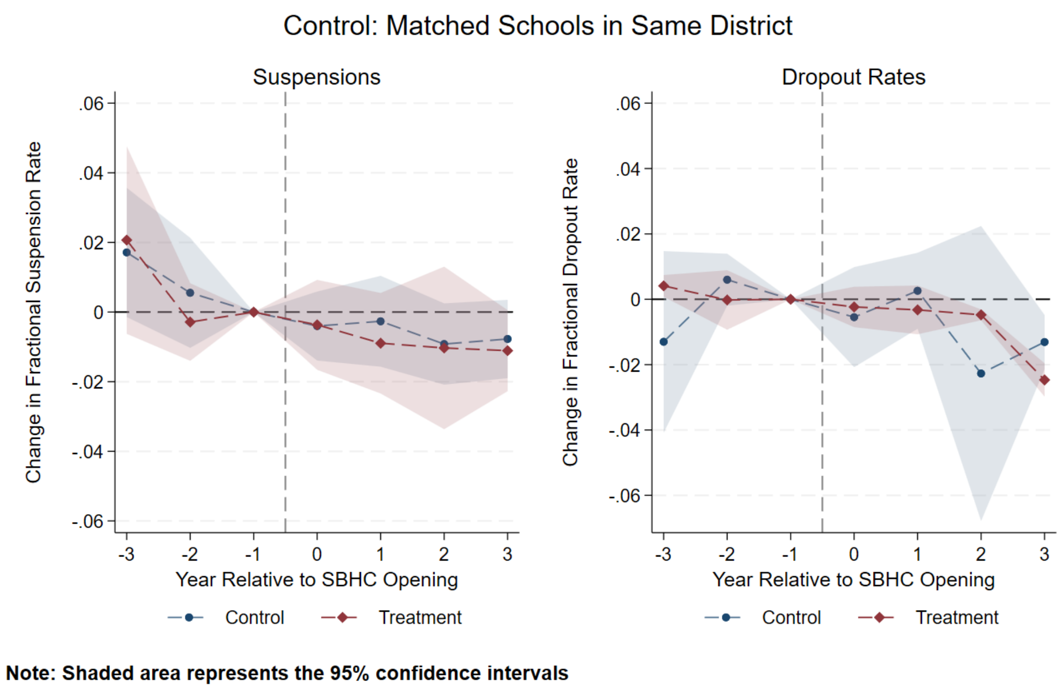

Figure D.2: This figure plots the *Event Time* coefficients from treated and control group event studies for suspension rates **(left)** and dropout rates **(right)** using control schools that are selected through 1:1 propensity-score matching from *the same school district* as each treated school. Both sub-sample event studies control for school fixed effects and a vector of school characteristics that includes fraction of Free and Reduced Price Meal (FRPM) students, fraction of underrepresented minority students, and total school enrollment. All lags prior to event time -3 and all leads after event time 3 are dropped from the estimation sample. Standard errors are clustered at the school level.

Finally, we might consider whether the restriction of matching to districts with at least one SBHC is necessary at all. The major concern with selecting concerns from districts that *never* open an SBHC during the study window is that these districts may be meaningfully different on unobservables and policies than districts that ever open an SBHC. In particular, there is a concern that if control schools in these districts have a similar predicted probability of opening an SBHC in year y but opt to not open one, this could be indiction of some alternate policy or program that was implemented in lieu of a school-based health center. If this is the case, parallel pre-trends between the treatment and control groups may be insufficient to satisfy the assumption that the trajectory of outcomes in the control schools represents the correct counterfactual for the expected trajectory of outcomes in treated schools in the absence of treatment. Table D.1 compares the 2012 sample means for a set of school and district-level covariates between districts that ever open an SBHC and districts that never open an SBHC. The sample sizes for each mean are in brackets.

Table D.1: Summary Statistics: District & School Characteristics (2012 Data)

	Ever-Treated Districts	Never-Treated Districts	p-value
Fraction FRPM	0.55	0.66	0.000
	[6531]	[1764]	
Fraction Minority	0.56	0.72	0.000
	[6455]	[1768]	
School Size (Total Enrollment)	607.43	629.18	0.133
	[6455]	[1770]	
Zip-Code Level Median Income	29703.77	27948.37	0.000
	[5966]	[1638]	
Number of Schools	17.29	75.05	0.000
	[6606]	[1779]	
Number of Elementary Schools	11.27	47.29	0.000
	[5942]	[1711]	
Number of Middle Schools	3.18	10.13	0.000
	[4900]	[1619]	
Number of High Schools	5.44	18.66	0.000
	[2214]	[863]	

p-values are from a t-test that the treated and un-treated means are equal

Number of observations in brackets under means

This table reveals that ever-treated districts are significantly different from never-treated districts across all covariates. Specifically, ever-treated district tend to be larger on average, with over double the number of high schools, over three times the number of middle schools, and nearly five times the number of elementary schools. While the average school-size is similar, the average school in a treated district has 11 percentage points more students qualifying for Free or Reduced-Price Meals, 16 percentage points more minority students and a lower zip-code-level median household income level (of around \$1,560). These differences on observable characteristics raise further concerns about the number of unobservable characteristics on which these two types of districts could differ. Although the results in section 5 are robust to matching from the sample of all never-treated schools regardless of district, baseline differences between ever-treated and never-treated districts suggest that the comparison group used in my main specification may be more theoretically comparable.

E California Department of Education Data Descriptions

E.1 Suspension Offense Categories

CDE Data Category	Offense	California Edu. Codes
Violent Incident (Injury)	Sexual Battery/Assault	48915(c)(4), 48900(n)
	Caused Physical Injury	48915(a)(1)(A)
	Committed Assault or Battery on a School Employee	48915(a)(1)(E)
	Used Force or Violence	48900(a)(2)
	Committed an act of Hate Violence	48900.3
	Hazing	48900(q)
Weapons Possession	Possession, Sale, Furnishing a Firearm	48915(c)(1)
	Possession, Sale, Furnishing a Firearm or Knife	48900(b)
	Brandishing a Knife	48915(c)(2)
	Possession of a Knife or Dangerous Object	48915(a)(1)(B)
	Possession of an Explosive	48915(c)(5)
Illicit Drug-Related	Sale of Controlled Substance	48915(c)(3)
	Possession of Controlled Substance	48915(a)(1)(C)
	Possession, Use, Sale, or Furnishing a Controlled Substance, Alcohol, Intoxicant	48900(c)
	Offering, Arranging, or Negotiating Sale of Controlled Substances, Alcohol, Intoxicants	48900(d)
	Offering, Arranging, or Negotiating Sale of Drug Paraphernalia	48900(j)
	Offering, Arranging, or Negotiating Sale of Soma	48900(p)
Other Reasons	Possession of an Imitation Firearm	48900(m)
	Possession or Use of Tobacco Products	48900(h)
	Property Damage	48900(f)
	Robbery or Extortion	48915(a)(1)(D)
	Property Theft	48900(g)
	Received Stolen Property	48900(l)
Defiance-Only	Disruption, Defiance	48900(k)(1)

Table E.1: This table shows the various offenses that are included in each "category" of suspensions defined by the California Department of Education. The third column shows the corresponding codes from *California Education Code* §48900 - 48927. The original data definitions can be found at: https://www.cde.ca.gov/ds/ad/fssd.asp

E.2 California Healthy Kids Survey

The core module of the California Healthy Kids Survey (CHKS) consists of around 155 questions that are selected to assess three pillars of developmental suppors that research has linked to "positive academic, psychosocial, and health outcomes among youth, even in high-risk environments": positive academic relationships; high expectations (academic and behavioral); and opportunities for meaningful participation and decision-making.[43]. Several papers have attempted to validate the psychometric properties of subsets of CHKS questions. One such paper comes from researchers at WestEd, the organization that lead the development of the CHKS (Mahecha and Hanson, 2020). This paper proposes the construction of a set of nine indices as weighted averages of subsets of the CHKS questions and verifies the internal consistency reliability and item bias of the constructs. I focus on four indices that are most likely to be correlated with mental health: caring staff-student relationships, school connectedness, delinquency, and substance use at school. Table E.2 partially reproduces a table from Mahecha and Hanson (2020) that lists the questions included in each index and the weight assigned to each question.[44]

In order to construct an index for each construct that takes on the same values as the questions with the index, I scale all weights to sum to one prior to taking a weighted average across all question in the index. Equation E.2 shows the equation for a given index, c, as a weighted average of a set of questions $\{Q_{c,i}\}$

$$I_c = \sum_{\forall Q_{c,i}} (Q_{c,i}) \frac{\omega_{c,i}}{\sum_{\forall c,j} \omega_{c,j}}$$

$\frac{\omega_i}{\sum_{\forall j} \omega_j}$ represents the scaled weight and $Q_{c,i}$ is the value for question i of construct c.

Responses that are missing answers for all questions in an index are assigned an index value of *missing*. For cases where an individual response contains missing answers for some but not all questions in an index, the index is re-scaled by dividing the value of the index by the sum of the weights on the questions with non-missing responses. This amounts to rescaling the weights on the questions that are answered to add up to one. I verify that this rescaling does not bias the index values by comparing the rescaled indices to indices constructed for the subset of responses with no missing questions.

Finally, for my two measures of mental health, I use two questions that are similar to the types of questions commonly used in other surveys measuring mental health. In particular, discussions with researchers who have worked closely with the CHKS suggest that the two questions on the CHKS that directly ask about mental health are drawn from similar surveys such as the Youth Risk Behavior Surveillance System (YRBSS). For these two questions, I use the individual question values as there is no obvious precedent for the conversion of these questions into a weighted index.

[43] *Source:* https://calschls.org/about/the-surveys/

[44] In Mahecha and Hanson (2020) each question is assigned a standardized factor loading from a confirmatory factor analysis model. As is standard in CFA models, the factor loading for each question comes from the correlation between that question and the underlying construct being measured. In constructing an index from a set of questions, each question is weighted by the factor loading to account for differences in how well each question captures the underlying construct.

California Healthy Kids Survey Item	Weight
Caring Staff-Student Relationships	
teacher or adult who really cares about me	0.806
teacher or adult who tells me when I do a good job	0.836
teacher or adult who notices when I'm not there	0.737
teacher or adult who always wants me to do my best	0.864
teacher or adult who listens to me when I have something to say	0.851
teacher or adult who believes that I will be a successful student	0.873
Caring Staff-Student Relationships	
I feel close to people at this school	0.649
I am happy to be at this school	0.835
I feel like I am part of this school	0.855
The teachers at this school treat students fairly	0.710
I feel safe in my school	0.735
Delinquency	
been in a physical fight at school (12 months)	0.681
been offered, sold, or given an illegal drug at school (12 months)	0.707
damaged school property on purpose at school (12 months)	0.745
carried a gun at school (12 months)	0.846
carried any other weapon at school (12 months)	0.778
been threatened or injured with a weapon at school (12 months)	0.870
seen someone carrying a gun, knife, or other weapon at school (12 months)	0.720
been threatened with harm or injury at school (12 months)	0.885
Substance Use at School	
cigarettes on school property (30 days)	0.939
smokeless tobacco on school property (30 days)	0.930
electronic cigarettes, e-cigarette on school property (30 days)	0.864
at least one drink of alcohol on school property (30 days)	0.874
marijuana on school property (30 days)	0.910
any other drug, pill, or medicine to get "high"... on school property (30 days)	0.936

Table E.2: This table lists the "item" and associated weight for each of the four indices constructed to measure school climate and socioemotional well-being. The table structure and contents are a reproduction of the table on pages 38-40 of Mahecha and Hanson (2020).