NBER WORKING PAPER SERIES

BREAKING THE EARLY BELL:
LESSONS FROM THE FIRST STATEWIDE MANDATE ON SCHOOL START TIMES

Jialu (Gloria) Dou
Rania Gihleb
Osea Giuntella
Jakub Lonsky

Working Paper 35184
http://www.nber.org/papers/w35184

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
May 2026, Revised October 2026

We thank seminar participants at the University of Pittsburgh, Carnegie Mellon University, West Virginia University, Masaryk University, and the Slovak Academy of Sciences, as well as conference participants at the Society of Labor Economists Meetings, American Society of Health Economists Conference, European Society for Population Economics, Brucchi Luchino Workshop, and Eastern Economic Association for helpful comments and suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2026 by Jialu (Gloria) Dou, Rania Gihleb, Osea Giuntella, and Jakub Lonsky. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Breaking the Early Bell: Lessons from the First Statewide Mandate on School Start Times
Jialu (Gloria) Dou, Rania Gihleb, Osea Giuntella, and Jakub Lonsky
NBER Working Paper No. 35184
May 2026, Revised October 2026
JEL No. I0, I1, I20, J0

ABSTRACT

California's Senate Bill 328 is the first statewide mandate on school start times. Using difference-in-differences designs across four data sources, we find that the mandate increased the share of students sleeping at least 8 hours by 13%; mental health effects point toward improvement but are imprecise. High schools in counties with earlier pre-policy start times gained 1.7–1.8 percentage points in the share meeting state standards in math, English, and science, largest among Hispanic students. Suggestive evidence shows adolescents' sports participation fell and shifted later; parents arrived at work later and were more likely to be at work after 4 p.m.

Jialu (Gloria) Dou
University of Pittsburgh
Department of Economics
jid39@pitt.edu

Rania Gihleb
University of Pittsburgh
Department of Economics
and IZA
gihleb@pitt.edu

Osea Giuntella
University of Pittsburgh
Department of Economics
and IZA
and also NBER
osea.giuntella@gmail.com

Jakub Lonsky
University of Edinburgh
School of Economics
Jakub.Lonsky@ed.ac.uk

An online appendix is available at http://www.nber.org/data-appendix/w35184

1 Introduction

Sleep is now a focus of both popular books and academic research in public health and social sciences (Huffington, 2016; Walker, 2017; Attia, 2023; Haidt, 2024). A growing body of evidence shows that sleep influences cognitive function, productivity, and long-term health (Lim and Dinges, 2010; Cappuccio et al., 2010; Grandner, 2022; Gibson and Shrader, 2018; Giuntella and Mazzonna, 2019; Giuntella et al., 2026). For adolescents, sleep is especially important, supporting physical growth, emotional regulation, and cognitive development (Twenge et al., 2017; Lo et al., 2016; Giddens et al., 2022; Creswell et al., 2023). The American Academy of Sleep Medicine accordingly recommends that adolescents aged 13–18 get 8–10 hours of sleep per night (Paruthi et al., 2016). Yet most fall short of these guidelines, due to biological, social, and environmental factors (Pallesen et al., 2011; Carskadon et al., 1998; Carskadon, 2002; Haidt, 2024).

A key barrier to adequate sleep for adolescents is their delayed circadian phase, which shifts natural sleep-wake cycles later into the night (Crowley et al., 2007). This biological pattern conflicts with early school start times, generating chronic sleep deficits among teens (Hafner et al., 2017). Electronic device use compounds the problem by further delaying sleep onset (Twenge et al., 2019; Billari et al., 2018). The shift toward later sleep timing also differs by sex: from late adolescence through early adulthood, men's sleep on free days is centered later in the night than women's, and their chronotype continues to delay for about a year longer (Roenneberg et al., 2004; Fischer et al., 2017). Sleep deficits are also larger among Hispanic, Black, and low-income adolescents (Guglielmo et al., 2018; Giddens et al., 2022), potentially reflecting household constraints (e.g., parents' nonstandard work schedules, longer commutes, crowded housing, and norms about bedtimes) that make an early bell harder to accommodate. Differences in biological sleep timing and household circumstances suggest that the effects of later school start times may vary across students.

A substantial literature has examined how school start times affect adolescent sleep and academic performance (Heissel and Norris, 2018; Carrell et al., 2011; Groen and Pabilo-

nia, 2019; Hafner et al., 2017). With few exceptions, this evidence comes from localized quasi-experiments: district-level or institution-specific schedule changes whose effects depend on the surrounding schedules and institutions. Scaling such a policy from voluntary local adoption to a statewide mandate may change its effects for three reasons. First, voluntary adopters may differ from other districts in their expected benefits, implementation costs, or capacity to change schedules. A mandate extends the policy to districts that would not otherwise adopt it. Second, broader implementation may induce adjustments in transportation, athletics, after-school programs, and family schedules that are less feasible when adoption is isolated. Third, coordinated changes across districts may ease scheduling conflicts faced by individual adopters. Estimates from local adoptions therefore need not predict the effect of a statewide mandate. Recent work stresses that effects estimated in local experiments often change when a policy is taken to scale (List, 2022, 2024). Effects at scale could be larger or smaller than those estimated from local adoption. California's statewide class-size reduction, for instance, produced smaller gains than those observed in the Tennessee STAR experiment, as scaling changed who was treated and the quality of the teachers hired to deliver it (Jepsen and Rivkin, 2009; Adusumilli et al., 2024).

California's Senate Bill 328 (SB 328) is the first U.S. statewide mandate requiring middle schools to start no earlier than 8:00 a.m. and high schools no earlier than 8:30 a.m. According to our own survey, the mandate was binding for roughly 80% of the non-rural California districts that responded[1] and required all affected districts to change at the same time. This broad implementation allows us to examine both the effects of the mandate and whether transportation, extracurricular activities, and parents' work routines adjusted to later school start times.

We draw on four data sources. The Youth Risk Behavior Survey (YRBS) measures sleep duration and mental health outcomes for high school students. The American Time Use Survey (ATUS) provides sleep timing and duration from the time diaries of respondents

[1]Based on our survey of pre-policy start times; see Section 2.1 and 3.3 for details.

aged 15–18, together with their sports participation and its timing, paid work, and phone use, and the work and commuting schedules of their parents. Children under 13 do not complete diaries, but adult respondents report when the first child under 13 in the household woke up and when the last one went to bed, which we use as measures of children's wake and bedtimes. For academic outcomes, we use a within-state design that accounts for common California-wide shocks. We compare schools in counties where the mandate required a larger shift in schedules with schools in counties that already started late or were exempt,[2] proxying pre-policy start times with teachers' commute arrival times in the American Community Survey (ACS) and measuring outcomes with school-level records from the California Department of Education (CDE).

We first show that teachers and students arrive at school later after the mandate takes effect, consistent with a shift to later school start times. Using the YRBS, we find that SB 328 increased the share of students sleeping at least 8 hours per night, the recommended minimum for this age group, by 13%. In the ATUS, among teenagers aged 15–18, the share awake by 6 a.m. declined by 61% and the share sleeping at least 8 hours rose by 14%. For children under 13, later wake times contributed to increases in sleep duration, with the share awake by 6 a.m. declining by 63% and the share sleeping at least 8 hours increasing by 12%. Estimates for mental health outcomes in the full sample are imprecise. Turning to academic outcomes, because the survey response rate was low, we use teachers' arrival times at work to proxy for pre-policy school start times. Using administrative data from the California Department of Education, we find that in schools in counties where pre-policy start times were earlier the share of students meeting state standards rose by about 1.7–1.8 percentage points in math, English, and science at the high school level, with gains accumulating over 2023–2025, comparable gains in math and English and a smaller gain in science in grade 8, and suggestive declines in dropout and suspension rates. Finally, we examine how extracurricular activities and family schedules adjusted to the mandate. Time-

[2]SB 328 explicitly exempted rural school districts from its provisions.

use diaries suggest that adolescents' weekday sports participation declined and shifted later in the day, while parents of affected children woke up and arrived at work slightly later, with no change in commuting time.

We also examine differences by gender and ethnicity, motivated by the discussion above, as exploratory analyses. Using YRBS data, we find suggestive evidence that boys experienced larger improvements in sleep duration than girls, although the differences are not precisely estimated. For mental health, effects on the likelihood of reporting difficulty concentrating are similar across genders, but boys show significant reductions in the likelihood of reporting sadness/hopelessness and suicidal intentions. Hispanic students show the largest reduction in difficulty concentrating, but not the largest sleep gains. Within California, Hispanic students also experienced the largest improvements in test performance, and economically disadvantaged students gained significantly, while gains among White students were smaller and less precisely estimated. These patterns are consistent with the view that a later start time relaxes a more binding constraint for groups with larger pre-policy sleep deficits, but we view them as suggestive: the heterogeneity was not pre-specified, and the subgroup contrasts are not precisely estimated.

In the YRBS, event-study estimates show no evidence of differential pre-trends, lending support to the parallel trends assumption underlying our identification strategy. In the smaller ATUS samples, a few early bins differ from zero, but they do not exhibit a systematic trend toward the post-policy estimates. Permutation tests that randomly reassign the treatment state and adoption year place the estimated California effects toward the tails of the placebo distributions. As further checks, we find no evidence that SB 328 affected school-level enrollment, racial or gender composition, or the share of economically disadvantaged students. We also find no effects on outcomes unlikely to respond to the policy, such as vision and hearing difficulties. These results provide reassurance against compositional changes and broader health trends as explanations for our findings.

We contribute to three related literatures. First, on school start times and academic

achievement, existing quasi-experimental evidence comes almost entirely from localized variation (voluntary district-level adoption, institution-specific schedule changes, or naturally occurring geographic variation (Hinrichs, 2011; Carrell et al., 2011; Edwards, 2012; Heissel and Norris, 2018; Bastian and Fuller, 2023; James et al., 2023)). SB 328 is a statewide mandate and, because adoption was simultaneous, we can trace how households and schedules outside the school adjusted. We provide the first evidence from a statewide mandate and document its effects on sleep, mental health, and achievement in math, English, and science. Second, on sleep, human capital, and productivity, prior work has documented the importance of sleep for cognitive performance and earnings (Giuntella et al., 2017; Gibson and Shrader, 2018; Jagnani, 2021; Bessone et al., 2021; Giuntella et al., 2026). We extend this literature by estimating the effects of a policy that increases adolescent sleep at statewide scale and by documenting the household adjustments that accompany it. Third, we contribute to research on disparities in adolescent sleep and well-being. Prior work has documented greater sleep deficits among boys and Hispanic youth (Giddens et al., 2022; Scott et al., 2021). We find suggestive evidence of larger improvements in sadness and suicidal ideation among boys and in difficulty concentrating among Hispanic students, although these patterns are not uniform and subgroup differences are imprecisely estimated.

School start times remain an active policy issue in other states. Florida enacted a similar statewide mandate in 2023 (House Bill 733), but legislation enacted in 2025 (CS/CS/SB 296) allowed districts to comply by submitting a report to the Department of Education rather than changing their schedules. Florida's rollback highlights the implementation challenges surrounding statewide start-time mandates and the importance of understanding how schools and households adjust.[3]

The remainder of the paper proceeds as follows. Section 2 reviews SB 328 and the

[3]Legislation proposing statewide minimum school start times was introduced in New York during the 2025–2026 session (S. 2631A, requiring public secondary schools to start no earlier than 8:30 a.m., with state aid tied to compliance) and Missouri in 2026 (S.B. 1710, proposing an 8:30 a.m. minimum for public high schools, with hardship waivers). The Florida Senate's analysis of the state's 2025 rollback documented district concerns about bus-driver shortages, transportation costs, child care, and after-school activities.

prior evidence on school start times, including a PRISMA-style systematic review of quasi-experimental studies. Section 3 describes the four data sources and the cross-state and within-state empirical specifications. Section 4 presents the main results on sleep, mental health, and academic outcomes, together with permutation tests and placebo checks, and shows how adolescents' activities and parents' schedules adjusted. Section 5 documents heterogeneity by gender and ethnicity. Section 6 concludes.

2 Background

2.1 SB 328

Senate Bill 328 (SB 328), signed into law by Governor Gavin Newsom on October 13, 2019 and codified at California Education Code §46148, requires California middle schools to start no earlier than 8:00 a.m. and high schools no earlier than 8:30 a.m. Districts were required to comply by July 1, 2022, or if applicable, upon the expiration of any collective bargaining agreement in effect as of January 1, 2020, whichever was later. The legislation includes an exemption for rural school districts, which face distinct logistical challenges including longer transportation times and fewer resources for schedule adjustments. We exploit this exemption in the within-state analysis below.

To assess how binding the mandate was at the outset, we conducted a survey of California school districts on their pre-policy start times. Among the 114 non-rural respondents reporting whether they shifted schedules in response to SB 328, approximately 80% reported a shift, indicating that the mandate required schedule changes for the large majority of high school districts.[4] SB 328 makes California the first state in the U.S. to enact a statewide

[4]Of the 674 relevant California school districts, 421 are high school-relevant (76 high school districts and 345 unified districts that span both elementary and secondary grades); the remaining 253 are elementary districts which operate no schools subject to SB 328's high school threshold, but they do include middle schools or middle school-level grades (hence still subject to SB 328's middle school threshold). Of the 114 non-rural respondents reporting whether they shifted schedules, 91 (approximately 80%) reported shifting. Consistently, among 85 non-rural respondents with usable pre-policy start times, 78 (92%) had start times before the 8:30 a.m. threshold. Rural districts are statutorily exempt and excluded from both calculations.

mandate on school start times.

2.2 Previous evidence on school start times

Prior research on school start times and student outcomes ranges from cross-sectional correlational analyses to randomized experiments. To organize this evidence we conducted a PRISMA-style systematic search of the quasi-experimental literature (42 studies, 33 of which report parametric standard errors; 18 provide effect sizes in a common metric and enter the pooled forest plots); we summarize the main findings here, the Appendix reports the search protocol, the list of studies, and the distribution of estimated effects (Figure A.1, Table A.23, and the forest plots in Figures A.2 and A.10; Table A.1 summarizes representative study-level estimates).

The most consistent finding in the literature is that later start times increase sleep duration.[5] Two meta-analyses that cover 20 and 28 studies confirm this (Bowers and Moyer, 2017; Yip et al., 2022), and Yip et al. (2022) find better sleep and development outcomes for school start times between 8:30 and 8:59 a.m. than for 8:00–8:29 a.m. Quasi-experimental estimates of sleep gains in non-boarding settings range from 19 to 43 minutes (Owens et al., 2017; Widome et al., 2020; Jung and Fenelon, 2024), though the gains from a South Korean regional mandate largely faded within a year (Jung and Fenelon, 2024).

Evidence on standardized tests is mixed. Hinrichs (2011) uses policy changes in the Twin Cities metropolitan area and finds no effect on academic achievement. Results for non-test outcomes are more consistent: Lenard et al. (2020) document increases in absences, tardiness, and dropout rate when start times are moved earlier, and James et al. (2023) find improved attendance and reduced tardiness, with GPA gains of up to 0.17 points emerging in the second year of follow-up.

[5] Pooling the quasi-experimental sleep estimates in Appendix Figure A.2 with inverse-variance random-effects (REML) weights gives approximately 39 additional minutes of sleep per one-hour delay in start time (95% CI: 26–51; $k = 10$), with substantial heterogeneity across settings ($I^2 = 90\%$): non-boarding studies cluster around 20–45 minutes while boarding-school settings yield 70–90. The corresponding pooled test-score effect is 0.03 SD per hour (95% CI: 0.01–0.05; $k = 18$ study–outcome estimates from 9 studies; $I^2 = 73\%$). Study-level estimates are in Appendix Table A.23.

Several quasi-experimental studies do find positive test score effects. Heissel and Norris (2018) find +0.06–0.08 SD per hour using variation in sunrise times across a time-zone boundary, with math effects concentrated among adolescents; Edwards (2012) finds gains of 1.5–3 percentile points per hour using variation driven by bus scheduling in North Carolina middle schools; Carrell et al. (2011) estimate that a 50-minute earlier start lowers grades by 0.13 SD (0.15 SD per hour), using random assignment to course schedules at the U.S. Air Force Academy. Groen and Pabilonia (2019), using time-diary data from the Child Development Supplement to the Panel Study of Income Dynamics (PSID-CDS), document suggestive reading gains of 0.16–0.28 SD for female students with no detectable effect for males, though their identification relies on OLS with selection bounds rather than a quasi-experimental design.[6]

Prior studies also report heterogeneity along the dimensions we examine. Bastian and Fuller (2023), using six years of statewide administrative data from North Carolina with student fixed effects, find that later start times reduce absences and suspensions among economically disadvantaged students, consistent with the larger academic gains among Hispanic students and the significant gains among economically disadvantaged students that we document. Evidence on gender is mixed: Groen and Pabilonia (2019) find gains only for girls, whereas Lusher and Yasenov (2016, 2018), exploiting a double-shift setting in which schools alternate monthly between morning and afternoon blocks, find that boys benefit from later starts while girls are unaffected, which explains up to 16% of the gender performance gap in morning classes. They attribute this to gender differences in sleep cycles, the mechanism we discussed above and one consistent with the somewhat larger, though imprecisely estimated, sleep gains we find for boys in the YRBS.

[6]Two studies provide evidence on the effect of sleep itself. Jagnani (2021), exploiting variation in sunset times across Indian districts that shifts children's sleep, finds that less sleep reduces human capital formation. Giuntella et al. (2026), in a randomized field experiment with U.S. university students, show that incentives increasing weeknight sleep by 19 minutes raise semester grades by 0.10–0.11 SD, with effects operating through cognitive and regularity channels rather than attendance. Because the latter varies sleep without changing schedules or light exposure, it supports sleep as a mechanism distinct from other features of later start times.

3 Data and empirical specification

3.1 Data

Our analysis draws on four data sources: the Youth Risk Behavior Surveillance System (YRBS), the American Time Use Survey (ATUS), administrative school-level records from the California Department of Education (CDE), and the American Community Survey (ACS). The YRBS and ATUS data support the cross-state difference-in-differences analysis of sleep and mental health. The CDE and ACS data support the within-state analysis of academic outcomes, with CDE providing school-level outcomes and ACS providing teachers' commute arrival times used as a proxy for pre-policy school start times. We also examine state-level SAT score reports from the College Board as a descriptive complement for high school outcomes as discussed below.

3.1.1 Youth Risk Behavior Surveillance System (YRBS)

The Youth Risk Behavior Surveillance System (YRBS), conducted biennially by the Centers for Disease Control and Prevention (CDC), collects data on a nationally representative sample of high school students in grades 9–12 (typically ages 14–18). The survey covers health behaviors including sleep duration, mental health, physical activity, and substance use. We use the national YRBS samples, which identify the respondent's state of residence; over our sample period they cover 47 states and Washington D.C. (South Dakota, Vermont, and Wyoming do not appear), with between 25 and 31 states represented in any given wave. We focus on the years 2007–2023. Summary statistics are reported in Table A.2.

3.1.2 American Time Use Survey

The American Time Use Survey (ATUS), conducted annually by the U.S. Bureau of Labor Statistics, provides self-reported data on daily activities and time allocation for a nationally representative sample of U.S. residents. We focus on weekdays and two analysis samples:

adolescents aged 15–18 and children aged 0–12, for whom we observe wake-up times, bedtimes, and sleep duration.[7] We use data from 2003 to 2024, excluding 2020 due to the COVID-19 pandemic. For children under 13, the ATUS instructs the respondent to report time use for either the eldest or youngest child in the household. To ensure that wake-up and bedtime reports can be attributed to the same child, we restrict the under 13 analysis to single-child households and report results including all households with at least one child under 13 as a robustness check in the Appendix. Summary statistics for both samples are reported in Table A.3. The share of adolescents sleeping at least 8 hours is much higher in the ATUS (70%) than in the YRBS (27%): the diary records the actual sleep episodes of one specific day, including weekday diaries from school holidays, whereas the YRBS asks students to classify their sleep on a usual school night into hour categories, and self-reported sleep is known to differ systematically from diary- and actigraphy-based measures (Lauderdale et al., 2008). Our estimates compare changes within each survey, so the level difference does not affect them. For the analysis of adjustment margins in Section 4.5, we construct from the adolescents' diaries indicators for any sports or exercise on the diary day and for its timing, for any paid work, and for time spent on the phone, messaging, and social media, and we use the diaries of parents aged 25–64 whose eldest own child in the household is at least 11 and whose youngest is at most 18 for their wake-up and bedtimes and the hours at which they are at work. The ACS, described below, provides the time at which parents arrive at work and their commute duration.

3.1.3 California Department of Education (CDE)

For academic outcomes, we draw on administrative data from the California Department of Education (CDE), which publishes annual school-level records for all public middle and high schools in California. These data provide a rich set of outcomes spanning academic achievement and disciplinary measures. Although SB 328 applies to middle and high schools,

[7] As a robustness check, we further exclude observations from the summer and other holidays.

younger students may be affected indirectly, through district-wide bell-schedule changes that move elementary start times together with secondary ones, especially in K–8 or 6–12 school configurations, and through shared household routines (Edwards, 2012; Groen and Pabilonia, 2019; Foster and Kalil, 2007).[8]

For academic outcomes, we use school-level data on standardized test scores from the California Assessment of Student Performance and Progress (CAASPP). Specifically, we examine the log of the mean scale score and the share of students meeting or exceeding the state standard in mathematics, English language arts/literacy (ELA), and science. Math and English outcomes are measured for Grade 11 at the high school level and Grade 8 at the middle school level. California Science Test (CAST) is administered by middle schools in Grade 8, while high schools can choose to administer it in Grade 10, 11, or 12. For dropout rates and disciplinary outcomes, we use school-level data on four-year cohort dropout rate, suspension rate, and the number of suspensions per 1,000 pupils. The dropout rate is available only for high schools, while suspensions sample covers both school levels combined.

The CDE data span the academic years 2011–12 through 2024–25, but data availability varies by outcome.[9] For assessment outcomes, we exclude the school years 2019–20 and 2020-21 due to the disruption to standardized testing caused by the COVID-19 pandemic.[10] The resulting panel covers approximately 1,700–2,500 schools for assessment outcomes and roughly 6,200 schools for school-wide suspension outcomes.

As a descriptive complement, we also use state-level SAT scores published by the College Board.[11] The SAT was redesigned in 2016, and we use scores for the 2018 and 2021–2023 graduating cohorts. For students with multiple test attempts, the data report the most

[8]Our district survey did not ask about elementary start times. The ACS arrival times of elementary and middle school teachers, who share an occupation code, do shift later after 2022 (Table A.4), consistent with some elementary schedules moving as well. We return to the interpretation of the under-13 estimates in Section 4.

[9]Suspensions data is available for academic years 2011–12 through 2024–25; math and English assessment data is available for academic years 2014–15 through 2024–25; dropout data is available for academic years 2016–17 through 2024–25; science assessment data is available for academic years 2018–19 through 2024–25.

[10]Standardized testing in California generally takes place in the spring, with most school districts implementing it between March and May (source: California Department of Education, 2025).

[11]https://reports.collegeboard.org/sat-suite-program-results

recent section scores. SAT data are available only at the state level, precluding heterogeneity analysis. We interpret the SAT results as descriptive rather than causal, since the test-optional admissions policies adopted during COVID-19 changed the composition of SAT takers in ways that confound simple pre-post comparisons; we discuss this further below.

3.1.4 American Community Survey (ACS)

The American Community Survey (ACS), conducted by the U.S. Census Bureau, records for each employed respondent the time they leave home for work, the duration of their commute, and their occupation, from which we infer the time they arrive at work. We use the ACS in three ways. First, the arrival times of K–12 teachers (IPUMS occupation codes 2310 for elementary and middle school teachers and 2320 for high school teachers) provide the proxy for pre-policy school start times that defines treatment in the within-state analysis; we use the 2015–2019 five-year sample to construct county-level pre-policy arrival-time distributions and the 2023 one-year sample to document the post-policy shift (Section 3.3). Second, for the analysis of household adjustment in Section 4.5, we use the 2015–2024 samples of parents aged 25–64 with a child aged 11–18 in the household to measure the time they arrive at work and the length of their commute, excluding 2020 (the experimental pandemic sample) and 2022 (for which the interview month is unavailable). Third, the ACS reports self-assessed vision and hearing difficulties for adolescents, which we use as placebo outcomes that school schedules should not affect. ACS person weights are used throughout.

3.2 Empirical specification

To identify the causal effects of SB 328 on sleep and mental health, we estimate a difference-in-differences (DID) model. We compare changes in sleep and mental health outcomes for students in California after the 2022 implementation of SB 328 to changes in other states over the same period.[12] We complement this cross-state DID with a within-state design

[12]The timing of treatment varies slightly across datasets, reflecting differences in data collection schedules. For the YRBS, the treatment period corresponds to the 2023 wave, the first survey conducted after the

that exploits variation in pre-policy compliance across California districts, described in Section 3.3.

Formally, we estimate the following DID model:

$$y_{ist} = \alpha + \beta\,\text{SB328}_{st} + \gamma X_{ist} + \tau_t + \eta_s + \epsilon_{ist} \tag{1}$$

where y_{ist} is a metric of sleep or mental health outcomes; β captures the effect of the SB 328 adoption in California, X_{ist} is a battery of individual controls (e.g. sex, age), and τ_t and η_s are year and state fixed effects, respectively. Standard errors are clustered at the state level. Identification rests on the parallel trends assumption: in the absence of SB 328, outcomes in California and other states would have followed parallel trends. We assess this assumption through event-study estimates and through permutation tests that reassign the treatment to other states. Because YRBS and ATUS are nationally representative but not state-representative, we weight observations using population data from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) Program, assigning weights to each respondent based on state of residence, survey year, age, sex, and race.[13]

To examine dynamic treatment effects and assess the plausibility of the parallel trends assumption more directly, we also estimate an event-study specification of the form:

$$y_{ist} = \alpha + \sum_{k \neq -1} \delta_k \cdot \mathbf{1}\{s = \text{CA}\} \cdot \mathbf{1}\{t - 2022 = k\} + \gamma X_{ist} + \tau_t + \eta_s + \epsilon_{ist} \tag{2}$$

where $\mathbf{1}\{s = \text{CA}\}$ is an indicator for California, $\mathbf{1}\{t - 2022 = k\}$ is an indicator for year k relative to the 2022 SB 328 implementation, and δ_k captures California's deviation from the comparison-state trend in event-time k. We omit $k = -1$ as the reference period. The

policy's implementation. The ATUS collects data continuously throughout the calendar year and includes month-level identifiers; we define the treatment period as beginning with the 2022–2023 academic year, when SB 328 took effect. For the SAT score data, the treatment period begins with the 2023 graduating cohort. For the CDE data, the treatment period begins with the 2022–2023 academic year, the first full school year under SB 328; the 2019–2020 and 2020–2021 school years are excluded due to COVID-19 disruptions to standardized testing.

[13]For the ATUS sample of children under 13, individual demographic characteristics are not consistently observed; we instead weight by the state-year population of individuals aged 0–12.

vector X_{ist} includes individual-level controls and state-level covariates capturing demographic composition and economic conditions.[14] Because of limited annual sample sizes in the ATUS and inconsistent state-year coverage in the YRBS, we aggregate observations into multi-year bins for the event-study analysis of these data, which improves power while preserving pre- and post-treatment dynamics.

3.3 Within-state analysis

For academic outcomes, and to address potential concerns about state-level confounders more generally, we use a within-state identification strategy that exploits variation in pre-existing compliance with SB 328 across California school districts. SB 328 only required changes from districts with pre-policy start times earlier than the new thresholds (8:00 a.m. for middle schools and 8:30 a.m. for high schools). Districts with earlier start times were therefore directly affected by the mandate, while already-compliant districts together with rural districts (exempt from the SB 328 provisions) serve as a within-California control group. In practice, we compare changes in school outcomes before and after 2022 in counties where the mandate required a larger shift (i.e., where teachers used to arrive earlier) with changes in counties whose schools already started late or were exempt. This design absorbs common statewide shocks (e.g., concurrent state education reforms, demographic trends, or statewide changes in schedules and transportation) that could otherwise confound the cross-state estimates, although it does not rule out differential shocks across counties.

Direct district-level data on pre-policy start times are not systematically available. The district survey described in Section 2 (approximately 80% of responding non-rural high school districts reported shifting schedules in response to SB 328) confirms that the mandate was binding for the large majority of high school districts subject to the law. However, although

[14]For YRBS, state-level covariates include population, population density, share of households with children, marital status and education variables, unemployment rate, median household income, and child poverty rate. For ATUS respondents aged 15–18, state-level covariates include the share of households with children, married share, school dropout rate (aged 16-19), unemployment rate, and child poverty rate. For children under 13, the vector adds total state population, share of divorced individuals, share of individuals with at least a high school degree, and median household income.

we contacted all 674 relevant California school districts, the response rate among the 421 high school-relevant districts (76 high school-only and 345 unified) was approximately 27% and likely selective, making the survey unsuitable as a systematic district-level measure of treatment intensity. We do, however, use the responding districts as a validation sample: estimates that define treatment from the surveyed start times are reported in Appendix Table A.16 and are in line with those based on the teacher-arrival proxy.

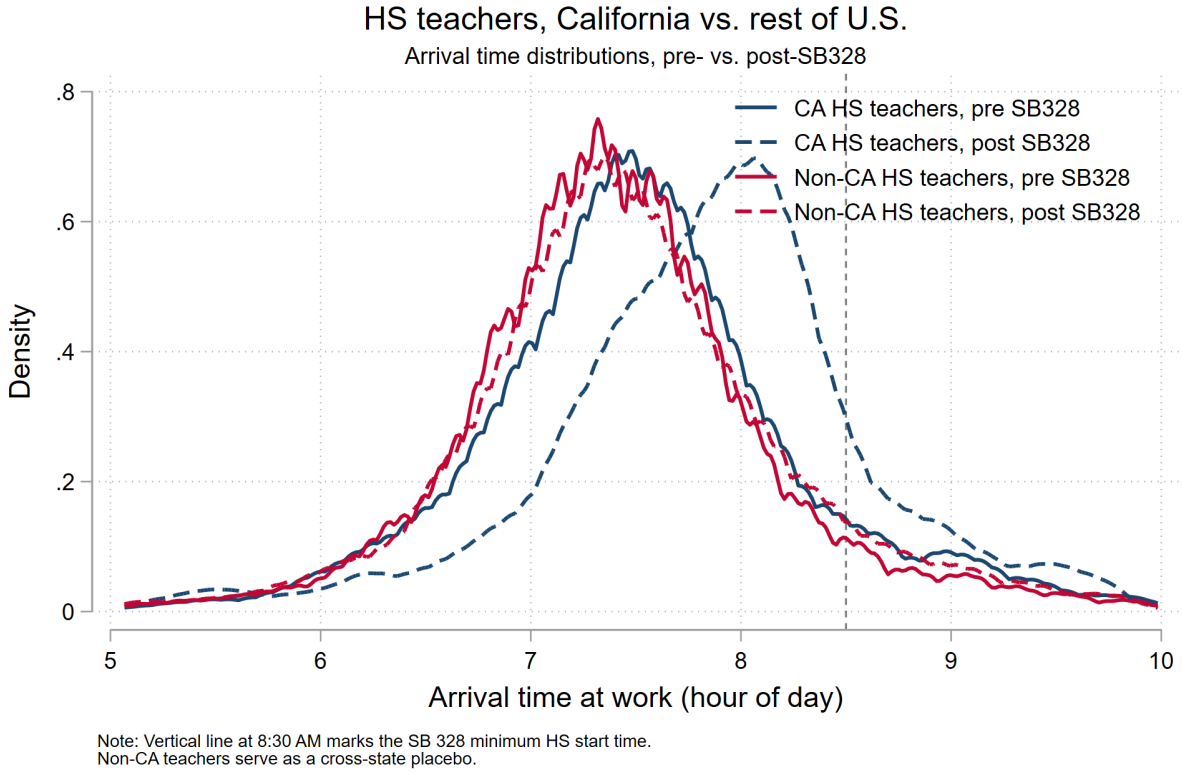
We therefore proxy for pre-policy school start times using teachers' commute arrival times from the American Community Survey. In districts with earlier school start times, teachers arrive earlier, so reported commute arrival times should reflect the underlying school schedule. We validate this proxy with two comparisons. First, the arrival-time distribution for California high school teachers shifted markedly toward later hours following SB 328's implementation in 2022, while distributions for teachers in other states were stable over the same period (Figure 1), indicating the shift was California-specific. Second, within California, commute schedules for workers in other occupations remained stable over the same period (Appendix Figure A.5), making broader labor-market shifts an unlikely alternative explanation. Together, these patterns support the use of pre-policy teacher arrival times as a proxy for pre-policy school start times.

Appendix Figures A.6 and A.7 provide analogous evidence for middle school teachers. Both show a post-implementation shift toward later arrivals, smaller in magnitude than for high school teachers. This is consistent with the mandate being less binding for middle schools: pre-policy middle school teacher arrivals were already concentrated near the 8:00 a.m. threshold.

Table A.4 formalizes this pattern, with treatment defined at the county level. Following SB 328, high school teachers' arrival times in treated California counties shifted approximately 19 minutes later ($p < 0.05$), and elementary and middle school teachers' arrival times by approximately 12 minutes ($p < 0.01$).

We further validate the proxy using self-reported school arrival times of 15–18 year olds

Figure 1: Validating the Teacher Arrival-Time Proxy: ACS Arrival Times, California High School Teachers vs. High School Teachers in Other States, Pre- vs. Post-SB 328



Notes: The figure presents kernel density estimates of self-reported arrival times at work among high school teachers, using data from the American Community Survey (ACS) for the years 2015–2019 and 2023. The pre-SB 328 period uses the 5-year ACS sample (2015–2019); the post-SB 328 period employs the 1-year 2023 ACS sample, the first full wave after the policy took effect on July 1, 2022. High school teachers are identified using IPUMS 2018 occupation code 2320. The sample is restricted to respondents with reported arrival times between 5:00 a.m. and 10:00 a.m. Observations are weighted by ACS person weights (`perwt`). The vertical dashed line at 8:30 a.m. marks the minimum high school start time mandated by SB 328. The figure shows a marked rightward shift in the distribution of California high school teacher arrival times following the policy, with the post-period mass concentrating immediately below the 8:30 a.m. threshold, while the distribution of arrival times for high school teachers in other states is essentially unchanged between the pre- and post-periods. This pattern supports the use of pre-policy teacher arrival times as a proxy for pre-policy school start times in the within-state analysis.

in the ATUS, restricted to metropolitan areas (Table 1).[15] We measure the time each student arrives at school as the first minute at school recorded in the diary. Across the estimation sample, 85% of high-school students in metropolitan areas who attended class were at school by 8:30 a.m., the earliest start time the law permits for high schools; after SB 328, that share fell by 38 percentage points in California relative to students in other states ($p < 0.01$). The share at school by 8:00 a.m. fell by 20 percentage points (sample mean 67%), while the share at school by 7:30 a.m. was unchanged, consistent with early-morning athletics and "zero period" classes continuing as before. The mandate thus moved arrivals from the hour before 8:30 a.m. to 8:30 a.m. or later.

We define a county's treatment intensity based on its pre-policy teacher arrival-time distribution, constructed from the ACS 2015–2019 5-year sample. For high school outcomes, counties in the bottom three quartiles of pre-policy mean teacher arrival time are classified as treated and the top quartile plus rural counties as control; for middle school outcomes, counties in the bottom two quartiles are classified as treated and the top two plus rural as control.[16] The asymmetry reflects the weaker bite of the mandate at the middle school level (Appendix Figure A.6).[17] The cut points are guided by the pre-policy arrival-time distributions in Figure 1 and Appendix Figures A.6–A.7, which show that teachers' arrival times differ sharply across school levels; results are robust to alternative cuts, including a top-versus-bottom-quartile contrast and an above-versus-below-mean measure (Appendix Tables A.14 and A.15; the sharper top-versus-bottom contrast is less precisely estimated).

[15] We restrict to respondents in metropolitan areas (IPUMS metro codes 1–3) to exclude rural areas, which approximate the rural districts exempted from SB 328 under the statute. Restricting further to central cities (metro code 1, Panel B) yields larger shifts: the share at school by 8:30 a.m. falls by 66 percentage points, consistent with the mandate binding most strongly in urban districts with the earliest pre-policy start times.

[16] Throughout the within-state analysis, "rural" counties are those that cannot be identified in the ACS public-use microdata because of their small population (the ACS identifies counties only when they contain at least one public-use microdata area, in practice counties with roughly 100,000 residents or more). These small, predominantly rural counties have no teacher arrival-time measure and are assigned to the control group. This is an approximation to the statute's exemption, which applies to rural *districts* as designated by the state rather than to counties: we have not matched the two classifications, and exempt districts located in identified counties are classified by their county's arrival times like any other school.

[17] For schoolwide suspension outcomes in a combined high school/middle school sample, we apply the high school treatment definition utilizing arrival times of all non-higher ed. teachers.

Table 1: SB 328 and the Time Students Arrive at School, ATUS (Ages 15–18)

	(1)	(2)	(3)
	Share of students already at school by		
	7:30 a.m.	8:00 a.m.	8:30 a.m.
Panel A: All metropolitan areas (metro ≤ 3)			
Policy (SB 328)	−0.048	−0.200***	−0.383***
	(0.066)	(0.058)	(0.023)
Observations	2,600	2,600	2,600
R-squared	0.100	0.107	0.116
Mean of dep. var.	0.355	0.665	0.849
Panel B: Metropolitan, central cities (metro = 1)			
Policy (SB 328)	−0.070	−0.440***	−0.662***
	(0.063)	(0.078)	(0.059)
Observations	708	708	708
R-squared	0.217	0.215	0.289
Mean of dep. var.	0.330	0.618	0.817

Notes: Data come from the 2003–2024 American Time Use Survey (ATUS). 2020 is omitted due to incomplete data collection during the COVID-19 pandemic, and observations are restricted to weekday diaries. Each column within a panel reports estimates from a separate difference-in-differences regression of an indicator for being at school by the stated time on a post-SB 328 × California interaction (post = September 2022 onward), with state, year, sex, age, month, SEER race and Hispanic-origin fixed effects and the state-level controls of Table 3; observations are weighted by SEER state-year-age-sex-race population weights. The time a student arrives at school is the first minute at school recorded in the diary, obtained from the minutes spent at school between 7:00 and 8:00 a.m. and between 8:00 and 8:30 a.m.; a student who reports no time at school before 8:30 a.m. arrived at 8:30 a.m. or later. The three indicators are cumulative: column (3) is one for every student at school by 8:30 a.m., the earliest start time SB 328 permits for high schools. The sample is restricted to ATUS respondents aged 15–18, enrolled full time in high school, who report attending class on the diary day. Panel A includes all metropolitan areas (IPUMS metro codes 1–3), excluding nonmetropolitan areas that most closely correspond to the rural districts exempted from SB 328; Panel B restricts to central cities of metropolitan areas (metro code 1). Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

We aggregate teacher arrival times to the county of work—the finest geography at which ACS cell sizes are large enough to construct stable pre-policy means—and assign every school the treatment status of its county. We then estimate two-way fixed effects models of the form:

$$y_{st} = \alpha + \beta \cdot \text{Policy}_{ct} + \gamma X_{st} + \tau_t + \eta_s + \epsilon_{st} \tag{3}$$

where y_{st} is an outcome of interest in school s and year t; Policy_{ct} is an indicator equal to one for non-compliant counties in post-implementation years; X_{st} is a vector of time-varying school-level controls; and τ_t and η_s are year and school fixed effects respectively. Standard errors are clustered at the county level.

We additionally estimate event-study specifications analogous to equation (2) to trace the dynamic evolution of treatment effects and assess pre-trend balance. Outcomes are the school-level CDE records described in Section 3.

To assess the validity of the within-state design, we conduct two placebo tests. First, in Table A.6 (cols. 3–4), we use the ACS-based treatment intensity measures to predict outcomes that should be unaffected by school schedules (self-reported serious vision and hearing difficulties). We find no effect, supporting the interpretation that our identifying variation captures schedule changes rather than broader correlated trends. Second, in Table A.7, we use the CDE data to test whether SB 328 affected school-level composition (total enrollment, racial composition, gender composition, share of economically disadvantaged students). The estimates are small relative to the means and precise enough to exclude meaningful compositional shifts: total enrollment changes by -14 students on a mean of 803 (95% CI -42 to $+14$), the Hispanic share by -0.4 percentage points on a mean of 56 (CI -1.8 to $+1.1$), the White share by $+0.3$ points (CI -0.7 to $+1.3$), the female share by $+0.1$ points (CI -0.1 to $+0.2$), and the socioeconomically disadvantaged share by $+0.7$ points on a mean of 61 (CI -1.6 to $+3.0$). Compositional change is therefore an unlikely driver of the academic results.

4 Results

4.1 Effects of SB 328 on sleep

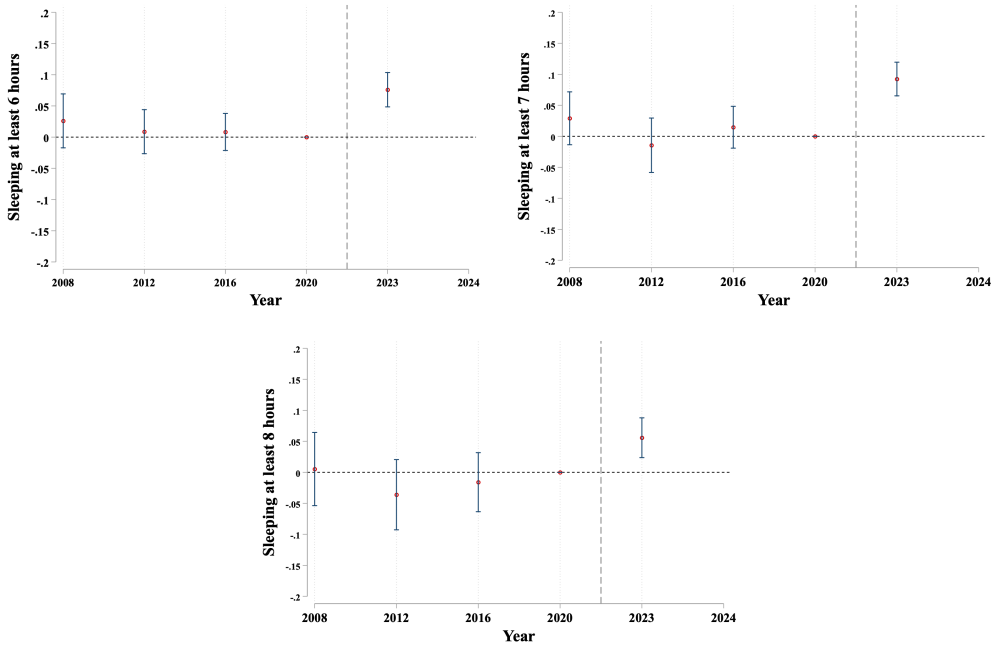
Table 2: Policy and Sleep Outcomes, YRBS (2007–2023)

	(1) Sleep $\geq$ 6 hours	(2) Sleep $\geq$ 7 hours	(3) Sleep $\geq$ 8 hours
CA SB 328	0.067***	0.070***	0.036**
	(0.017)	(0.017)	(0.017)
Observations	118,035	118,035	118,035
R-squared	0.024	0.037	0.037
Mean of dep. var.	0.792	0.558	0.271
Std. dev. of dep. var.	0.406	0.497	0.445

Notes: Data comes from the 2007-2023 national samples of the Youth Risk Behavior Survey (YRBS). Observations are weighted by population weights obtained from National Cancer Institute's SEER Program and assigned according to individual's state of residence, survey year, age, sex, and race. All specifications control for individual-level covariates (grade, race/ethnicity, sex) and state-level characteristics (population, population density, % households with children, marital status & education-related variables, unemployment rate, median household income, and child poverty rate). All regressions further include state and year fixed effects. Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.

This section reports effects on sleep at the individual level, using YRBS and ATUS data. Table 2 reports the effects of SB 328 on sleep duration in the YRBS. The policy increased the share of students sleeping 6 hours or more by 6.7 percentage points (an 8% increase relative to the mean; $p < 0.01$), the share sleeping 7 hours or more by 7 percentage points (13%; $p < 0.01$), and the share sleeping 8 hours or more by 3.6 percentage points (13%; $p < 0.05$). Figure 2 shows the corresponding event-study, comparing sleep among students in California and other states before and after SB 328. We find no evidence of differential pre-trends, supporting the parallel trends assumption, and the event-study estimates corroborate the difference-in-differences results in Table 2.

Figure 2: California SB 328 and Sleep, YRBS



Notes: The figure illustrates the differences in sleep behavior among students in California and other states before and after the adoption of SB 328. Data comes from the 2007-2023 national samples of the Youth Risk Behavior Survey (YRBS). Observations are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to individual's state of residence, survey year, age, sex, and race. Survey years were pooled as follows: (1) 2008 combines 2007 & 2009; (2) 2012 combines 2011 and 2013; (3) 2016 combines 2015 and 2017; (4) 2020 combines 2019 and 2021; (5) 2023 is the 2023 wave alone. Standard errors are clustered at the state level. Coefficient estimates are provided together with the 95% confidence intervals.

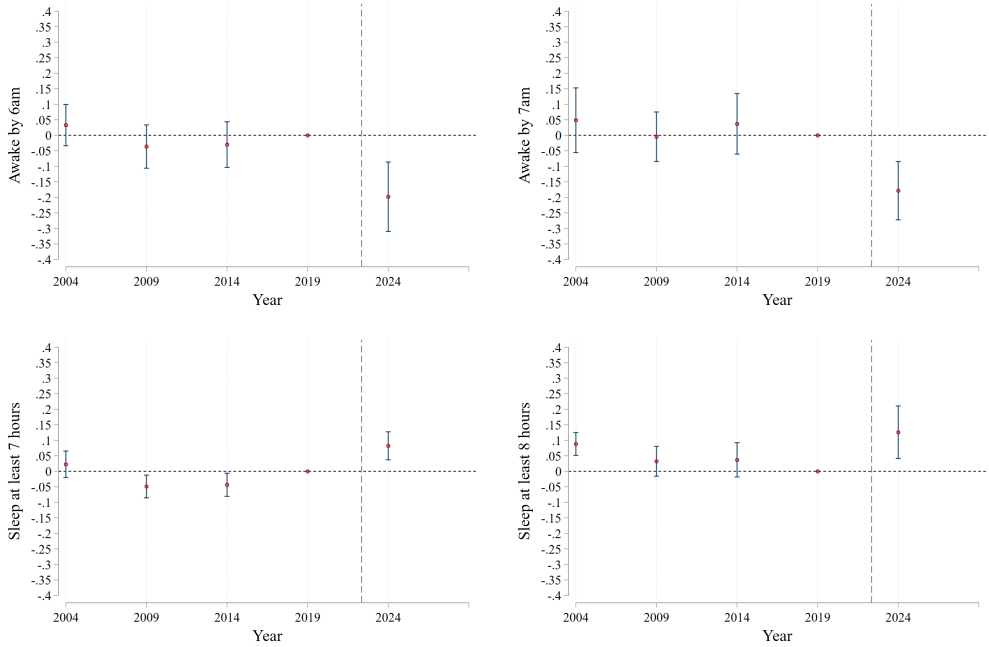
Table 3: SB 328, Wake-up Times and Sleep, Evidence from ATUS (2003–2024, Excluding 2020)

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Awake by			Bedtime			Sleep			
	6am	7am	8am	after 9pm	after 10pm	after 11pm	sleep hours	sleep≥6	sleep≥7	sleep≥8
Panel A: Children aged 15–18										
CA SB 328	-0.172***	-0.184***	-0.094***	-0.012	0.071*	0.034	0.464**	0.041*	0.115***	0.097**
	(0.045)	(0.028)	(0.033)	(0.034)	(0.037)	(0.045)	(0.207)	(0.024)	(0.024)	(0.044)
Observations	5,784	5,784	5,784	5,784	5,784	5,784	5,784	5,784	5,784	5,784
R-squared	0.069	0.153	0.148	0.054	0.055	0.047	0.058	0.020	0.027	0.042
Mean of dep. var.	0.283	0.591	0.714	0.693	0.354	0.225	9.120	0.942	0.862	0.703
Std. dev. of dep. var.	0.450	0.492	0.452	0.461	0.478	0.418	2.346	0.234	0.345	0.457
Panel B: Children under 13 (Single-Child Households)										
CA SB 328	-0.081***	-0.082***	-0.018	0.101***	-0.012	0.018	0.361***	0.043***	0.071***	0.106***
	(0.023)	(0.026)	(0.025)	(0.029)	(0.015)	(0.014)	(0.126)	(0.012)	(0.012)	(0.019)
Observations	11,822	11,822	11,822	11,822	11,822	11,822	11,822	11,822	11,822	11,822
R-squared	0.014	0.024	0.036	0.042	0.034	0.023	0.027	0.016	0.019	0.025
Mean of dep. var.	0.128	0.432	0.758	0.323	0.135	0.0656	9.657	0.965	0.938	0.876
Std. dev. of dep. var.	0.334	0.495	0.428	0.467	0.341	0.248	1.997	0.184	0.241	0.330

Notes: Data come from the 2003–2024 American Time Use Survey (ATUS). 2020 is omitted due to incomplete data collection during the COVID-19 pandemic, and observations are restricted to weekday diaries. Observations are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to individual's state of residence, survey year, age, sex, and race. Panel A restricts the analysis sample to children aged 15–18. Panel B restricts the sample to households with exactly one child under the age of 13. Panel A specifications include fixed effects for state, year, sex, age, month, SEER race, and Hispanic origin and control for state-level characteristics (% households with children, marital status and education-related variables, the share of 16–19 year-olds out of school, the unemployment rate, and the child poverty rate). Panel B specifications include fixed effects for state, year, month, SEER race, and Hispanic origin and control for state-level characteristics (total population, % households with children, share divorced, share with at least a high school degree, the unemployment rate, median household income, and the child poverty rate). Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table 3 Panel A turns to the ATUS time diaries of respondents aged 15–18 on weekdays. The pattern mirrors the YRBS: the share awake by 6 a.m. fell by 17.2 percentage points (61% of the mean; column 1), the share awake by 7 a.m. by 18.4 points (31%; column 2), and the share awake by 8 a.m. by 9.4 points (column 3); total sleep rose by about 28 minutes (column 7), and the shares sleeping at least 7 and 8 hours by 11.5 and 9.7 points (columns 9–10). Bedtimes moved little (columns 4–6).

Figure 3: California SB 328 and Sleep, ATUS (Adolescents Aged 15–18)

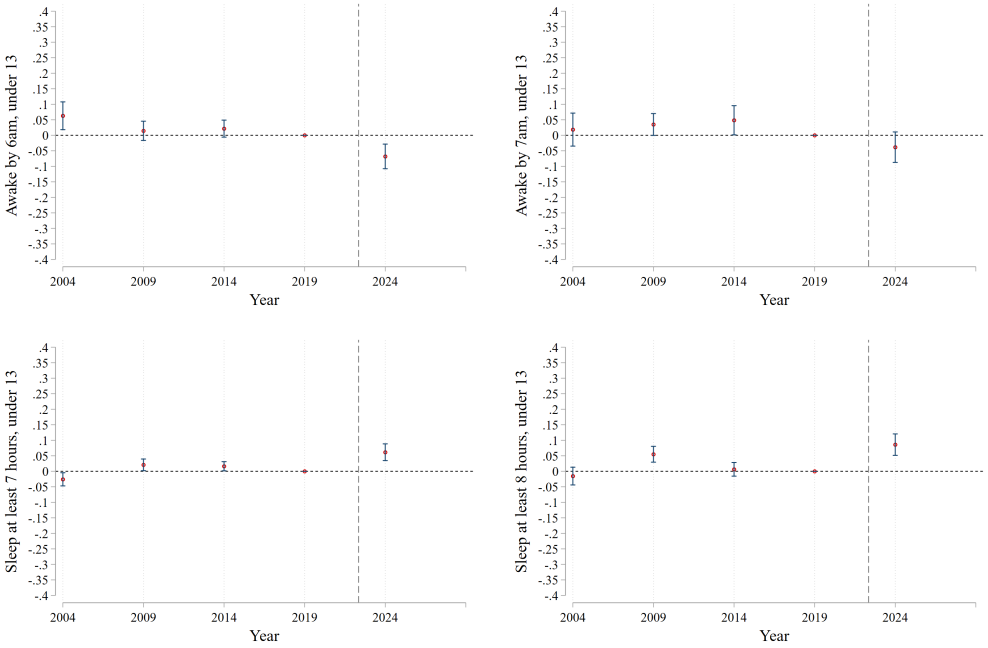


Notes: The figure illustrates the differences in sleep behavior among students in California and other states before and after the adoption of SB 328. Data come from the 2003–2024 American Time Use Survey (ATUS), excluding 2020. Observations are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to individual's state of residence, survey year, age, sex, and race. Survey years were pooled as follows: (1) 2004 combines 2003–2006; (2) 2009 combines 2007–2011; (3) 2014 combines 2012–2016; (4) 2019 combines 2017–2021 and January–August 2022 (excluding 2020); (5) 2024 combines September 2022–2024. Year fixed effects are at the bin level, whereas Table 3 uses annual year fixed effects, so the post-period coefficient and the pooled estimate need not coincide.

Figure 3 reports the corresponding event-study. Pre-period coefficients are small and show no trend toward the post-period estimates, although with five-year bins a few early bins differ from zero; the post-period estimates are of the same sign and similar magnitude as the difference-in-differences estimates.

We next examine sleep effects for children under 13, using ATUS data on the wake-up

Figure 4: California SB 328 and Sleep, ATUS (Children Under 13)



Notes: The figure illustrates the differences in sleep behavior among students in California and other states before and after the adoption of SB 328. Data come from the 2003–2024 American Time Use Survey (ATUS), excluding 2020. Observations are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to individual's state of residence, survey year, age, sex, and race. Survey years were pooled as follows: (1) 2004 combines 2003–2006; (2) 2009 combines 2007–2011; (3) 2014 combines 2012–2016; (4) 2019 combines 2017–2021 and January–August 2022 (excluding 2020); (5) 2024 combines September 2022–2024. Standard errors are clustered at the state level. Coefficient estimates are provided together with the 95% confidence intervals. Year fixed effects are at the bin level, whereas Table 3 uses annual year fixed effects, so the post-period coefficient and the pooled estimate need not coincide.

time and bedtime of the child in single-child households (Panel B of Table 3).[18] As in Panel A, we restrict to weekdays. Children under 13 were 8.1 percentage points less likely to be awake by 6 a.m. (a 63% decline relative to the mean; $p < 0.01$, column 1) and 8.2 percentage points less likely to be awake by 7 a.m. (19%; $p < 0.01$, column 2). We find no effect on the likelihood of being awake by 8 a.m., the time by which most elementary pupils must be up. Because the sample consists of single-child households, these shifts cannot reflect an older sibling's schedule. Two channels remain. First, districts may have moved elementary bell times together with secondary ones: the ACS arrival times of elementary and middle school teachers, who share an occupation code, shift later after 2022 (Table A.4), although we cannot separate the two groups. Second, the household's morning routine as a whole moved later, consistent with the later wake-up and work-arrival times among parents of school-age children documented in Section 4.5. The child's age is observed in single-child households, but only four treated diaries refer to children aged 11–12, too few to isolate the direct effect of the middle-school threshold; the estimates are driven by children aged 6–10, to whom the mandate did not apply.[19] We therefore read the under-13 results as evidence of household- and district-level schedule adjustment rather than of a direct effect of the mandate.

For sleep duration, we combine wake-up and bedtime reports within single-child households, where both can be confidently attributed to the same child. Sleep duration increased by approximately 22 minutes on average ($p < 0.01$, column 7). The share of children sleeping at least 7 hours rose by 7.1 percentage points (an 8% increase relative to the mean; $p < 0.01$, column 9), and the share sleeping at least 8 hours rose by 10.6 percentage points (12%; $p < 0.01$, column 10).

[18] Table A.8 reports results including all households with at least one child under 13. In these households, the reported wake-up time refers to the first child to rise and the bedtime to the last child to go to bed, so the two need not describe the same child and the implied sleep duration is mismeasured; the estimates are correspondingly attenuated and imprecise, and the sleep-duration coefficient turns negative. Table A.9 reports results excluding observations with imputed sleep below 2 or above 15 hours; the threshold estimates are very similar to the main estimates, while the estimates for hours of sleep are smaller and imprecise.

[19] Restricting to children aged 6–10 (17 treated diaries), the share awake by 6 a.m. falls by 16 percentage points and sleep duration rises by about 75 minutes; for ages 11–12 the estimates are imprecise and change sign across outcomes.

Table 4: SB 328 and Mental Health, YRBS (2007-2023)

	(1) Difficulty concentrating	(2) Felt sad/hopeless	(3) Suicidal intentions
CA SB 328	-0.054*	-0.024	-0.013
	(0.027)	(0.015)	(0.011)
Observations	48,199	134,396	119,283
R-squared	0.053	0.067	0.030
Mean of dep. var.	0.389	0.319	0.181
Std. dev. of dep. var.	0.487	0.466	0.385

Notes: Data comes from the 2007-2023 national samples of the Youth Risk Behavior Survey (YRBS). Observations are weighted by population weights obtained from National Cancer Institute's SEER Program and assigned according to individual's state of residence, survey year, age, sex, and race. All specifications control for individual-level covariates (grade, race/ethnicity, sex) and state-level characteristics (population, population density, % households with children, marital status & education-related variables, unemployment rate, median household income, and child poverty rate). All regressions further include state and year fixed effects. Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.

Figure 4 reports the corresponding event-study for children under 13. The pre-2022 coefficients show no trend toward the post-period estimates, with no sign of a California-specific secular trend in children's sleep behavior, although some early bins differ from zero. Post-period estimates show declines in the share awake by 6 a.m. and 7 a.m. and increases in the share sleeping at least 7 or 8 hours.[20]

As a robustness check, we restrict the sample to metropolitan areas, where the mandate applied; the estimates are essentially unchanged (Table A.10).[21] Results excluding summer and holiday diaries are broadly similar, although the smaller sample makes several estimates less precise and the under-13 estimate for awake by 8 a.m. changes sign (Table A.11). Because the ATUS is collected continuously through the year, we also re-estimate the models with academic-year (September–August) rather than calendar-year fixed effects; again the

[20]The event-study point estimates are somewhat smaller than the pooled DiD estimates because they reflect a single post-period bin (2022–2024) rather than averaging across the full sample.

[21]The ATUS contains no rural California diaries after 2022, so a within-California urban–rural contrast cannot be estimated; the table reports the California × post-SB 328 coefficient among metropolitan respondents, controlling for metropolitan status and its interactions with California and the post period.

estimates are very similar (Table A.12).

Putting these estimates on the per-hour footing used in Section 2 requires a measure of how much start times moved, which none of our data provide directly for the cross-state comparison. The within-California shift in teachers' arrival times (19 minutes between more- and less-exposed counties, Table A.4) and the movement of students' arrivals out of the hour before 8:30 a.m. (Table 1) both suggest an average delay well under an hour, on the order of 20–30 minutes. Scaled by that range, the ATUS gain of 28 minutes of sleep corresponds to roughly 55–85 minutes per hour of delay, and the outlier-trimmed estimate of 19 minutes (Table A.9) to about 40–60; these are at or above the 20–45 minutes per hour typical of non-boarding studies and rest on 23 treated diaries, so we read them as imprecise estimates to be interpreted cautiously rather than as evidence of unusually large effects.

4.2 Effects on mental well-being

Table 4 reports YRBS estimates for mental health outcomes. We find a marginally significant decline of approximately 5 percentage points in the likelihood of reporting difficulty concentrating or making decisions due to physical or mental health conditions (a 14% reduction relative to the mean; $p < 0.10$, column 1). Pooled effects on sadness/hopelessness (-2.4 pp; $p = 0.13$, column 2) and suicidal ideation (-1.3 pp; $p = 0.25$, column 3) are negatively signed but imprecisely estimated, and the corresponding event-study estimates do not reach statistical significance.

These muted average effects mask heterogeneity, which we document in Section 5. Boys exhibit significant declines in sadness/hopelessness (-2.3 pp; $p < 0.05$) and suicidal ideation (-3.8 pp; $p < 0.01$), in addition to a decline in difficulty concentrating (-5.6 pp; $p < 0.10$), and Hispanic students show the largest reduction in difficulty concentrating (-9.9 pp; $p < 0.01$). Whether these groups also experienced the largest sleep gains is less clear-cut, as we discuss later in the text.

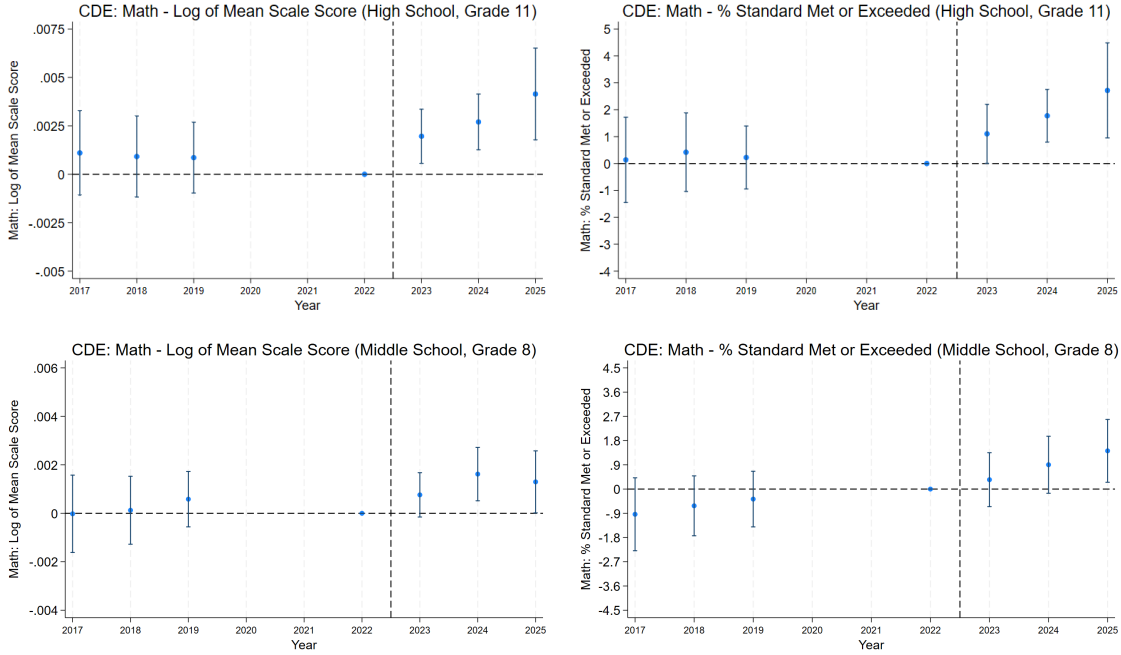
4.3 Effects on academic outcomes

We then turn to the effects of delayed school start times on academic outcomes. Because statewide test-score trends after 2022 reflect the uneven recovery from the pandemic across states, a cross-state comparison cannot cleanly attribute changes in California's scores to the mandate. We therefore use the within-state design described in Section 3.3: we proxy for pre-policy school start times using teachers' commute arrival times in the ACS and compare changes in school outcomes before and after 2022 in counties where the mandate required a larger shift—where teachers used to arrive earlier—with changes in counties whose schools already started late or were exempt. This design absorbs statewide shocks, including the pandemic and its aftermath, that would confound a cross-state comparison; it does not rule out differences in recovery across counties that happen to be correlated with pre-policy start times, a caveat we keep in mind when interpreting the estimates. Outcomes are drawn from school-level California Department of Education (CDE) records on academic performance and disciplinary measures, and the results are presented in Figures 5–9 and Table 5.

Figure 5 reports within-state event-study estimates for math achievement at the high school (Grade 11) and middle school (Grade 8) levels; estimation uses the full 2014–15 onward sample and the figures display coefficients from 2017 onward. Pre-period estimates are close to zero, consistent with parallel trends between more- and less-exposed districts within California. Post-period estimates turn positive and increase year-over-year through 2025. At the high school level, the share of students meeting or exceeding the state math standard rises by approximately 1.1 percentage points in 2023 and 2.7 percentage points by 2025; log mean scale scores show the same upward pattern, with 2025 estimates corresponding to a 0.4% gain.[22] Middle school effects follow a similar trajectory but are smaller in magnitude, with the share meeting or exceeding standards rising by approximately 1.4 percentage points by 2025, consistent with the weaker bite of the mandate at the middle school level. The

[22]The percentage-point movement on the share-meeting-standards is mechanically larger than the log-scale-score movement when test scores cluster near the proficiency threshold, so the two outcomes need not scale linearly.

Figure 5: Within-State Analysis – Effect of SB 328 on Math Achievement (CDE)

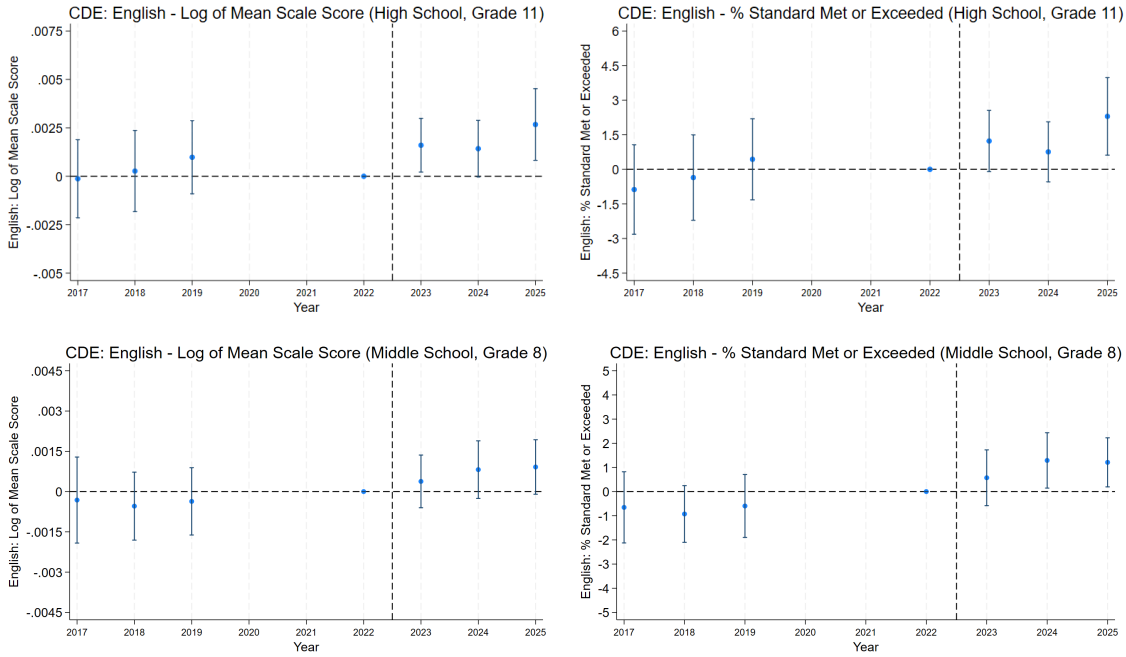


Notes: The figure illustrates the differences in math standardized test performance among high school (Grade 11) and middle school (Grade 8) students within California before and after the adoption of SB 328. Data come from the California Department of Education for the academic years 2014–15 through 2024–25. All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, and female students, and the share of socioeconomically disadvantaged students. Academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. Standard errors are clustered at the county level. Coefficient estimates are provided together with the 95% confidence intervals.

accumulating pattern across 2023–2025 is consistent with academic benefits compounding with exposure to the new schedule.

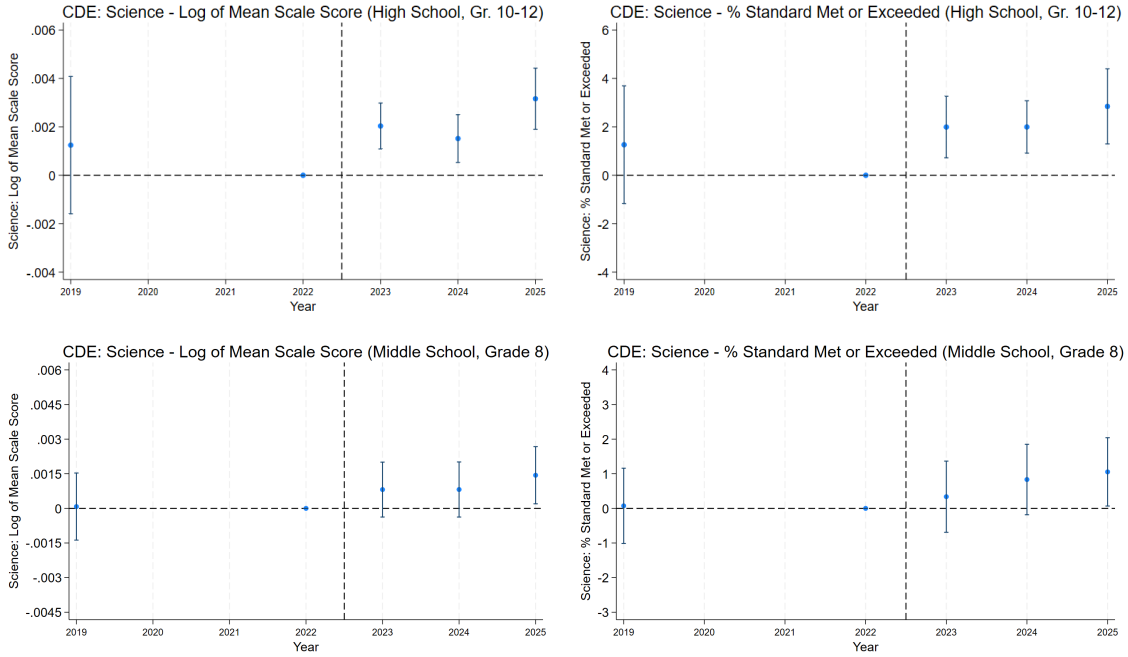
Figures 6 and 7 extend the within-state analysis to English language arts/literacy and science. As with math, pre-period estimates are close to zero and show no pre-trends. Following implementation, English scores improve at both school levels: the share meeting or exceeding standards rose by approximately 2.3 percentage points at the high school level and 1.2 percentage points at the middle school level by 2025, with log mean scale scores following a similar upward trajectory. Science gains are particularly notable at the high school level (Grades 10–12), where the share meeting or exceeding standards increased by approximately 2 percentage points in 2023–2024 and 2.9 percentage points by 2025, with all post-period estimates statistically distinguishable from zero; middle school science effects

Figure 6: Within-State Analysis – Effect of SB 328 on English Achievement (CDE)



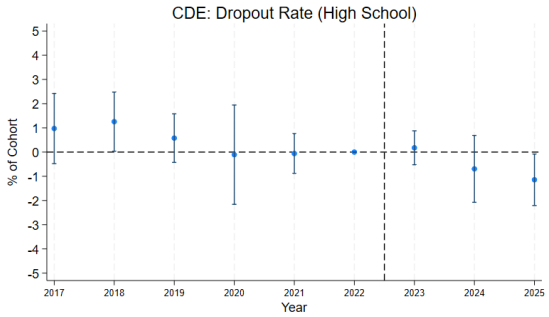
Notes: The figure illustrates the differences in English standardized test performance among high school (Grade 11) and middle school (Grade 8) students within California before and after the adoption of SB 328. Data come from the California Department of Education for the academic years 2014–15 through 2024–25. All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, and female students, and the share of socioeconomically disadvantaged students. Academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. Standard errors are clustered at the county level. Coefficient estimates are provided together with the 95% confidence intervals.

Figure 7: Within-State Analysis – Effect of SB 328 on Science Achievement (CDE)



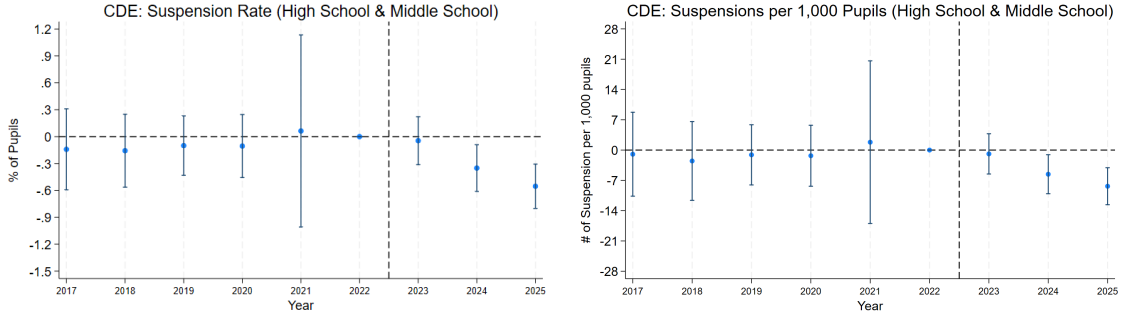
Notes: The figure illustrates the differences in science standardized test performance among high school (Grades 10-12) and middle school (Grade 8) students within California before and after the adoption of SB 328. Data come from the California Department of Education for the academic years 2018–19 through 2024–25. All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, and female students, and the share of socioeconomically disadvantaged students. Academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. Standard errors are clustered at the county level. Coefficient estimates are provided together with the 95% confidence intervals.

Figure 8: Within-State Analysis – Effect of SB 328 on Dropout Rate (CDE)



Notes: The figure illustrates the differences in high school cohort-specific dropout rate among students within California before and after the adoption of SB 328. Data come from the California Department of Education for the academic years 2016–17 through 2024–25. All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, and female students, and the share of socioeconomically disadvantaged students. Standard errors are clustered at the county level. Coefficient estimates are provided together with the 95% confidence intervals.

Figure 9: Within-State Analysis – Effect of SB 328 on Suspensions (CDE)



Notes: The figure illustrates the differences in school-level suspension outcomes among students within California before and after the adoption of SB 328. Estimation sample combines high schools and middle schools. Data come from the California Department of Education for the academic years 2011–12 through 2024–25. All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, and female students, and the share of socioeconomically disadvantaged students. Standard errors are clustered at the county level. Coefficient estimates are provided together with the 95% confidence intervals.

are positive but less precisely estimated.

Figures 8 and 9 report effects on dropout rates and suspensions, available only in the within-state CDE data. The high school dropout rate declines by approximately 1.3 percentage points by 2025, though we interpret this pattern cautiously given some positive pre-period drift. Suspension rates show a suggestive and accumulating decline through the post-period.

Table 5 reports the corresponding two-way fixed effects estimates. Panel A confirms positive effects on high school achievement: math, English, and science estimates are statistically significant across both outcome measures (log mean scale score and share meeting standards). Panel B shows broadly consistent results for middle school, with math, English, and science each yielding statistically significant effects on at least one measure. Panel C reports dropout rates and disciplinary outcomes: a marginally significant decline in the high school dropout rate of approximately 0.97 percentage points ($p < 0.10$), a marginally significant decline in the suspension rate of 0.32 percentage points ($p < 0.10$), and a corresponding reduction of approximately 6.8 suspensions per 1,000 students ($p < 0.10$).

As a descriptive complement for high school seniors, we examine state-level SAT scores (Appendix Table A.5). California's average total SAT score rose by approximately 25 points

(a 2% gain relative to the mean) relative to the national trend following SB 328's implementation, with gains of roughly 11 points on math and 13 points on English, Reading, and Writing (ERW). We interpret this pattern as descriptive rather than causal. SAT data are available only at the state level, and the composition of SAT test-takers changed substantially during the sample period as colleges adopted test-optional admissions in response to COVID-19; California saw both a sharp decline in SAT participation and a rise in average scores after 2020, consistent with selection into testing by higher-achieving students. More generally, the uneven recovery from the pandemic across states means that cross-state comparisons of test scores after 2022 cannot isolate the policy. The within-state CAASPP estimates for grade 11 are not affected by test-optional selection and net out statewide trends, although they remain exposed to differential trends across counties.

4.4 Robustness checks

We evaluate our findings using permutation tests (Figure 10). In each of 2,000 draws we assign a placebo treatment to a randomly chosen state and a randomly chosen adoption year between 2007 and 2023, re-estimate equation (1) with the placebo indicator in place of SB 328, and compare the observed California coefficient with the distribution of placebo coefficients. This addresses inference with a single treated state. We focus on the outcomes for which the baseline specifications detected statistically significant effects.

For sleep, the observed effects on sleeping at least 6 and at least 7 hours lie in the extreme tails of the placebo distribution (one-sided empirical p-values below 0.001 and of 0.001); the effect on sleeping at least 8 hours has an empirical p-value of 0.06, and the effect on difficulty concentrating 0.03. Appendix Figures A.3 and A.4 report permutation tests for the ATUS sleep outcomes.

We supplement the permutation tests with two placebo checks. First, we test whether SB 328 affected school-level characteristics that should not respond to the policy (total enrollment, racial composition, gender composition, and the share of economically disadvantaged

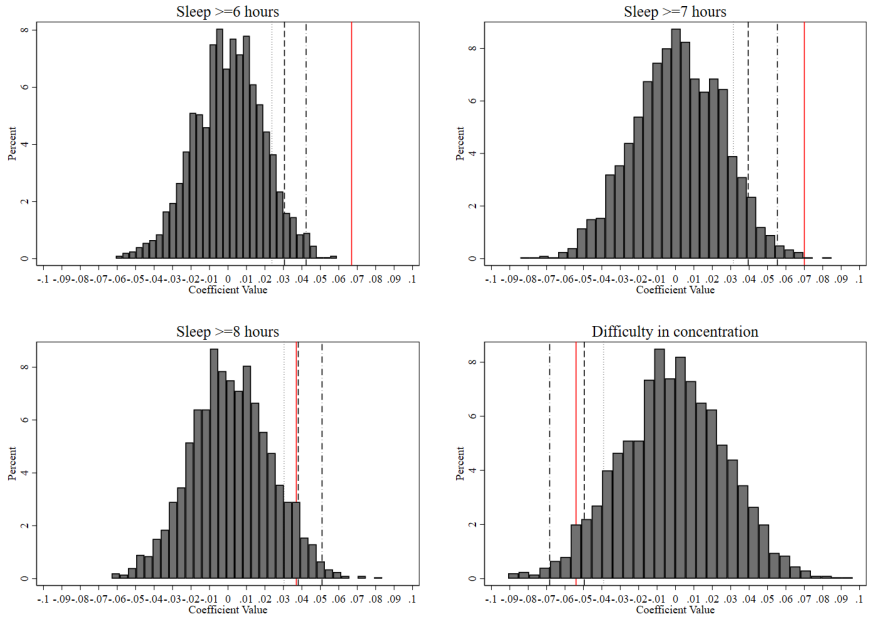
Table 5: Within-State Analysis – Effect of SB 328 on Academic and School Outcomes, CDE

	(1)	(2)	(3)	(4)	(5)	(6)
	Math		English		Science	
	ln score	% met	ln score	% met	ln score	% met
Panel A: High School (Grade 11 / Grades 10–12)						
Policy (SB 328)	0.0021**	1.823**	0.0016**	1.745***	0.0017***	1.729***
	(0.001)	(0.686)	(0.001)	(0.591)	(0.001)	(0.627)
Observations	13,653	13,660	13,668	13,676	8,014	8,014
R-squared	0.163	0.092	0.076	0.064	0.245	0.143
Mean of dep. var.		24.81		50.93		26.32
Std. dev. of dep. var.		20.47		23.46		18.27
Panel B: Middle School (Grade 8)						
Policy (SB 328)	0.0012*	1.700***	0.0010*	1.697***	0.0010**	0.693*
	(0.001)	(0.483)	(0.001)	(0.539)	(0.000)	(0.395)
Observations	17,236	17,243	17,246	17,252	10,116	10,116
R-squared	0.893	0.875	0.867	0.852	0.034	0.041
Mean of dep. var.		33.95		48.09		30.17
Std. dev. of dep. var.		20.79		19.37		18.33

	(1)	(2)	(3)
	Dropout Rate	Suspension Rate	No. Suspensions per 1,000
Panel C: Dropout rates and Disciplinary Outcomes			
Policy (SB 328)	-0.967*	-0.324*	-6.776*
	(0.562)	(0.178)	(3.799)
Observations	14,507	75,541	75,541
R-squared	0.023	0.168	0.105
Mean of dep. var.	8.700	3.624	60.56
Std. dev. of dep. var.	12.28	5.279	117.8

Notes: Data come from the California Department of Education for the academic years 2014–15 through 2024–25 (Panels A and B, cols. 1-4); academic years 2018–19 through 2024–25 (Panels A and B, cols. 5-6); academic years 2016–17 through 2024–25 (Panel C, col. 1); and academic years 2011–12 through 2024–25 (Panel C, cols. 2-3). In Panels A and B, academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. Panel C contains school-wide outcomes, with high school-only sample in col. (1), and a combined high school/middle school sample in cols. (2)-(3). All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, and female students, and the share of socioeconomically disadvantaged students. Regressions in Panel A (and Panel C, col. 1) use teacher-arrival treatment intensity defined for high school teachers (IPUMS occupation 2320; bottom three quartiles vs. top quartile plus rural counties). Regressions in Panel B use teacher-arrival treatment intensity defined for elementary/middle school teachers (IPUMS occupation 2310; bottom two quartiles vs. top two quartiles plus rural counties). Panel C (cols 2-3) regressions use a combined HS+ES/MS teacher-arrival treatment intensity (IPUMS occupations 2320 and 2310; bottom three quartiles vs. top quartile plus rural counties). Standard errors in parentheses, clustered at the county level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Figure 10: Permutation Tests



Notes: The figure illustrates the results of permutation tests for various outcomes, where the policy adoption year was randomly reassigned to other states to create placebo scenarios. Sleep and mental health outcomes come from the YRBS. The red vertical line represents the observed treatment effect, while the dashed lines indicate the 1%, 5%, and 10% significance thresholds derived from the placebo distribution. By comparing the observed effect to the distribution of placebo effects, the test evaluates whether the observed policy effect is statistically significant or could plausibly arise by random chance. Standard errors are clustered at the state level.

students) using CDE data (Table A.7). As discussed in Section 3.3, the estimates are small relative to the means and their confidence intervals exclude changes larger than about three percentage points in any demographic share, making compositional change an unlikely driver of the academic results.

Second, we test whether the SB 328 treatment predicts adolescent health outcomes that should be unaffected by school schedules, specifically self-reported vision and hearing difficulties in the ACS (Table A.6). The placebo test uses both levels of variation: the cross-state comparison (columns 1–2) addresses the concern that California-specific health trends might confound our YRBS and ATUS estimates, and the within-state comparison (columns 3–4) addresses the concern that our ACS-based teacher-arrival proxy might be capturing broader correlated trends rather than schedule changes. We find no significant effects in either specification.

One concern these checks address only partly is the pandemic: the post-period coincides with the recovery from COVID-19, and if adolescents' sleep or mental health in California recovered differently from other states for reasons unrelated to the mandate, the cross-state estimates would absorb that difference. The permutation tests show that few other states display changes of the observed size over the same period, but we cannot rule this out entirely. Together, these robustness checks make random chance, compositional shifts, and broader correlated trends unlikely explanations for our results.

4.5 Adjustment margins: adolescents' activities and parents' schedules

As argued in the introduction, a statewide mandate has effects beyond school schedules. Table 6 uses the ATUS to document how adolescents' days changed (Appendix Figure A.8 reports the corresponding event studies), and Appendix Table A.13 turns to their parents, combining ATUS diaries with commuting data from the ACS.

Among adolescents aged 15–18, weekday participation in sports and exercise declined by

13 percentage points (from a mean of 34%; $p < 0.10$), and among those who did exercise, the share doing so after 4 p.m. rose by 19 percentage points (from a mean of 66%; $p < 0.01$): sport became less frequent and moved later in the day. Paid work did not change. Time on the phone, messaging, and social media rose by about 22 minutes (from a mean of 15). The diaries do not distinguish school-based from outside activities, so we cannot tell whether the decline reflects fewer activities overall or activities moved into the school day.

Parents of children aged 11–18 adjusted in parallel (Appendix Table A.13). In the ATUS they were 6.5 percentage points less likely to be awake by 6 a.m. (from a mean of 58%) and 5.8 points more likely to be still at work after 4 p.m. The ACS, with close to two million parent-year observations, shows the corresponding shift at the start of the workday: the share arriving at work between 7 and 8 a.m. fell by 1.1 percentage points (4% of the mean) and the share arriving between 8 and 9 a.m. rose by 1.2 points (6%), while commuting time did not change (+0.02 minutes on a mean of 28). Parents' workdays thus started and ended slightly later rather than becoming shorter. Appendix Figure A.9 reports the corresponding event-study estimates; pre-period coefficients are close to zero.

The ATUS estimates rest on few treated diaries (23 adolescents) and we therefore read the time-use evidence as descriptive of the direction of household adjustment rather than as precise magnitudes; the ACS commuting estimates are precisely estimated but small. Taken together, they indicate that the mandate's sleep gains came with some rescheduling—less, and later, weekday sport for adolescents and a slightly later workday for parents—that a cost–benefit assessment of statewide adoption needs to weigh.

5 Heterogeneity

There are reasons to expect the effects of a later start time to differ across groups. Sex differences in circadian timing make early schedules more binding for boys: from late adolescence men's sleep on free days is centered later in the night than women's, and their

Table 6: SB 328 and Adolescents' Non-School Time Use, Evidence from ATUS

	(1)	(2)	(3)	(4)
	Sports activities		Paid work	Phone/screen
	any	after 4pm | any	any	minutes
CA SB 328	-0.127*	0.191***	-0.034	21.86***
	(0.074)	(0.059)	(0.025)	(3.57)
Observations	5,784	1,936	5,784	5,784
R-squared	0.051	0.065	0.088	0.031
Mean of dep. var.	0.335	0.663	0.149	14.50
Std. dev. of dep. var.	0.472	0.473	0.356	45.59

Notes: Data come from the 2003–2024 American Time Use Survey (ATUS), excluding 2020 and restricted to weekday diaries of respondents aged 15–18. Each column reports a separate difference-in-differences regression of the outcome on a California × post-SB 328 indicator (post = September 2022 onward). Observations are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to an individual's state of residence, survey year, age, sex, and race. All specifications include state, year, interview month, sex, age, SEER race and Hispanic-origin fixed effects, and control for state-level characteristics (% households with children, marital status & education-related variables, the unemployment rate, and the child poverty rate). Sleep outcomes for this sample are reported in Table 3. Column (1) is an indicator for any participation in sports or exercise on the diary day; column (2) is, conditional on any sports or exercise that day, an indicator for sports or exercise after 4pm; column (3) is an indicator for any paid work; and column (4) is defined as minutes spent on telephone calls, email, messaging, and social media. Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.

chronotype continues to delay for about a year longer (Roenneberg et al., 2004; Fischer et al., 2017). Hispanic and economically disadvantaged adolescents are less likely to obtain sufficient sleep than their White and more affluent peers (Guglielmo et al., 2018; Giddens et al., 2022), reflecting household constraints (i.e., parents' work schedules, longer commutes, shared bedrooms, and norms about bedtimes and evening activities) that an earlier start time compounds and a later one relaxes. On this view, the groups with the most binding morning constraints before the policy should gain the most sleep from it. Whether those gains translate into mental-health and academic improvements depends in addition on baseline levels and on complementary inputs at school and at home, so we do not expect a one-to-one mapping. Two caveats apply throughout this section. First, the heterogeneity analysis was not pre-specified and we present it as exploratory. Second, the ATUS treated cell is too small to split by gender or ethnicity, so the subgroup analysis relies on the YRBS, with several thousand California respondents per wave, and on the CDE academic data.

5.1 Effects by gender

Table A.17 reports the heterogeneity of the effects on sleep and mental wellbeing by gender using YRBS data, with separate rows for girls and boys. Among girls in grades 9–12, SB 328 led to a 6.9 percentage point increase in the likelihood of sleeping at least 6 hours (+9%, $p < 0.01$, column 1); a 5.2 percentage point increase in the likelihood of sleeping at least 7 hours (+10%, $p < 0.05$, column 2); and a non-statistically significant 3.2 percentage point increase in the likelihood of sleeping at least 8 hours (column 3). We also find a 5.5 percentage point decline in the likelihood that girls reported difficulties making decisions or concentrating because of physical or mental health problems (−12%, $p < 0.10$, column 4) and a non-statistically significant 3.4 percentage point decline (column 5) in the likelihood of reporting feelings of sadness and hopelessness. There is no evidence of a decline in suicidal intentions.

Effects on sleep appear larger in magnitude among boys (Boys rows), although the gender difference is small relative to the standard errors. Specifically, SB 328 led to a 6.8 percentage point increase in the likelihood of sleeping at least 6 hours among boys in grades 9–12 (+8.4%, $p < 0.01$, column 1); an 8.6 percentage point increase in the likelihood of sleeping at least 7 hours (+15%, $p < 0.01$, column 2); and a 4.1 percentage point increase in the likelihood of sleeping at least 8 hours (+14%, $p < 0.05$, column 3). On mental well-being, we find a marginally significant 5.6 percentage point decline in the likelihood of reporting difficulties concentrating among boys (−19%, $p < 0.10$, Boys rows, column 4). We estimate a 2.3 percentage point decline in the likelihood of reporting feelings of sadness and hopelessness (−10%, $p < 0.05$, column 5), and a significant decline of 3.8 percentage points in suicidal intentions (−31%, $p < 0.01$, column 6). We caution that the cross-gender contrast on sadness and hopelessness reflects differences in precision rather than in point estimates: the girls' coefficient (−3.4 pp) is in fact larger in magnitude than the boys' (−2.3 pp), but is less precisely estimated, and the two are not statistically distinguishable from one another.

The CDE estimates by gender (Tables A.18 and A.19) show positive academic effects

for both groups; at the high school level the effects on the share meeting standards are statistically significant in math, English, and science for both girls and boys, while at the middle school level they are significant across subjects for girls and for math only among boys, with point estimates somewhat larger for girls. The gender differences lie within the standard error bands, so we do not read them as evidence that one group gained more.

5.2 Heterogeneity by ethnicity

Hispanic students represent approximately 56% of California's public school enrollment, making heterogeneity across ethnic groups especially relevant for the policy's distributional implications.

Hispanic students. In the YRBS (Table A.20), Hispanic students show significant sleep gains: a 5.7 percentage point increase in the likelihood of sleeping at least 6 hours (7%; $p < 0.10$, column 1) and a 4 percentage point increase in the likelihood of sleeping at least 7 hours (7%; $p < 0.01$, column 2). On mental health, the likelihood of reporting difficulty concentrating fell by 9.9 percentage points (-25%; $p < 0.01$, column 4), substantially larger than the overall estimate. Effects on sadness, hopelessness, and suicidal intentions are not statistically significant in the YRBS. Hispanic students thus show the largest reduction in difficulty concentrating, but their sleep gains are, if anything, smaller than those of White students (columns 1–3).

Academic gains are also concentrated among Hispanic students. The within-state CDE estimates show that Hispanic students have the largest and most consistently significant gains across all subjects at both school levels (Tables A.21 and A.22): at the high school level, the share meeting or exceeding standards rises by approximately 1.9 percentage points in both math and English, slightly larger than the overall estimates of 1.8 points in math and 1.7 in English.

White students. White students show somewhat larger sleep gains in the YRBS but smaller mental health and academic effects. The likelihood of sleeping at least 6 hours increased by 8 percentage points (+10%; $p < 0.01$), at least 7 hours by 9.3 percentage points (+16%; $p < 0.01$), and at least 8 hours by 5.6 percentage points (+20%; $p < 0.05$) in Table A.20. On mental health, only the decline in sadness or hopelessness is marginally significant (-3.1 pp, -10%; $p < 0.10$, column 5); the concentration effect documented for Hispanic students is not present for White students. Academic effects in the CDE data are smaller and less precisely estimated than for Hispanic students.

Other groups and economic status. Sleep and mental health effects for other ethnic groups are not statistically significant (bottom rows of Table A.20). Beyond ethnicity, socioeconomically disadvantaged students also show significant academic gains in the CDE data (Panel C of Tables A.21 and A.22): at the high school level, the share meeting standards rises by 1.5 percentage points in math and English and by 1.9 points in science, with a significant decline in the dropout rate. These gains are larger than those of White students, though smaller than the full-sample estimates in math and English. Across groups, the concentration and academic gains are consistently largest for Hispanic students, whereas sleep gains in the YRBS are not. We view the heterogeneity as exploratory.

6 Conclusions

California's Senate Bill 328 increased the share of students sleeping at least 8 hours per night by 13%, with suggestive evidence of larger gains among boys. Time-use evidence further reveals that wake-up times shifted later for adolescents and for younger children in affected households, and that parents of school-age children woke up and arrived at work later, indicating that the mandate shifted the schedule of the whole household rather than only that of the students it covered. Average mental health effects are imprecisely estimated. Boys show statistically significant declines in sadness/hopelessness and suicidal ideation, and

Hispanic students show a large reduction in difficulty concentrating; we read these subgroup patterns as exploratory.

The policy improved academic outcomes. Within California, high schools in counties with earlier pre-policy start times saw increases of about 1.7–1.8 percentage points in the share of students meeting state standards in math, English, and science relative to schools in less-exposed counties. We find comparable gains in math and English and a smaller gain in science in grade 8, with academic gains accumulating over 2023–2025. We also find a suggestive decline in high school dropout rates, though we interpret this cautiously given some pre-period drift. Academic gains are largest in magnitude among Hispanic students and statistically significant among economically disadvantaged students, suggesting that SB 328 may also narrow achievement gaps.

The mandate's broad implementation also allows us to examine adjustments beyond school schedules. We find suggestive evidence that adolescents were less likely to participate in sports on weekdays and, when they did, participated later in the day. Parents of affected children arrived at work slightly later, with no detectable change in commuting time. These household adjustments, together with districts' transition costs, should be considered alongside the sleep and achievement gains when assessing statewide adoption.

With Florida's 2023 mandate substantially weakened by 2025 legislation and other states still weighing comparable measures, California's experience provides the first evidence on the effects of a statewide mandate. Longer follow-up will establish whether the academic gains persist or grow and whether they translate into changes in college enrollment and labor market entry.

References

Adusumilli, K., Agostinelli, F., and Borghesan, E. (2024). Heterogeneity and endogenous compliance: Implications for scaling class size interventions. Technical report, National

Bureau of Economic Research.

Attia, P. (2023). *Outlive: The science and art of longevity*. Harmony.

Bastian, K. C. and Fuller, S. C. (2018). Answering the bell: High school start times and student academic outcomes. *Aera Open*, 4(4):2332858418812424.

Bastian, K. C. and Fuller, S. C. (2023). Late but right on time? school start times and middle grade students' engagement and achievement outcomes in north carolina. *American Journal of Education*, 129(2):177–203.

Bessone, P., Rao, G., Schilbach, F., Schofield, H., and Toma, M. (2021). The economic consequences of increasing sleep among the urban poor. *The Quarterly Journal of Economics*, 136(3):1887–1941.

Billari, F. C., Giuntella, O., and Stella, L. (2018). Broadband internet, digital temptations, and sleep. *Journal of Economic Behavior & Organization*, 153:58–76.

Boergers, J., Gable, C. J., and Owens, J. A. (2014). Later school start time is associated with improved sleep and daytime functioning in adolescents. *Journal of Developmental & Behavioral Pediatrics*, 35(1):11–17.

Bowers, J. M. and Moyer, A. (2017). Effects of school start time on students' sleep duration, daytime sleepiness, and attendance: a meta-analysis. *Sleep health*, 3(6):423–431.

Cappuccio, F. P., D'Elia, L., Strazzullo, P., and Miller, M. A. (2010). Sleep duration and all-cause mortality: A systematic review and meta-analysis of prospective studies. *Sleep*, 33(5):585–592.

Carrell, S. E., Maghakian, T., and West, J. E. (2011). A's from Zzzz's? The causal effect of school start time on the academic achievement of adolescents. *American Economic Journal: Economic Policy*, 3(3):62–81.

Carskadon, M. A. (2002). Factors influencing sleep patterns of adolescents.

Carskadon, M. A., Wolfson, A. R., Acebo, C., Tzischinsky, O., and Seifer, R. (1998). Adolescent sleep patterns, circadian timing, and sleepiness at a transition to early school days. *Sleep*, 21(8):871–881.

Chan, C. S., Poon, C. Y., Leung, J. C., Lau, K. N., and Lau, E. Y. (2018). Delayed school start time is associated with better sleep, daytime functioning, and life satisfaction in residential high-school students. *Journal of adolescence*, 66:49–54.

Chan, C. S., Tang, M. C., Leung, J. C., Poon, C. Y., and Lau, E. Y. (2024). Delayed school start time is associated with better sleep, mental health, and life satisfaction among residential high-school students: a prospective study. *Sleep*, 47(11):zsae171.

Cortes, K. E., Bricker, J., and Rohlfs, C. (2012). The role of specific subjects in education production functions: Evidence from morning classes in chicago public high schools. *The BE Journal of Economic Analysis & Policy*, 12(1).

Creswell, J. D., Tumminia, M. J., Price, S., Sefidgar, Y., Cohen, S., Ren, Y., Brown, J., Dey, A. K., Dutcher, J. M., Villalba, D., et al. (2023). Nightly sleep duration predicts grade point average in the first year of college. *Proceedings of the National Academy of Sciences*, 120(8):e2209123120.

Crowley, S. J., Acebo, C., and Carskadon, M. A. (2007). Sleep, circadian rhythms, and delayed phase in adolescence. *Sleep medicine*, 8(6):602–612.

Dewald, J. F., Meijer, A. M., Oort, F. J., Kerkhof, G. A., and Bögels, S. M. (2010). The influence of sleep quality, sleep duration and sleepiness on school performance in children and adolescents: A meta-analytic review. *Sleep Medicine Reviews*, 14(3):179–189.

Dunster, G. P., de la Iglesia, L., Ben-Hamo, M., Nave, C., Fleischer, J. G., Panda, S., and de la Iglesia, H. O. (2018). Sleepmore in Seattle: Later school start times are associ-

ated with more sleep and better performance in high school students. *Science Advances*, 4(12):eaau6200.

Edwards, F. (2012). Early to rise? the effect of daily start times on academic performance. *Economics of Education Review*, 31(6):970–983.

Fischer, D., Lombardi, D. A., Marucci-Wellman, H., and Roenneberg, T. (2017). Chronotypes in the US – influence of age and sex. *PLoS ONE*, 12(6):e0178782.

Foster, E. M. and Kalil, A. (2007). Living arrangements and children's development in Low-Income White, Black, and Latino families. *Child Development*, 78(6):1657–1674.

Fuller, S. C. and Bastian, K. C. (2024). Saved by the bell schedule? the effects of a later school start time on high schoolers in an urban district. *Urban Education*, 59(5):1552–1585.

Gibson, M. and Shrader, J. (2018). Time use and labor productivity: The returns to sleep. *Review of Economics and Statistics*, 100(5):783–798.

Giddens, N. T., Juneau, P., Manza, P., Wiers, C. E., and Volkow, N. D. (2022). Disparities in sleep duration among american children: effects of race and ethnicity, income, age, and sex. *Proceedings of the National Academy of Sciences*, 119(30):e2120009119.

Giuntella, O., Han, W., and Mazzonna, F. (2017). Circadian rhythms, sleep, and cognitive skills: Evidence from an unsleeping giant. *Demography*, 54(5):1715–1742.

Giuntella, O. and Mazzonna, F. (2019). Sunset time and the economic effects of social jetlag: Evidence from US time zone borders. *Journal of Health Economics*, 65:210–226.

Giuntella, O., Saccardo, S., and Sadoff, S. (2026). Sleep: Educational impact and habit formation. *Journal of Political Economy*, 134(8).

Grandner, M. A. (2022). Sleep, health, and society. *Sleep medicine clinics*, 17(2):117–139.

Groen, J. A. and Pabilonia, S. W. (2019). Snooze or lose: High school start times and academic achievement. *Economics of Education Review*, 72:204–218.

Guglielmo, D., Gazmararian, J. A., Chung, J., Rogers, A. E., and Hale, L. (2018). Racial/ethnic sleep disparities in us school-aged children and adolescents: a review of the literature. *Sleep health*, 4(1):68–80.

Hafner, M., Stepanek, M., and Troxel, W. M. (2017). The economic implications of later school start times in the united states. *Sleep Health*, 3(6):451–457.

Haidt, J. (2024). *The anxious generation: How the great rewiring of childhood is causing an epidemic of mental illness*. Random House.

Heissel, J. A. and Norris, S. (2018). Rise and shine the effect of school start times on academic performance from childhood through puberty. *Journal of Human Resources*, 53(4):957–992.

Hinrichs, P. (2011). When the bell tolls: The effects of school starting times on academic achievement. *Education Finance and Policy*, 6(4):486–507.

Huffington, A. (2016). *The sleep revolution: Transforming your life, one night at a time*. Harmony.

Jagnani, M. (2021). Children's sleep and human capital production. *The Review of Economics and Statistics, Forthcoming*.

James, S. A., Erickson, D. J., Lammert, S., and Widome, R. (2023). School start time delays and high school educational outcomes: Evidence from the start/learn study. *Journal of adolescence*, 95(4):751–763.

Jepsen, C. and Rivkin, S. (2009). Class size reduction and student achievement: The potential tradeoff between teacher quality and class size. *Journal of Human Resources*, 44(1):223–250.

Jung, J. and Fenelon, A. (2024). Do later school start times improve adolescents' sleep and substance use? a quasi-experimental study. *Preventive medicine*, 185:108028.

Lauderdale, D. S., Knutson, K. L., Yan, L. L., Liu, K., and Rathouz, P. J. (2008). Self-reported and measured sleep duration: How similar are they? *Epidemiology*, 19(6):838–845.

Lenard, M., Morrill, M. S., and Westall, J. (2020). High school start times and student achievement: Looking beyond test scores. *Economics of Education Review*, 76:101975.

Lim, J. and Dinges, D. F. (2010). A meta-analysis of the impact of short-term sleep deprivation on cognitive variables. *Psychological bulletin*, 136(3):375.

List, J. A. (2022). *The voltage effect: How to make good ideas great and great ideas scale*. Crown Currency.

List, J. A. (2024). Optimally generate policy-based evidence before scaling. *Nature*, 626(7999):491–499.

Lo, J. C., Lee, S. M., Lee, X. K., Sasmita, K., Chee, N. I. Y. N., Tandi, J., Cher, W. S., Gooley, J. J., and Chee, M. W. L. (2018). Sustained benefits of delaying school start time on adolescent sleep and well-being. *Sleep*, 41(6):zsy052.

Lo, J. C., Ong, J. L., Leong, R. L., Gooley, J. J., and Chee, M. W. (2016). Cognitive performance, sleepiness, and mood in partially sleep deprived adolescents: the need for sleep study. *Sleep*, 39(3):687–698.

Luong, P., Lusher, L., and Yasenov, V. (2019). Does schedule irregularity affect productivity? Evidence from random assignment into college classes. *Labour Economics*, 60:115–128.

Lusher, L. and Yasenov, V. (2016). Double-shift schooling and student success: Quasi-experimental evidence from europe. *Economics Letters*, 139:36–39.

Lusher, L. and Yasenov, V. (2018). Gender performance gaps: Quasi-experimental evidence on the role of gender differences in sleep cycles. *Economic Inquiry*, 56(1):252–262.

Oster, E. (2019). Unobservable selection and coefficient stability: Theory and evidence. *Journal of Business & Economic Statistics*, 37(2):187–204.

Owens, J. A., Belon, K., and Moss, P. (2010). Impact of delaying school start time on adolescent sleep, mood, and behavior. *Archives of pediatrics & adolescent medicine*, 164(7):608–614.

Owens, J. A., Dearth-Wesley, T., Herman, A. N., Oakes, J. M., and Whitaker, R. C. (2017). A quasi-experimental study of the impact of school start time changes on adolescent sleep. *Sleep health*, 3(6):437–443.

Pallesen, S., Saxvig, I. W., Molde, H., Sørensen, E., Wilhelmsen-Langeland, A., and Bjorvatn, B. (2011). Brief report: behaviorally induced insufficient sleep syndrome in older adolescents: prevalence and correlates. *Journal of adolescence*, 34(2):391–395.

Paruthi, S., Brooks, L. J., D'Ambrosio, C., Hall, W. A., Kotagal, S., Lloyd, R. M., Malow, B. A., Maski, K., Nichols, C., Quan, S. F., et al. (2016). Consensus statement of the american academy of sleep medicine on the recommended amount of sleep for healthy children: methodology and discussion. *Journal of clinical sleep medicine*, 12(11):1549–1561.

Pope, N. G. (2016). How the time of day affects productivity: Evidence from school schedules. *Review of Economics and Statistics*, 98(1):1–11.

Roenneberg, T., Kuehnle, T., Pramstaller, P. P., Ricken, J., Havel, M., Guth, A., and Merrow, M. (2004). A marker for the end of adolescence. *Current biology*, 14(24):R1038–R1039.

Sadikova, E., Widome, R., Robinson, E., Aris, I. M., and Tiemeier, H. (2024). Delaying high school start times impacts depressed mood among students: Evidence from a natural experiment. *Social Psychiatry and Psychiatric Epidemiology*, 59(11):2073–2082.

Scott, A. J., Webb, T. L., Martyn-St James, M., Rowse, G., and Weich, S. (2021). Improving sleep quality leads to better mental health: A meta-analysis of randomised controlled trials. *Sleep medicine reviews*, 60:101556.

Temkin, D. A., Princiotta, D., Ryberg, R., and Lewin, D. S. (2018). Later start, longer sleep: Implications of middle school start times. *Journal of School Health*, 88(5):370–378.

Twenge, J. M., Hisler, G. C., and Krizan, Z. (2019). Associations between screen time and sleep duration are primarily driven by portable electronic devices: Evidence from a population-based study of us children ages 0–17. *Sleep medicine*, 56:211–218.

Twenge, J. M., Krizan, Z., and Hisler, G. (2017). Decreases in self-reported sleep duration among us adolescents 2009–2015 and association with new media screen time. *Sleep medicine*, 39:47–53.

Wahlstrom, K., Dretzke, B., Gordon, M., Peterson, K., Edwards, K., and Gdula, J. (2014). Examining the impact of later high school start times on the health and academic performance of high school students: A multi-site study. Technical report, Center for Applied Research and Educational Improvement, University of Minnesota.

Wahlstrom, K. L., Davison, M. L., Choi, J., and Ross, J. N. (2001). Minneapolis Public Schools start time study: Executive summary. Technical report, Center for Applied Research and Educational Improvement, University of Minnesota.

Walker, M. (2017). *Why we sleep: Unlocking the power of sleep and dreams*. Simon and Schuster.

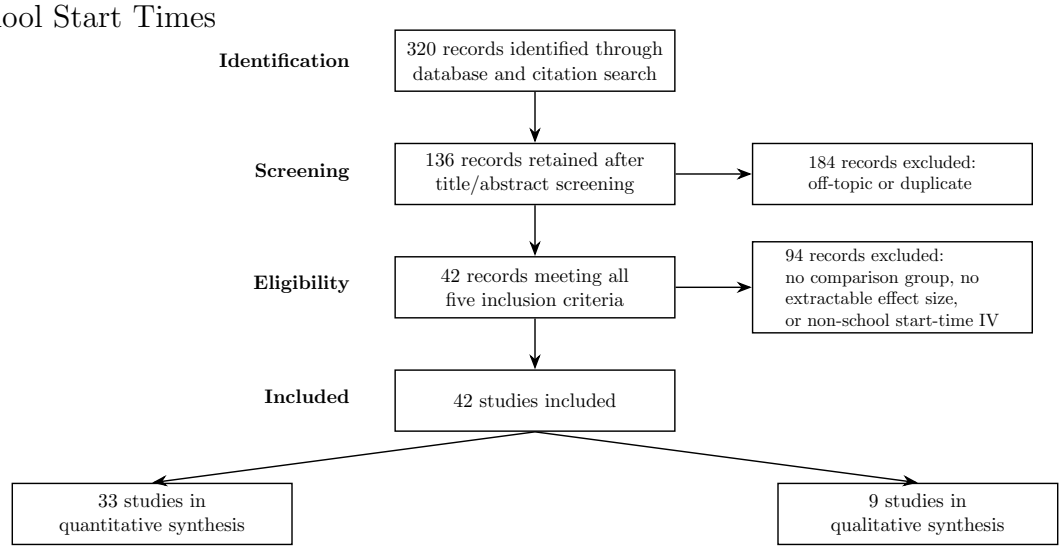
Widome, R., Berger, A. T., Iber, C., Wahlstrom, K., Laska, M. N., Kilian, G., Redline, S., and Erickson, D. J. (2020). Association of delaying school start time with sleep duration, timing, and quality among adolescents. *JAMA pediatrics*, 174(7):697–704.

Yang, D. and Choi, J. (2021). The effect of delaying school start time on adolescents' time use and health: Evidence from a policy change in south korea. *Asian Population Studies*, 17(1). Published online 23 July 2020, doi:10.1080/17441730.2020.1794312.

Yip, T., Wang, Y., Xie, M., Ip, P. S., Fowle, J., and Buckhalt, J. (2022). School start times, sleep, and youth outcomes: a meta-analysis. *Pediatrics*, 149(6):e2021054068.

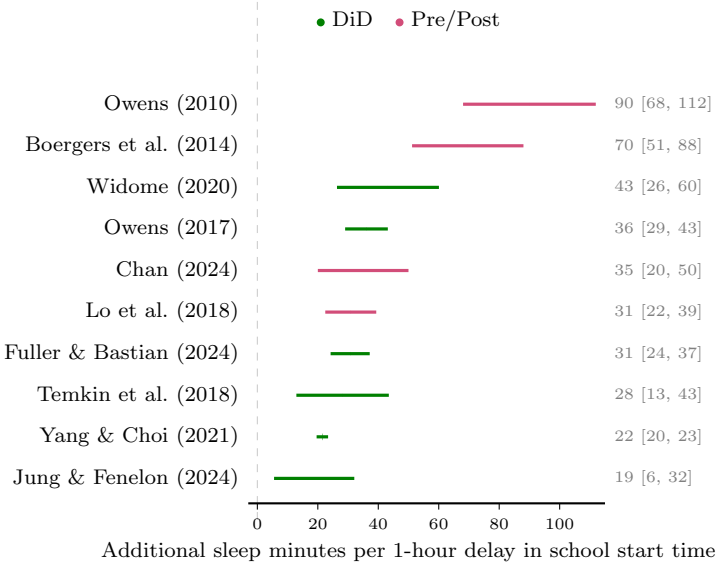
Appendix

Figure A.1: PRISMA Flow Diagram: Systematic Review of Quasi-Experimental Evidence on School Start Times



Notes: PRISMA-style flow diagram for the systematic review described in Section 2. Records were identified through backward and forward citation searches via OpenAlex, anchored on Bowers and Moyer (2017), Yip et al. (2022), Carrell et al. (2011), and Heissel and Norris (2018), and supplemented with keyword searches of the NBER and IZA working paper series. Inclusion criteria: (i) school start time as the independent variable; (ii) sleep, academic, or mental-health outcome; (iii) a design that provides a counterfactual (a comparison-group design for academic and mental-health outcomes; for sleep duration also pre/post designs measuring the same students before and after a change), or a selection-on-observables design with formal omitted-variable bounds; (iv) peer-reviewed publication in English; (v) extractable effect sizes. The 9 qualitative studies report effects without parametric standard errors and so do not enter the inverse-variance pooled estimates reported in the footnote of Section 2.

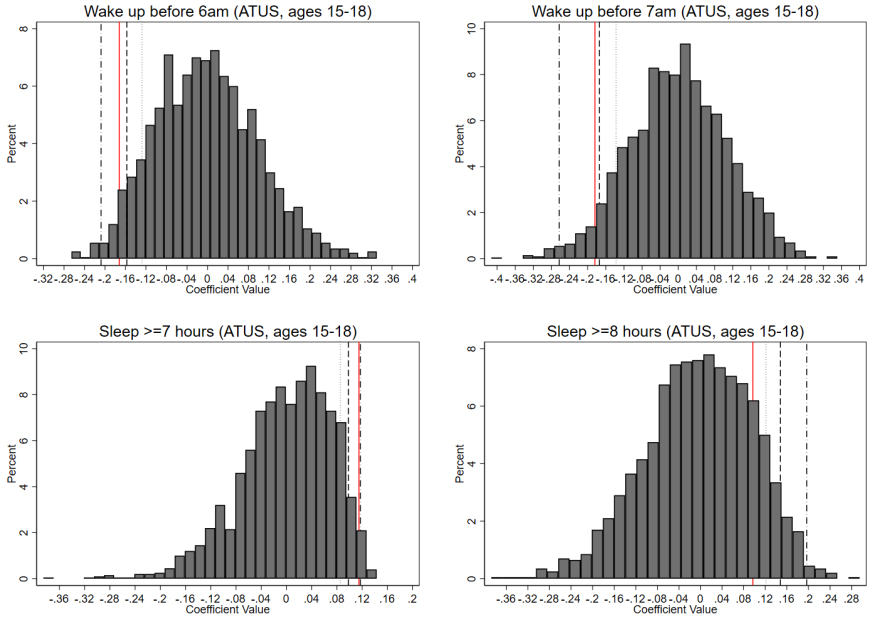
Figure A.2: Meta-Analytic Forest Plot: Sleep Duration Effects of Later School Start Times



Notes: Each point represents a study-level estimate of additional sleep minutes per one-hour delay in school start time, with 95% confidence interval. Marker size is proportional to inverse-variance weight. Colors indicate identification strategy. The figure displays the full distribution of effects rather than an inverse-variance-pooled point estimate (see the footnote in Section 2.2 of the text for a single rough-benchmark pooled value computed for comparability with prior meta-analyses). Chan et al. (2018) and Chan et al. (2024) report on the same Hong Kong residential boarding-school intervention (7:30 a.m. $\rightarrow$ 8:30 a.m. SST change implemented September 2017) with overlapping Cohort 1 baseline samples; only the longer-term Chan (2024) follow-up is shown to avoid double-counting.

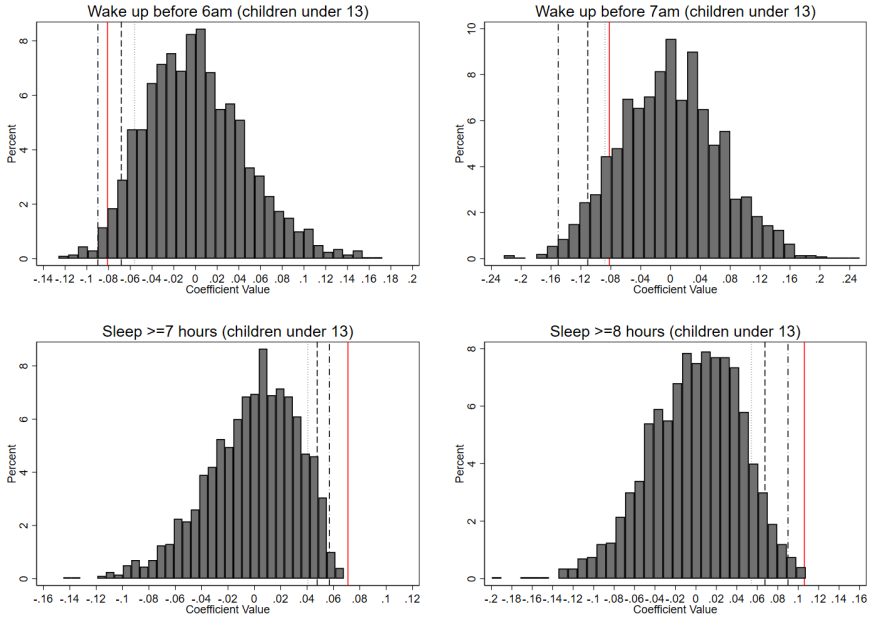
Study-specific disclosures. Yang & Choi (2021): the plotted point is the weekday-only sleep gain (raw +16.1 min, rescaled to +22 min/hr); the same source reports a partially offsetting weekend decline of −7.6 min/night, so the weekly-average effect is smaller ($\approx$ +9.3 min/day before rescaling). *Widome et al. (2020):* the plotted point is the school-night-only effect at follow-up 1 (raw +41 min); the same source reports concurrent weekend-night declines of −24 min (FU1) and −34 min (FU2). The per-hour rescale uses the midpoint of the two treated schools' delays (50 and 65 min). *Lo et al. (2018):* the plotted point is one-month time-in-bed (TIB); the same source reports a smaller +10 min total-sleep-time effect at nine-month follow-up. *Jung & Fenelon (2024):* the plotted point is the year-0 effect from the event-study specification; the same source documents that this gain fades to +7 min within one year as bedtimes shift later (see Table A.23 footnote *h*).

Figure A.3: Permutation Tests, Adolescents Aged 15–18, ATUS



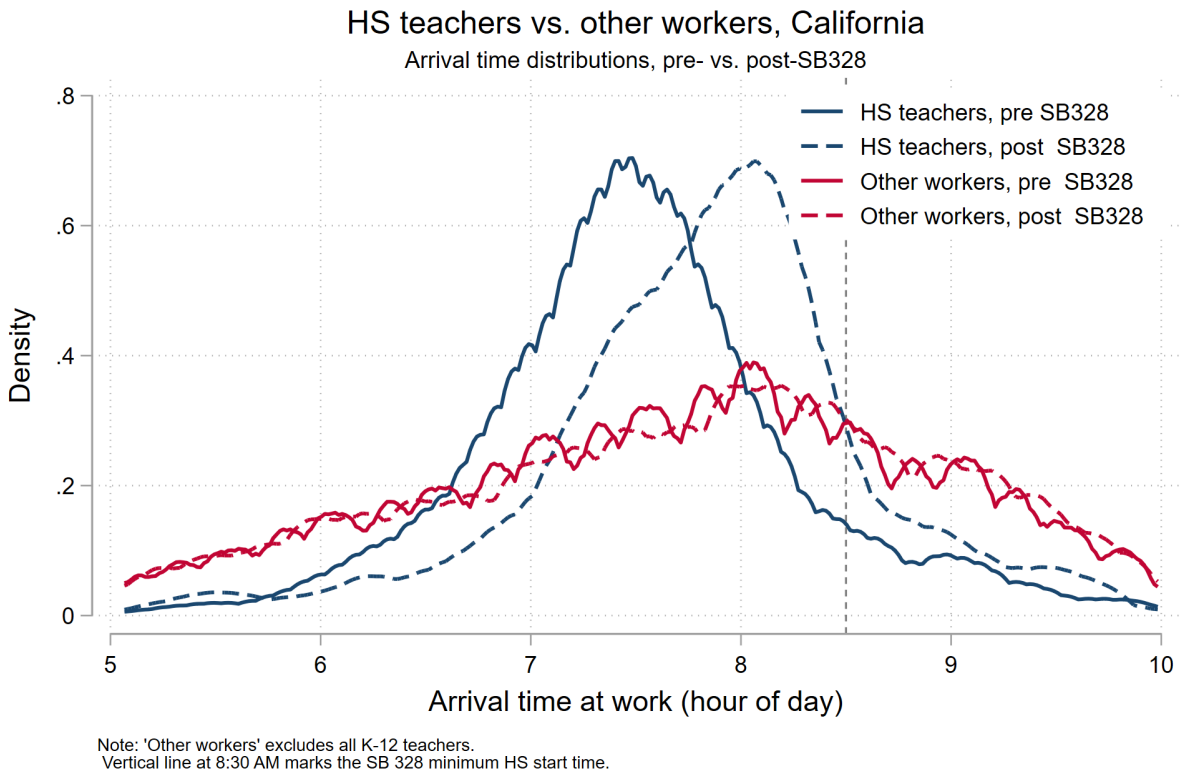
Notes: The figure illustrates the results of permutation tests for various outcomes from ATUS respondents aged 15-18, where the policy adoption year was randomly reassigned to other states to create placebo scenarios. The red vertical line represents the observed treatment effect, while the dashed lines indicate the 1%, 5%, and 10% significance thresholds derived from the placebo distribution. By comparing the observed effect to the distribution of placebo effects, the test evaluates whether the observed policy effect is statistically significant or could plausibly arise by random chance. Standard errors are clustered at the state level.

Figure A.4: Permutation Tests, Children Under 13, ATUS



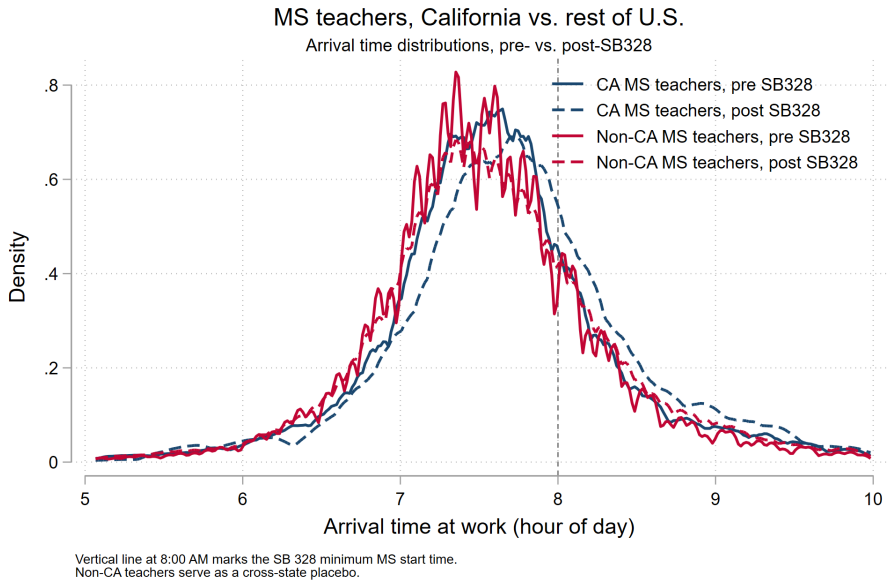
Notes: The figure illustrates the results of permutation tests for various outcomes from ATUS respondents under the age of 13, where the policy adoption year was randomly reassigned to other states to create placebo scenarios. The red vertical line represents the observed treatment effect, while the dashed lines indicate the 1%, 5%, and 10% significance thresholds derived from the placebo distribution. By comparing the observed effect to the distribution of placebo effects, the test evaluates whether the observed policy effect is statistically significant or could plausibly arise by random chance. Standard errors are clustered at the state level.

Figure A.5: Validating the Teacher Arrival-Time Proxy: ACS Arrival Times, California High School Teachers vs. Other Workers, Pre- vs. Post-SB 328



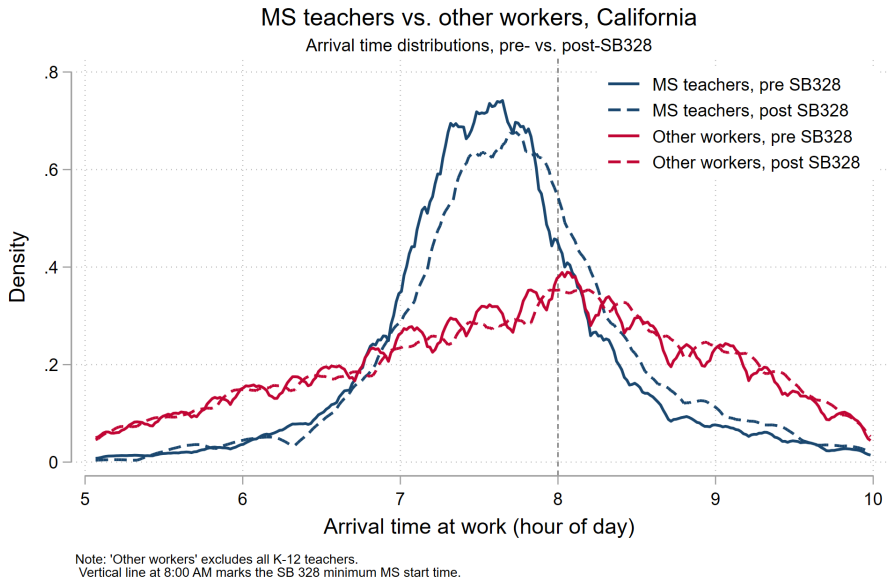
Notes: The figure presents kernel density estimates of self-reported arrival times at work among California workers, using data from the American Community Survey (ACS) for the years 2015–2019 and 2023. The pre-SB 328 period uses the 5-year ACS sample (2015–2019); the post-SB 328 period employs the 1-year 2023 ACS sample, the first full wave after the policy took effect on July 1, 2022. High school teachers are identified using IPUMS 2018 occupation code 2320. "Other workers" excludes all K–12 teachers (IPUMS codes 2310 and 2320). The sample is restricted to California residents with reported arrival times between 5:00 a.m. and 10:00 a.m. Observations are weighted by ACS person weights. The vertical dashed line at 8:30 a.m. marks the minimum high school start time mandated by SB 328. The figure shows a marked rightward shift in the distribution of high school teacher arrival times following the policy, with the post-period mass concentrating immediately below the 8:30 a.m. threshold, while the distribution of arrival times for non-teacher workers is essentially unchanged between the pre- and post-periods. This pattern supports the use of pre-policy teacher arrival times as a proxy for pre-policy school start times in the within-state analysis.

Figure A.6: Teacher Arrival Time Proxy, Middle School Teachers: ACS Arrival Times, California vs. Other States, Pre- vs. Post-SB 328



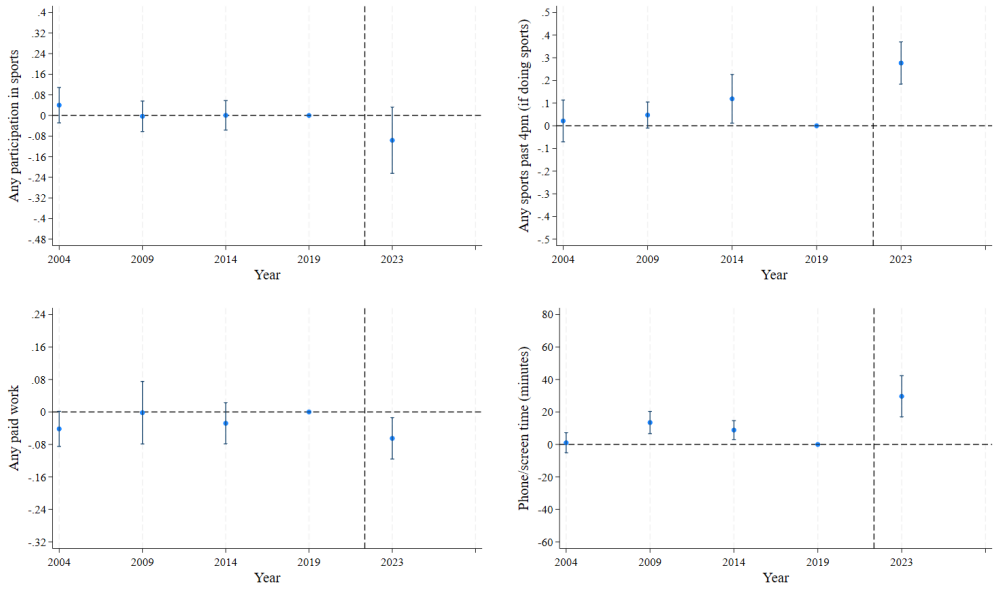
Notes: The figure presents kernel density estimates of self-reported arrival times at work among middle school teachers, using data from the American Community Survey (ACS) for the years 2015–2019 and 2023. The pre-SB 328 period uses the 5-year ACS sample (2015–2019); the post-SB 328 period employs the 1-year 2023 ACS sample, the first full wave after the policy took effect on July 1, 2022. Note that the underlying ACS variable OCC2010 (value 2310) does not actually distinguish between elementary and middle school teachers. Hence, the graph captures the combined elementary school/middle school arrival time distributions. The sample is restricted to respondents with reported arrival times between 5:00 a.m. and 10:00 a.m. Observations are weighted by ACS person weights. The vertical dashed line at 8:00 a.m. marks the minimum middle school start time mandated by SB 328. The figure shows a modest rightward shift in the distribution of arrival times for California elementary/middle school teachers following the policy, while the corresponding distribution for elementary/middle school teachers in other states remains largely unchanged. The smaller magnitude of the shift reflects both the fact that pre-policy arrival times for middle school teachers were already concentrated closer to the 8:00 a.m. threshold, as well as the contamination of the sample with elementary school teachers.

Figure A.7: Teacher Arrival Time Proxy, Middle School Teachers: ACS Arrival Times, California Middle School Teachers vs. Other Workers, Pre- vs. Post-SB 328



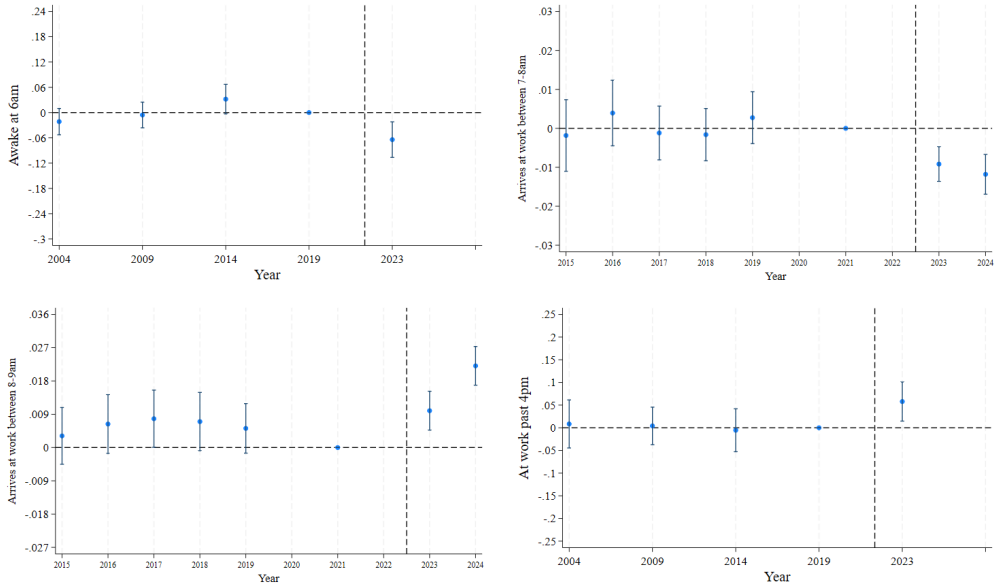
Notes: The figure presents kernel density estimates of self-reported arrival times at work among California workers, using data from the American Community Survey (ACS) for the years 2015–2019 and 2023. The pre-SB 328 period uses the 5-year ACS sample (2015–2019); the post-SB 328 period employs the 1-year 2023 ACS sample, the first full wave after the policy took effect on July 1, 2022. Note that the underlying ACS variable OCC2010 (value 2310) does not actually distinguish between elementary and middle school teachers. Hence, the graph captures the combined elementary school/middle school arrival time distributions. "Other workers" excludes all K–12 teachers (IPUMS codes 2310 and 2320). The sample is restricted to California residents with reported arrival times between 5:00 a.m. and 10:00 a.m. Observations are weighted by ACS person weights. The vertical dashed line at 8:00 a.m. marks the minimum middle school start time mandated by SB 328. The post-period shift is directionally consistent with the SB 328 mandate but smaller in magnitude than for high school teachers, reflecting both the fact that pre-policy MS teacher arrivals were already concentrated before the 8:00 a.m. threshold, as well as the contamination of the sample with elementary school teachers.

Figure A.8: SB 328 and Adolescents' Non-School Time Use, Evidence from ATUS (2003-24)



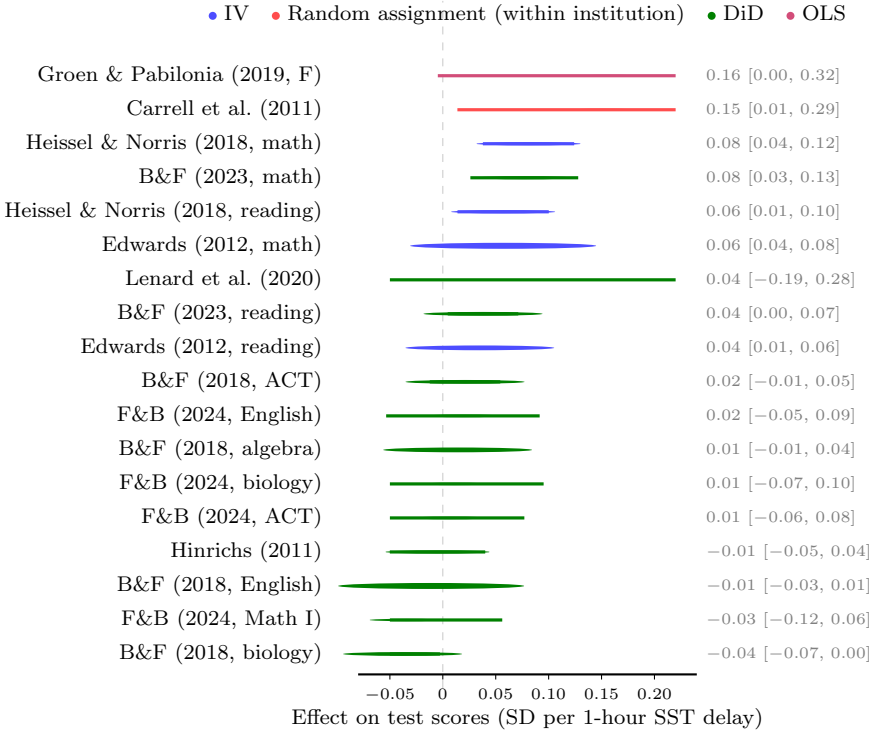
Notes: The figure illustrates the differences in non-school time use among adolescents, ages 15-18, in California and other states before and after the adoption of SB 328. Data come from the 2003–2024 American Time Use Survey (ATUS). 2020 is omitted due to incomplete data collection during the COVID-19 pandemic, and observations are restricted to weekday diaries. Observations are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to an individual's state of residence, survey year, age, sex, and race. The survey years were pooled as follows: (1) 2004 combines 2003-2006; (2) 2009 combines 2007-2011; (3) 2014 combines 2012-2016; (4) 2019 combines 2017-2021 and January–August 2022 (excluding 2020); (5) the post-period point, labelled 2023 in the figure, combines September 2022–2024. Top-left graph depicts an indicator for any participation in sports or exercise on the diary day; top-right graph shows an indicator for sports or exercise after 4pm (conditional on any sports or exercise that day); bottom-left graph depicts an indicator for any paid work; and bottom-right graph shows the number of minutes spent on telephone calls, email, messaging, and social media. Standard errors are clustered at the state level. Coefficient estimates are provided together with the 95% confidence intervals.

Figure A.9: SB 328 and Parental Work-Related Time Use, ATUS (2003-24), ACS (2015-24)



Notes: The figure illustrates the differences in work-related time use among parents (ages 25-64) of children (ages 11-18), who worked away from home, in California and other states before and after the adoption of SB 328. Data come from the 2003–2024 American Time Use Survey (ATUS): top-left and bottom-right graphs; and the 2015–2024 American Community Survey (ACS): top-right and bottom-left graphs. ATUS sample excludes 2020 due to incomplete data collection during the COVID-19 pandemic. ACS sample excludes 2020 (experimental COVID-19 sample) and 2022 (interview month is unavailable). Observations in ATUS analyses are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to individual's state of residence, survey year, age, sex, and race. In ATUS analyses, the survey years were pooled as follows: (1) 2004 combines 2003-2006; (2) 2009 combines 2007-2011; (3) 2014 combines 2012-2016; (4) 2019 combines 2017-2021 and January–August 2022 (excluding 2020); (5) the post-period point, labelled 2023 in the figure, combines September 2022–2024. Standard errors are clustered at the state level. Coefficient estimates are provided together with the 95% confidence intervals.

Figure A.10: Meta-Analytic Forest Plot: Academic Effects of Later School Start Times



Notes: Each point is a study-level effect estimate placed on a common metric: every study's coefficient has been expressed in outcome standard-deviation (SD) units and linearly rescaled to a one-hour delay in school start time, so estimates are comparable across studies despite differing original outcome scales and treatment doses. Bars are 95% confidence intervals; marker area is proportional to inverse-variance weight and colors indicate identification strategy. Confidence intervals extending beyond the axis are clipped at the boundary, with the full interval shown at right. We display the full distribution of estimates rather than a pooled point estimate; a single rough-benchmark pooled value is reported in the footnote in Section 2.2 for comparability with prior meta-analyses. James et al. (2023) is omitted because its outcome is GPA points, not SD-normalized test scores. For Bastian & Fuller (2023) the school-FE specification is shown (the student-FE alternative is smaller but similar). Several rows rest on documented rescaling or interpretive assumptions—Fuller & Bastian (2024) ACT (SD imputed from B&F 2018), Carrell et al. (2011) (contrast SE is an upper bound), Heissel & Norris (2018) (morning-sunlight IV read as a start-time delay), and Lenard et al. (2020) (advance rescaled to a delay)—each detailed in Appendix Table A.23.

Table A.1: Meta-Analytic Summary: Quasi-Experimental Evidence on School Start Times

Study	Design	Scope	Sleep (min./hr)	Academic (SD/hr)	MH
Panel A: Sleep duration					
Owens et al. (2017)	Pre/post QE	1 district	+36.1 (3.6)		
Widome et al. (2020)	DiD + actigraphy	1 metro	+43 (8.6)		
Lo et al. (2018)	Pre/post QE	1 school	+30.9 (4.3)		+
Jung & Fenelon (2024)	DiD (mandate)	Multi-state	+19 (6.8)		
Panel B: Academic achievement					
Hinrichs (2011)	Cross-district DiD	Multi-district		0.00 (0.02)	
Edwards (2012)	Transportation IV	1 district		+0.06 (0.01)	
Carrell et al. (2011)	Random assignment (within institution)	1 academy		+0.15 (0.07)[a]	
Heissel & Norris (2018)	Time-zone IV	1 state		+0.08 (0.02)[b]	
Groen & Pabilonia (2019)	OLS + Oster bounds	National		+0.16 (0.08)[c]	
Bastian & Fuller (2023)	Panel DiD	1 state		+0.08 (0.03)[g]	
James et al. (2023)	DiD + actigraphy	1 metro		+0.17 GPA pts (0.03)[d]	
Panel C: Mental health					
Chan et al. (2024)	Pre/post boarding	1 school	+35 (7.7)		+[e]
Sadikova et al. (2024)	IV (SST→sleep)	1 metro			+[f]

Notes: Sleep effects in additional minutes per one-hour delay in school start time (Panel A entries normalized uniformly to per-hour scale; raw source treatments: Owens 2017 50 min, Widome 2020 57 min, Lo 2018 45 min). Academic effects in SD units per one-hour delay. Standard errors in parentheses. MH = + indicates improvement in mental health outcomes. Panels A–C display representative studies; a single rough-benchmark random-effects pooled estimate is reported in the footnote in Section 2.2 of the text for comparability with prior meta-analyses. See Appendix Table A.23 for the principal studies discussed in the text; Figure A.1 reports the PRISMA flow.

[a]Per-hour-equivalent of the 7:00 vs. 7:50 a.m. first-period contrast at USAFA, rescaled from the source paper's 50-min contrast. [b]Adolescent (pubescent) subsample; near zero for younger children. [c]Female subsample; OLS with Oster (2019) bounds (0.16–0.28 SD); males null. [d]Effect reported in GPA points (not SD units); excluded from the pooled academic estimate. GPA at two-year follow-up; attendance improves immediately. [e]Significant reductions in depression, anxiety, and insomnia (d = 0.29–0.67); the d = 0.72 figure in Chan et al. (2024) refers to total sleep time, not mental health. [f]IV-estimated: a policy-induced 30-min sleep gain yields −0.78 depression points (SE 0.27) and −1.36 fatigue points (SE 0.38). [g]School-fixed-effect specification shown; the student-fixed-effect alternative reports math +0.039 (SE 0.012) and reading +0.026 (SE 0.009), both reported by the authors as robust.

Table A.2: Summary Statistics - YRBS (2007–2023)

	Mean	Std. Dev.	N
Covariates			
Age categories	5.00	1.25	139,021
Female	0.50	0.50	139,703
Asian	0.04	0.20	139,703
Black	0.16	0.36	139,703
Hispanic	0.25	0.44	139,703
White	0.45	0.50	139,703
Outcomes			
Sleep $\geq$ 6 hours	0.79	0.41	120,919
Sleep $\geq$ 7 hours	0.56	0.50	120,919
Sleep $\geq$ 8 hours	0.27	0.44	120,919
Concentration difficulties	0.39	0.49	49,362
Sadness/hopelessness feelings	0.32	0.47	137,950
Suicidal intentions	0.18	0.39	122,451

Notes: Summary statistics reported for the full YRBS sample (2007–2023), including main covariates and mental health outcomes. Age categories are coded from 1 ($\leq$12 years) to 7 ($\geq$18 years). Outcomes are based on self-reported responses.

Table A.3: Summary Statistics – ATUS (2003–2024, Excluding 2020)

Panel A: Children aged 15–18	Mean	Std. Dev.	N
Awake by 6am	0.283	0.450	5,785
Awake by 7am	0.591	0.492	5,785
Awake by 8am	0.714	0.452	5,785
Bedtime after 9pm	0.694	0.461	5,785
Bedtime after 10pm	0.354	0.478	5,785
Bedtime after 11pm	0.226	0.418	5,785
Sleep hours	9.120	2.346	5,785
Sleep at least 6 hours	0.942	0.234	5,785
Sleep at least 7 hours	0.862	0.345	5,785
Sleep at least 8 hours	0.703	0.457	5,785
Panel B: Children under 13	Mean	Std. Dev.	N
Awake by 6am	0.128	0.334	11,825
Awake by 7am	0.432	0.495	11,825
Awake by 8am	0.758	0.428	11,825
Bedtime after 9pm	0.322	0.467	11,825
Bedtime after 10pm	0.135	0.341	11,825
Bedtime after 11pm	0.066	0.248	11,825
Sleep hours	9.657	1.997	11,825
Sleep at least 6 hours	0.965	0.184	11,825
Sleep at least 7 hours	0.938	0.241	11,825
Sleep at least 8 hours	0.876	0.330	11,825

Notes: Summary statistics are reported for the ATUS sample for key sleep-related outcomes. Panel A presents results for children aged 15–18. Panel B presents results for children under 13, based on parental reports and linked child-level sleep data. Means and standard deviations are rounded to three decimal places.

Table A.4: SB 328 and Teacher Arrival Time in California (ACS Evidence)

	(1) *High School Teachers*	(2) *Middle & Elemen. School Teachers*
	Arrival time	Arrival time
Policy (SB 328)	18.93**	12.30***
	(8.32)	(3.50)
Observations	8,387	25,808
R-squared	0.085	0.027

Notes: The table reports difference-in-differences estimates of the effect of SB 328 on teacher arrival-at-work time, using individual-level data from the American Community Survey (ACS) for the years 2015-2024 (2020 wave is dropped due to COVID-19; 2022 wave is dropped since the month of interview is not reported in ACS). Column (1) restricts to high school teachers (IPUMS 2018 occupation code 2320), while column (2) restricts to elementary & middle school teachers (IPUMS 2018 occupation code 2310). Note that the occupation code 2310 does not distinguish between elementary and middle school teachers. Hence, col. (2) captures a combined effect on elementary school and middle school teachers. The dependent variable is the arrival time at work (in minutes from midnight). Regression in col. (1) uses teacher-arrival treatment intensity defined for high school teachers (IPUMS occupation 2320; bottom three quartiles vs. top quartile plus rural counties). Regression in col. (2) uses teacher-arrival treatment intensity defined for elementary/middle school teachers (IPUMS occupation 2310; bottom two quartiles vs. top two quartiles plus rural counties). All specifications include county-of-work and year fixed effects, individual-level controls (sex, marital status, number of children, race/ethnicity, education), and ACS person weights. Standard errors in parentheses, clustered at the county-of-work level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.5: CA SB 328 and High School Senior SAT Scores

	(1) SAT	(2) Math	(3) Evidence-based reading and writing score
CA SB 328	24.987***	10.673***	13.313***
	(4.793)	(2.288)	(2.535)
Observations	204	204	204
Mean of dep. var.	1,112	549.7	562.3
Std. dev. of dep. var.	95.05	49.49	45.87

Notes: Data on SAT scores were published by the College Board. Data include 50 US states and Washington, D.C., for the years 2018, 2021, 2022, and 2023. All estimates include state and year fixed effects. Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.

Table A.6: ACS Placebo Test (Adolescents, Ages 11-18) – Vision and Hearing Difficulties

	(1)	(2)	(3)	(4)
	State Level		California	
	Vision	Hearing	Vision	Hearing
Policy (SB 328)	0.0005	0.0004	-0.0013	0.0001
	(0.0005)	(0.0004)	(0.0021)	(0.0010)
Observations	2,551,938	2,551,938	306,465	306,465
R-squared	0.001	0.000	0.002	0.001
Mean of dep. var.	0.011	0.006	0.010	0.006
Std. dev. of dep. var.	0.104	0.078	0.100	0.076

Notes: This table reports placebo estimates of the effect of the ACS-based teacher arrival-time treatment on outcomes that should be unaffected by the schedule mandate, namely DIFFEYE (whether respondent is blind or has a serious vision difficulty) and DIFFHEAR (whether respondent is deaf or has a serious hearing difficulty). Data comes from the 2015-2024 American Community Survey (ACS). 2020 wave is dropped due to COVID-19; 2022 wave is dropped since the month of interview is not reported in ACS. Columns (1)–(2) use the full U.S. sample of teens aged 11–18; columns (3)–(4) use within-California variation with a sample of teens aged 11–18. Regressions in cols. (1)–(2) use the California x post-SB 328 interaction term. Regressions in cols. (3)–(4) use teacher-arrival treatment intensity defined for high school teachers (IPUMS occupation 2320; bottom three quartiles vs. top quartile plus rural counties). All specifications include state/county and year fixed effects, individual-level controls (age, sex, race/ethnicity), and ACS person weights. Analyses in cols. (1)–(2) further employ state-level time-varying socioeconomic covariates. Standard errors in parentheses, clustered at the state level (cols. 1–2) or county level (cols. 3–4). *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.7: CDE Placebo Test on School Characteristics

	(1)	(2)	(3)	(4)	(5)
	Total Enrollment	White Share	Hispanic Share	Female Share	Socioecon. Disadvantaged Share
Policy (SB 328)	-13.92	0.283	-0.383	0.079	0.693
	(14.46)	(0.516)	(0.748)	(0.083)	(1.189)
Observations	63,450	64,024	76,551	63,450	63,450
R-squared	0.024	0.345	0.192	0.002	0.216
Mean of dep. var.	803.2	25.78	56.02	48.21	61.31
Std. dev. of dep. var.	703.7	21.21	27.58	4.514	26.05

Notes: This table reports placebo regressions using school-level characteristics from CDE data for the academic years 2014–15 through 2024–25. The dependent variables include total enrollment and demographic shares (race/ethnicity, sex, and socioeconomically disadvantaged students). All regressions use a combined HS+ES/MS teacher-arrival treatment intensity (IPUMS occupations 2320 and 2310; bottom three quartiles vs. top quartile plus rural counties). All specifications include school and year fixed effects, as well as time-varying demographic covariates unrelated to the outcome tested. Standard errors in parentheses, clustered at the county level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.8: Robustness Check for SB 328, Wake-up Times and Sleep, ATUS All Households with Children Under 13

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Awake by			Bedtime			Sleep			
	6am	7am	8am	after 9pm	after 10pm	after 11pm	sleep hours	sleep≥6	sleep≥7	sleep≥8
CA SB 328	-0.001	0.016	0.017	0.069***	0.027*	0.024***	-0.168***	-0.003	-0.001	0.000
	(0.007)	(0.027)	(0.021)	(0.015)	(0.016)	(0.008)	(0.049)	(0.003)	(0.014)	(0.010)
Observations	40,289	40,289	40,289	40,289	40,289	40,289	40,289	40,289	40,289	40,289
R-squared	0.014	0.047	0.078	0.065	0.038	0.024	0.035	0.012	0.014	0.020
Mean of dep. var.	0.145	0.511	0.814	0.373	0.154	0.0705	9.287	0.960	0.922	0.839
Std. dev. of dep. var.	0.352	0.500	0.389	0.483	0.361	0.256	1.974	0.197	0.268	0.367

Notes: Data come from the 2003–2024 American Time Use Survey (ATUS). 2020 is omitted due to incomplete data collection during the COVID-19 pandemic, and observations are restricted to weekday diaries. Observations are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to the individual's state of residence, survey year, age, sex, and race. The estimation sample comprises children under 13 (including those from households with more than one child). All specifications control for state, year, and month fixed effects, individual-level covariates (race/ethnicity of the parent), and state-level characteristics (population, % households with children, marital status and education-related variables, unemployment rate, median household income, and child poverty rate). Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.9: Robustness Check for SB 328, Wake-up Times and Sleep, ATUS Without Outliers

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Awake by			Bedtime			Sleep			
	6am	7am	8am	after 9pm	after 10pm	after 11pm	sleep hours	sleep≥6	sleep≥7	sleep≥8
Panel A: Children aged 15–18										
CA SB 328	-0.161***	-0.174***	-0.086**	-0.007	0.080**	0.044	0.316	0.028*	0.103***	0.087**
	(0.041)	(0.028)	(0.033)	(0.034)	(0.037)	(0.039)	(0.190)	(0.016)	(0.018)	(0.042)
Observations	5,681	5,681	5,681	5,681	5,681	5,681	5,681	5,681	5,681	5,681
R-squared	0.070	0.154	0.149	0.055	0.057	0.049	0.065	0.020	0.028	0.042
Mean of dep. var.	0.285	0.598	0.723	0.700	0.356	0.225	9.024	0.943	0.863	0.700
Std. dev. of dep. var.	0.451	0.490	0.448	0.458	0.479	0.418	2.120	0.231	0.344	0.458
Panel B: Children under 13 (Single-Child Households)										
CA SB 328	-0.069***	-0.071***	-0.021	0.106***	-0.009	0.022*	0.112	0.031**	0.059***	0.095***
	(0.024)	(0.026)	(0.026)	(0.031)	(0.016)	(0.013)	(0.116)	(0.012)	(0.012)	(0.020)
Observations	11,673	11,673	11,673	11,673	11,673	11,673	11,673	11,673	11,673	11,673
R-squared	0.014	0.025	0.037	0.041	0.032	0.020	0.028	0.013	0.017	0.023
Mean of dep. var.	0.122	0.429	0.758	0.319	0.129	0.0591	9.686	0.972	0.945	0.882
Std. dev. of dep. var.	0.327	0.495	0.428	0.466	0.335	0.236	1.699	0.164	0.227	0.322

Notes: Data come from the 2003–2024 American Time Use Survey (ATUS). 2020 is omitted due to incomplete data collection during the COVID-19 pandemic, and observations are restricted to weekday diaries. Observations with sleep hours below 2 and above 15 are dropped as outliers. Observations are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to the individual's state of residence, survey year, age, sex, and race. Panel A restricts the analysis sample to children aged 15–18. Panel B restricts the analysis sample to children under the age of 13. All specifications control for state, year, and month fixed effects. Panel A regressions further control for individual-level covariates (age, sex, race/ethnicity) and state-level characteristics (% households with children, marital status and education-related variables, unemployment rate, and child poverty rate). Panel B regressions further control for individual-level covariates (race/ethnicity of the parent) and state-level characteristics (population, % households with children, marital status and education-related variables, unemployment rate, median household income, and child poverty rate). Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.10: Robustness Check for SB 328, Wake-up Times and Sleep, ATUS Metropolitan Sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Awake by			Bedtime			Sleep			
	6am	7am	8am	after 9pm	after 10pm	after 11pm	sleep hours	sleep≥6	sleep≥7	sleep≥8
Panel A: Children aged 15–18										
CA SB 328	-0.197***	-0.184***	-0.087**	-0.003	0.059	0.028	0.514**	0.046*	0.129***	0.106**
	(0.046)	(0.025)	(0.035)	(0.035)	(0.046)	(0.055)	(0.206)	(0.027)	(0.029)	(0.049)
Observations	5,730	5,730	5,730	5,730	5,730	5,730	5,730	5,730	5,730	5,730
R-squared	0.071	0.154	0.149	0.055	0.056	0.048	0.059	0.020	0.028	0.042
Mean of dep. var.	0.283	0.590	0.713	0.694	0.354	0.226	9.119	0.942	0.862	0.702
Std. dev. of dep. var.	0.451	0.492	0.452	0.461	0.478	0.418	2.345	0.234	0.345	0.457
Panel B: Children under 13 (Single-Child Households)										
CA SB 328	-0.084***	-0.086***	-0.028	0.101***	-0.013	0.015	0.380***	0.043***	0.073***	0.111***
	(0.025)	(0.028)	(0.027)	(0.031)	(0.018)	(0.014)	(0.137)	(0.012)	(0.013)	(0.021)
Observations	11,732	11,732	11,732	11,732	11,732	11,732	11,732	11,732	11,732	11,732
R-squared	0.015	0.025	0.037	0.042	0.034	0.023	0.029	0.016	0.020	0.026
Mean of dep. var.	0.128	0.432	0.759	0.323	0.135	0.0656	9.655	0.965	0.938	0.876
Std. dev. of dep. var.	0.334	0.495	0.428	0.468	0.342	0.248	1.996	0.184	0.241	0.330

Notes: Data come from the 2003–2024 American Time Use Survey (ATUS). 2020 is omitted due to incomplete data collection during the COVID-19 pandemic, and observations are restricted to weekday diaries. The ATUS contains no rural California diaries (IPUMS metro code 4) after September 2022, so a California × post-SB 328 × non-rural triple interaction is collinear with California × post-SB 328 and cannot be estimated. The table therefore reports the California × post-SB 328 coefficient among respondents in metropolitan areas (metro codes 1–3), controlling for a non-rural indicator and its interactions with California and with post-SB 328. Panel A restricts the analysis sample to children aged 15–18, weighted by SEER population weights based on state of residence, survey year, age, sex, and race; specifications include fixed effects for state, year, sex, age, month, SEER race, and Hispanic origin and state-level controls (% households with children, marital status and education-related variables, the share of 16–19 year-olds out of school, the unemployment rate, and the child poverty rate). Panel B restricts to households with exactly one child under 13, weighted by SEER state-year population aged 0–12; specifications include fixed effects for state, year, month, SEER race, and Hispanic origin and state-level controls (total population, % households with children, share divorced, share with at least a high school degree, the unemployment rate, median household income, and the child poverty rate). Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.11: Robustness Check for SB 328, Wake-up Times and Sleep, ATUS Excluding Summer and Holidays

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Awake by			Bedtime			Sleep			
	6am	7am	8am	after 9pm	after 10pm	after 11pm	sleep hours	sleep≥6	sleep≥7	sleep≥8
Panel A: Children aged 15–18										
CA SB 328	-0.172***	-0.172***	-0.100***	-0.042	0.065	0.035	0.427*	0.052*	0.119***	0.070
	(0.036)	(0.028)	(0.029)	(0.040)	(0.047)	(0.051)	(0.246)	(0.028)	(0.028)	(0.043)
Observations	4,856	4,856	4,856	4,856	4,856	4,856	4,856	4,856	4,856	4,856
R-squared	0.066	0.134	0.110	0.052	0.056	0.049	0.054	0.021	0.028	0.041
Mean of dep. var.	0.305	0.642	0.768	0.679	0.332	0.210	8.972	0.939	0.854	0.686
Std. dev. of dep. var.	0.460	0.479	0.422	0.467	0.471	0.407	2.282	0.239	0.353	0.464
Panel B: Children under 13 (Single-Child Households)										
CA SB 328	-0.078**	-0.056*	0.047**	0.106***	-0.008	-0.000	0.245*	0.045***	0.077***	0.103***
	(0.030)	(0.029)	(0.023)	(0.028)	(0.015)	(0.015)	(0.134)	(0.012)	(0.015)	(0.019)
Observations	9,894	9,894	9,894	9,894	9,894	9,894	9,894	9,894	9,894	9,894
R-squared	0.017	0.025	0.028	0.032	0.030	0.022	0.027	0.015	0.018	0.023
Mean of dep. var.	0.127	0.443	0.779	0.305	0.124	0.0611	9.671	0.966	0.943	0.882
Std. dev. of dep. var.	0.333	0.497	0.415	0.461	0.330	0.240	1.956	0.180	0.232	0.323

Notes: Data come from the 2003–2024 American Time Use Survey (ATUS). 2020 is omitted due to incomplete data collection during the COVID-19 pandemic, and observations are restricted to weekday diaries. Observations are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to the individual's state of residence, survey year, age, sex, and race. Panel A restricts the analysis sample to children aged 15–18. Panel B restricts the sample to households with exactly one child under the age of 13. Panel A specifications include fixed effects for state, year, sex, age, month, SEER race, and Hispanic origin and control for state-level characteristics (% households with children, marital status and education-related variables, the share of 16–19 year-olds out of school, the unemployment rate, and the child poverty rate). Panel B specifications include fixed effects for state, year, month, SEER race, and Hispanic origin and control for state-level characteristics (total population, % households with children, share divorced, share with at least a high school degree, the unemployment rate, median household income, and the child poverty rate). Both panels further exclude observations from the summer and all other holidays. Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.12: Robustness Check for SB 328, Wake-up Times and Sleep, ATUS With Academic Year Fixed Effect

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	Awake by			Bedtime			Sleep			
	6am	7am	8am	after 9pm	after 10pm	after 11pm	sleep hours	sleep≥6	sleep≥7	sleep≥8
Panel A: Children aged 15–18										
CA SB 328	-0.156***	-0.205***	-0.087**	-0.060	0.063	0.020	0.652**	0.045	0.120***	0.116**
	(0.051)	(0.031)	(0.034)	(0.045)	(0.043)	(0.049)	(0.276)	(0.033)	(0.029)	(0.056)
Observations	5,784	5,784	5,784	5,784	5,784	5,784	5,784	5,784	5,784	5,784
R-squared	0.081	0.159	0.155	0.059	0.065	0.063	0.065	0.024	0.034	0.051
Mean of dep. var.	0.283	0.591	0.714	0.693	0.354	0.225	9.120	0.942	0.862	0.703
Std. dev. of dep. var.	0.450	0.492	0.452	0.461	0.478	0.418	2.346	0.234	0.345	0.457
Panel B: Children under 13 (Single-Child Households)										
CA SB 328	-0.083***	-0.091***	-0.025	0.104***	-0.012	0.018	0.370***	0.044***	0.073***	0.108***
	(0.023)	(0.027)	(0.025)	(0.028)	(0.015)	(0.014)	(0.130)	(0.012)	(0.013)	(0.019)
Observations	11,822	11,822	11,822	11,822	11,822	11,822	11,822	11,822	11,822	11,822
R-squared	0.016	0.029	0.041	0.045	0.036	0.025	0.029	0.018	0.022	0.028
Mean of dep. var.	0.128	0.432	0.758	0.323	0.135	0.0656	9.657	0.965	0.938	0.876
Std. dev. of dep. var.	0.334	0.495	0.428	0.467	0.341	0.248	1.997	0.184	0.241	0.330

Notes: Data come from the 2003–2024 American Time Use Survey (ATUS). 2020 is omitted due to incomplete data collection during the COVID-19 pandemic, and observations are restricted to weekday diaries. This table estimates the main table for ATUS with academic year fixed effect. Observations are weighted by population weights obtained from the National Cancer Institute's SEER Program and assigned according to individual's state of residence, survey year, age, sex, and race. Panel A restricts the analysis sample to children aged 15–18. Panel B restricts the sample to households with exactly one child under the age of 13. Panel A specifications include fixed effects for state, year, sex, age, month, SEER race, and Hispanic origin and control for state-level characteristics (% households with children, marital status and education-related variables, the share of 16–19 year-olds out of school, the unemployment rate, and the child poverty rate). Panel B specifications include fixed effects for state, year, month, SEER race, and Hispanic origin and control for state-level characteristics (total population, % households with children, share divorced, share with at least a high school degree, the unemployment rate, median household income, and the child poverty rate). Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.13: SB 328 & Parental Work-Related Time Use, ATUS (2003–24), ACS (2015-2024)

	(1) Awake by 6am	(2) Arrival at work (7-8am)	(3) Arrival at work (8-9am)	(4) Commuting time	(5) Working past 4pm
CA SB 328	-0.065***	-0.011***	0.012***	0.021	0.058**
	(0.022)	(0.002)	(0.003)	(0.181)	(0.022)
Observations	14,000	1,973,940	1,973,940	1,973,940	14,000
R-squared	0.024	0.021	0.023	0.036	0.035
Mean of dep. var.	0.583	0.301	0.218	28.04	0.485
Std. dev. of dep. var.	0.493	0.459	0.413	23.78	0.500

Notes: Data come from the 2003–2024 American Time Use Survey (ATUS): cols. (1) and (5); and the 2015–2024 American Community Survey (ACS): cols. (2)-(4). ATUS sample excludes 2020 due to incomplete data collection during COVID-19 pandemic. ACS sample excludes 2020 (experimental COVID-19 sample) and 2022 (interview month is unavailable). Both samples are restricted to parents (ages 25-64) of children (ages 11-18), who worked away from home. ATUS analysis is restricted to weekdays and observations are weighted by population weights obtained from the National Cancer Institute's SEER Program (assigned according to individual's state of residence, survey year, age, sex, and race). All specifications include fixed effects for state, year, while ATUS regressions also include month fixed effects. All specifications control for the following individual-level characteristics: age, sex, race/ethnicity, education, marital status, and the number of children in the household. All specifications further control for the following state-level characteristics: total population, % households with children, share married, share divorced, share with at least a high school degree, share of 16–19 year-olds out of school, child poverty rate, unemployment rate, and median household income. Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.14: Above and Below Mean – Effect of SB 328 on Academic and School Outcomes (CDE)

	(1) Log Math Score	(2) Math Std. Met Share	(3) Log English Score	(4) English Std. Met Share	(5) Log Science Score	(6) Science Std. Met Share	(7) Dropout Rate	(8) Suspension Rate	(9) No. Suspensions per 1,000
Panel A: High School									
Policy (SB 328)	0.0023***	2.126***	0.0017***	1.690***	0.0017***	1.747***	-0.868*	-0.174	-3.439
	(0.001)	(0.496)	(0.001)	(0.507)	(0.000)	(0.501)	(0.443)	(0.160)	(3.878)
Observations	13,653	13,660	13,668	13,676	8,014	8,014	14,507	75,541	75,541
R-squared	0.165	0.095	0.077	0.065	0.246	0.144	0.023	0.168	0.104
Mean of dep. var.		24.81		50.93		26.32	8.700	3.624	60.56
Std. dev. of dep. var.		20.47		23.46		18.27	12.28	5.279	117.8
Panel B: Middle School									
Policy (SB 328)	0.0014*	1.898***	0.0013**	1.862***	0.0011**	0.838**			
	(0.001)	(0.478)	(0.001)	(0.539)	(0.000)	(0.396)			
Observations	17,236	17,243	17,246	17,252	10,116	10,116			
R-squared	0.893	0.875	0.867	0.852	0.034	0.041			
Mean of dep. var.		33.95		48.09		30.17			
Std. dev. of dep. var.		20.79		19.37		18.33			

Notes: Results of a robustness check where teacher-arrival treatment intensity is defined as above vs. below mean (and not based on quartiles). Regressions in Panel A (cols. 1-7), use teacher-arrival treatment intensity defined for high school teachers (IPUMS occupation 2320); regressions in Panel B (cols. 1-6) use teacher-arrival treatment intensity defined for elementary/middle school teachers (IPUMS occupation 2310); regressions in Panel A (cols. 8-9) use a combined HS+ES/MS teacher-arrival treatment intensity (IPUMS occupations 2320 and 2310). Data come from the California Department of Education for the academic years 2014–15 through 2024–25 (Panels A and B, cols. 1-4); academic years 2018–19 through 2024–25 (Panels A and B, cols. 5-6); academic years 2016–17 through 2024–25 (Panel A, col. 7); and academic years 2011–12 through 2024–25 (Panel A, cols. 8-9). In Panels A and B (cols. 1-6), academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. Panel A (cols. 7-9) contain school-wide outcomes, with high school-only sample in col. (7), and a combined high school/middle school sample in cols. (8)-(9). All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, and female students, and the share of socioeconomically disadvantaged students. Standard errors in parentheses, clustered at the county level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.15: Top and Bottom Quartile – Effect of SB 328 on Academic and School Outcomes (CDE)

	(1) Log Math Score	(2) Math Std. Met Share	(3) Log English Score	(4) English Std. Met Share	(5) Log Science Score	(6) Science Std. Met Share	(7) Dropout Rate	(8) Suspension Rate	(9) No. Suspensions per 1,000
Panel A: High School									
Policy (SB 328)	0.0017	1.903**	0.0014*	1.865**	0.0011*	0.997	-1.311*	-0.307	-7.437
	(0.001)	(0.864)	(0.001)	(0.775)	(0.001)	(0.625)	(0.670)	(0.198)	(4.553)
Observations	5,558	5,560	5,558	5,560	3,277	3,277	5,974	38,119	38,119
R-squared	0.190	0.100	0.092	0.069	0.196	0.098	0.028	0.179	0.119
Mean of dep. var.		24.04		49.91		25.66	8.526	4.152	69.14
Std. dev. of dep. var.		20.33		23.47		18.21	11.95	5.398	115.4
Panel B: Middle School									
Policy (SB 328)	0.0007	1.515**	0.0002	0.878	0.0009	0.560			
	(0.001)	(0.679)	(0.001)	(0.696)	(0.001)	(0.445)			
Observations	8,470	8,474	8,476	8,480	5,055	5,055			
R-squared	0.894	0.876	0.860	0.846	0.038	0.045			
Mean of dep. var.		32.91		47.10		29.29			
Std. dev. of dep. var.		21.03		19.56		18.36			

Notes: Results of a robustness check where teacher-arrival treatment intensity is defined as top vs. bottom quartile plus rural counties. Regressions in Panel A (cols. 1-7), use teacher-arrival treatment intensity defined for high school teachers (IPUMS occupation 2320); regressions in Panel B (cols. 1-6) use teacher-arrival treatment intensity defined for elementary/middle school teachers (IPUMS occupation 2310); regressions in Panel A (cols. 8-9) use a combined HS+ES/MS teacher-arrival treatment intensity (IPUMS occupations 2320 and 2310). Data come from the California Department of Education for the academic years 2014–15 through 2024–25 (Panels A and B, cols. 1-4); academic years 2018–19 through 2024–25 (Panels A and B, cols. 5-6); academic years 2016–17 through 2024–25 (Panel A, col. 7); and academic years 2011–12 through 2024–25 (Panel A, cols. 8-9). In Panels A and B (cols. 1-6), academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. Panel A (cols. 7-9) contain school-wide outcomes, with high school-only sample in col. (7), and a combined high school/middle school sample in cols. (8)-(9). All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, and female students, and the share of socioeconomically disadvantaged students. Standard errors in parentheses, clustered at the county level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.16: Survey Responses – Effect of SB 328 on Academic and School Outcomes (CDE)

	(1) Log Math Score	(2) Math Std. Met Share	(3) Log English Score	(4) English Std. Met Share	(5) Log Science Score	(6) Science Std. Met Share	(7) Dropout Rate
Panel A: High School							
Policy (SB 328)	0.0019***	1.305**	0.0015**	1.520**	0.0017***	1.755***	-1.024*
	(0.001)	(0.544)	(0.001)	(0.621)	(0.001)	(0.584)	(0.609)
Observations	13,653	13,660	13,668	13,676	8,014	8,014	14,507
R-squared	0.163	0.091	0.076	0.064	0.246	0.143	0.023
Mean of dep. var.		24.81		50.93		26.32	8.700
Std. dev. of dep. var.		20.47		23.46		18.27	12.28
Panel B: Middle School							
Policy (SB 328)	0.0013**	1.350***	0.0013**	1.645***	0.0009*	0.433	
	(0.001)	(0.418)	(0.001)	(0.552)	(0.001)	(0.382)	
Observations	17,236	17,243	17,246	17,252	10,116	10,116	
R-squared	0.893	0.875	0.867	0.852	0.034	0.041	
Number of schools					2,230	2,230	
Mean of dep. var.		33.95		48.09		30.17	
Std. dev. of dep. var.		20.79		19.37		18.33	

Notes: Results of a robustness check where the ACS-based teacher-arrival treatment intensity (as defined in Table 5) is replaced with compliance information from the direct survey of California districts (for those districts which responded to the survey). Data come from the California Department of Education for the academic years 2014–15 through 2024–25 (Panels A and B, cols. 1-4); academic years 2018–19 through 2024–25 (Panels A and B, cols. 5-6); and academic years 2016–17 through 2024–25 (Panel A, col. 7). In Panels A and B (cols. 1-6), academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. Panel A, col. (7) contains the school-wide dropout rate for the high school sample. All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, and female students, and the share of socioeconomically disadvantaged students. Standard errors in parentheses, clustered at the county level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.17: SB 328, Sleep and Mental Health by Gender (YRBS)

	(1) Sleep ≥ 6	(2) Sleep ≥ 7	(3) Sleep ≥ 8	(4) Difficulty concentrating	(5) Felt sad/hopeless	(6) Suicidal intentions
Girls						
CA SB 328	0.069***	0.052**	0.032	-0.055*	-0.034	0.011
	(0.018)	(0.022)	(0.020)	(0.031)	(0.034)	(0.021)
Observations	59,288	59,288	59,288	24,144	67,424	59,782
R-squared	0.020	0.028	0.025	0.047	0.045	0.016
Mean of dep. var.	0.778	0.532	0.253	0.478	0.411	0.239
Std. dev. of dep. var.	0.416	0.499	0.435	0.500	0.492	0.426
Boys						
CA SB 328	0.068***	0.086***	0.041**	-0.056*	-0.023**	-0.038***
	(0.019)	(0.021)	(0.019)	(0.029)	(0.011)	(0.009)
Observations	58,747	58,747	58,747	24,055	66,972	59,501
R-squared	0.028	0.043	0.046	0.014	0.019	0.008
Mean of dep. var.	0.806	0.584	0.290	0.300	0.225	0.124
Std. dev. of dep. var.	0.395	0.493	0.454	0.458	0.418	0.329

Notes: Data comes from the 2007-2023 national Youth Risk Behavior Survey (YRBS). Observations are weighted by population weights obtained from National Cancer Institute's SEER Program and assigned according to individual's state of residence, survey year, age, sex, and race. All specifications control for individual-level covariates (grade, race/ethnicity) and state-level characteristics (population, population density, % households with children, marital status & education-related variables, unemployment rate, median household income, and child poverty rate). All regressions further include state and year fixed effects. Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.

Table A.18: Within-State Analysis by Gender – Effect of SB 328 on High School Academic and School Outcomes (CDE)

	(1) Log Math Score	(2) Math Std. Met Share	(3) Log English Score	(4) English Std. Met Share	(5) Log Science Score	(6) Science Std. Met Share	(7) Dropout Rate	(8) Suspension Rate	(9) No. Suspensions per 1,000
Girls									
Policy (SB 328)	0.0020**	2.061***	0.0017**	1.770**	0.0014***	1.563**	-1.133*	-0.198	-3.492*
	(0.001)	(0.644)	(0.001)	(0.670)	(0.001)	(0.595)	(0.634)	(0.123)	(2.015)
Observations	12,398	12,403	12,434	12,439	7,205	7,205	13,711	75,541	75,541
R-squared	0.219	0.131	0.059	0.052	0.228	0.136	0.015	0.089	0.057
Mean of dep. var.		25.58		57.43		26.28	6.977	2.172	32.89
Std. dev. of dep. var.		20.50		23.36		18.67	11.14	4.178	83.71
Boys									
Policy (SB 328)	0.0018**	1.699***	0.0011*	1.084*	0.0015***	1.481**	-1.069*	-0.432*	-9.559*
	(0.001)	(0.540)	(0.001)	(0.583)	(0.001)	(0.588)	(0.573)	(0.239)	(5.617)
Observations	12,507	12,515	12,514	12,521	7,264	7,264	13,773	75,531	75,531
R-squared	0.109	0.062	0.071	0.051	0.191	0.100	0.026	0.185	0.116
Mean of dep. var.		26.24		46.86		27.68	9.586	4.936	85.23
Std. dev. of dep. var.		21.07		22.98		18.75	12.91	6.483	151.0

Notes: Top panel contains regression estimates with outcomes measured for girls, while bottom panel contains regression estimates with outcomes measured for boys. Data come from the California Department of Education for the academic years 2014–15 through 2024–25 (cols. 1-4); academic years 2018–19 through 2024–25 (cols. 5-6); academic years 2016–17 through 2024–25 (col. 7); and academic years 2011–12 through 2024–25 (cols. 8-9). In cols. (1)-(6), academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. Cols. (7)-(9) show school-wide outcomes, with high school-only sample in col. (7), and a combined high school/middle school sample in cols. (8)-(9). Regressions in cols. (1)-(7) use teacher-arrival treatment intensity defined for high school teachers (IPUMS occupation 2320; bottom three quartiles vs. top quartile plus rural counties). Regressions in cols. (8)-(9) use a combined HS+ES/MS teacher-arrival treatment intensity (IPUMS occupations 2320 and 2310; bottom three quartiles vs. top quartile plus rural counties). All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, female students, and the share of socioeconomically disadvantaged students. Standard errors in parentheses, clustered at the county level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.19: Within-State Analysis by Gender – Effect of SB 328 on Middle School Academic and School Outcomes (CDE)

	(1) Log Math Score	(2) Math Std. Met Share	(3) Log English Score	(4) English Std. Met Share	(5) Log Science Score	(6) Science Std. Met Share
Girls						
Policy (SB 328)	0.0006	1.551***	0.0009	1.677***	0.0010	0.810*
	(0.001)	(0.534)	(0.001)	(0.551)	(0.001)	(0.471)
Observations	15,568	15,570	15,569	15,571	9,159	9,159
R-squared	0.881	0.857	0.854	0.828	0.033	0.047
Mean of dep. var.		34.54		54.14		29.19
Std. dev. of dep. var.		21.31		19.71		18.84
Boys						
Policy (SB 328)	0.0009	1.137**	0.0006	1.042	0.0005	0.530
	(0.001)	(0.546)	(0.001)	(0.677)	(0.001)	(0.405)
Observations	15,538	15,541	15,549	15,552	9,125	9,125
R-squared	0.886	0.865	0.853	0.836	0.042	0.037
Mean of dep. var.		33.84		42.48		31.43
Std. dev. of dep. var.		20.87		19.59		18.70

Notes: Top panel contains regression estimates with outcomes measured for girls, while bottom panel contains regression estimates with outcomes measured for boys. Data come from the California Department of Education for the academic years 2014–15 through 2024–25 (cols. 1-4); academic years 2018–19 through 2024–25 (cols. 5-6). In all regressions, academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. All specifications employ elementary/middle school teacher-arrival treatment intensity (IPUMS occupation 2310; bottom two quartiles vs. top two quartiles plus rural counties). All regressions further include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, female students, and the share of socioeconomically disadvantaged students. Standard errors in parentheses, clustered at the county level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.20: SB 328, Sleep and Mental Health by Race (YRBS)

	(1) Sleep ≥ 6	(2) Sleep ≥ 7	(3) Sleep ≥ 8	(4) Difficulty concentrating	(5) Felt sad/hopeless	(6) Suicidal intentions
Hispanics						
CA SB 328	0.057*	0.040***	0.012	-0.099***	-0.010	-0.025
	(0.030)	(0.014)	(0.009)	(0.027)	(0.021)	(0.018)
Observations	30,043	30,043	30,043	12,745	34,745	30,144
R-squared	0.021	0.037	0.030	0.057	0.065	0.030
Mean of dep. var.	0.799	0.569	0.286	0.404	0.343	0.179
Std. dev. of dep. var.	0.401	0.495	0.452	0.491	0.475	0.383
Whites						
CA SB 328	0.080***	0.093***	0.056**	-0.033	-0.031*	-0.016
	(0.023)	(0.026)	(0.026)	(0.034)	(0.018)	(0.012)
Observations	55,808	55,808	55,808	22,566	61,574	55,325
R-squared	0.020	0.034	0.038	0.053	0.068	0.030
Mean of dep. var.	0.818	0.590	0.281	0.379	0.307	0.183
Std. dev. of dep. var.	0.386	0.492	0.449	0.485	0.461	0.387
Others						
CA SB 328	0.017	-0.007	-0.009	-0.007	-0.016	0.020
	(0.021)	(0.019)	(0.016)	(0.019)	(0.028)	(0.017)
Observations	32,184	32,184	32,184	12,888	38,077	33,814
R-squared	0.023	0.032	0.033	0.062	0.064	0.037
Mean of dep. var.	0.740	0.491	0.241	0.391	0.316	0.181
Std. dev. of dep. var.	0.439	0.500	0.428	0.488	0.465	0.385

Notes: Data comes from the 2007-2023 national Youth Risk Behavior Survey (YRBS). Observations are weighted by population weights obtained from National Cancer Institute's SEER Program and assigned according to individual's state of residence, survey year, age, sex, and race. All specifications control for individual-level covariates (grade, sex) and state-level characteristics (population, population density, % households with children, marital status & education-related variables, unemployment rate, median household income, and child poverty rate). All regressions further include state and year fixed effects. Standard errors in parentheses, clustered at the state level. *** Significant at the 1% level ** Significant at the 5% level * Significant at the 10% level.

Table A.21: Within-State Analysis by Race – Effect of SB 328 on High School Academic and School Outcomes (CDE)

	(1) Log Math Score	(2) Math Std. Met Share	(3) Log English Score	(4) English Std. Met Share	(5) Log Science Score	(6) Science Std. Met Share	(7) Dropout Rate	(8) Suspension Rate	(9) No. Suspensions per 1,000
Panel A: Hispanics									
Policy (SB 328)	0.0034***	1.921***	0.0025***	1.928***	0.0020***	2.254***	-1.272*	-0.357**	-5.995
	(0.001)	(0.473)	(0.001)	(0.584)	(0.001)	(0.655)	(0.680)	(0.171)	(3.785)
Observations	11,725	11,728	11,792	11,796	6,982	6,982	13,062	75,541	75,541
R-squared	0.186	0.083	0.067	0.045	0.243	0.136	0.026	0.144	0.096
Mean of dep. var.		18.24		45.88		19.96	8.721	3.585	57.99
Std. dev. of dep. var.		14.77		21.13		14.23	12.22	5.303	113.0
Panel B: Whites									
Policy (SB 328)	0.0006	1.437**	0.0002	0.379	0.0009	1.150	-0.298	-0.207	-5.860
	(0.001)	(0.661)	(0.001)	(0.679)	(0.001)	(0.773)	(0.442)	(0.167)	(3.716)
Observations	9,338	9,344	9,375	9,380	5,324	5,324	10,981	75,425	75,425
R-squared	0.084	0.058	0.040	0.041	0.120	0.070	0.012	0.104	0.069
Mean of dep. var.		35.11		62.99		37.97	7.842	3.632	61.65
Std. dev. of dep. var.		21.14		20.96		18.55	12.21	5.669	126.7
Panel C: Socioeconomically Disadvantaged									
Policy (SB 328)	0.0026***	1.502**	0.0015**	1.451**	0.0019***	1.861***	-1.234**	-0.352*	-7.071
	(0.001)	(0.627)	(0.001)	(0.675)	(0.001)	(0.506)	(0.612)	(0.202)	(4.482)
Observations	12,408	12,413	12,426	12,431	7,399	7,399	14,153	75,445	75,445
R-squared	0.150	0.071	0.048	0.036	0.234	0.135	0.029	0.167	0.106
Mean of dep. var.		19.07		45.12		21.15	9.210	4.162	70.13
Std. dev. of dep. var.		15.94		21.06		15.01	12.40	5.637	126.7

Notes: Panel A contains regression estimates with outcomes measured for Hispanic students; Panel B contains regression estimates with outcomes measured for White students; Panel C contains regression estimates with outcomes measured for socioeconomically disadvantaged students. Data come from the California Department of Education for the academic years 2014–15 through 2024–25 (cols. 1-4); academic years 2018–19 through 2024–25 (cols. 5-6); academic years 2016–17 through 2024–25 (col. 7); and academic years 2011–12 through 2024–25 (cols. 8-9). In cols. (1)-(6), academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. Cols. (7)-(9) show school-wide outcomes, with high school-only sample in col. (7), and a combined high school/middle school sample in cols. (8)-(9). Regressions in cols. (1)-(7) use teacher-arrival treatment intensity defined for high school teachers (IPUMS occupation 2320; bottom three quartiles vs. top quartile plus rural counties). Regressions in cols. (8)-(9) use a combined HS+ES/MS teacher-arrival treatment intensity (IPUMS occupations 2320 and 2310; bottom three quartiles vs. top quartile plus rural counties). All specifications include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, female students, and the share of socioeconomically disadvantaged students. Standard errors in parentheses, clustered at the county level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Table A.22: Within-State Analysis by Race – Effect of SB 328 on Middle School Academic Outcomes (CDE)

	(1) Log Math Score	(2) Math Std. Met Share	(3) Log English Score	(4) English Std. Met Share	(5) Log Science Score	(6) Science Std. Met Share
Panel A: Hispanics						
Policy (SB 328)	0.0025***	1.924***	0.0019**	1.816***	0.0015***	0.881*
	(0.001)	(0.542)	(0.001)	(0.606)	(0.001)	(0.451)
Observations	14,166	14,168	14,162	14,164	8,418	8,418
R-squared	0.813	0.774	0.784	0.756	0.033	0.044
Mean of dep. var.		24.70		40.01		21.67
Std. dev. of dep. var.		15.62		16.31		13.47
Panel B: Whites						
Policy (SB 328)	0.0004	0.864	-0.0001	0.720	-0.0008	-0.473
	(0.001)	(0.777)	(0.001)	(0.737)	(0.001)	(0.658)
Observations	11,744	11,746	11,769	11,771	6,617	6,617
R-squared	0.823	0.790	0.791	0.753	0.030	0.024
Mean of dep. var.		44.23		58.67		42.12
Std. dev. of dep. var.		20.52		18.53		18.22
Panel C: Socioeconomically Disadvantaged						
Policy (SB 328)	0.0014	1.260**	0.0010*	0.806	0.0008	0.543
	(0.001)	(0.620)	(0.001)	(0.575)	(0.000)	(0.337)
Observations	15,096	15,101	15,092	15,096	8,969	8,969
R-squared	0.804	0.763	0.762	0.727	0.034	0.037
Mean of dep. var.		24.47		38.54		21.73
Std. dev. of dep. var.		15.51		15.62		13.38

Notes: Panel A contains regression estimates with outcomes measured for Hispanic students; Panel B contains regression estimates with outcomes measured for White students; Panel C contains regression estimates with outcomes measured for socioeconomically disadvantaged students. Data come from the California Department of Education for the academic years 2014–15 through 2024–25 (cols. 1-4); academic years 2018–19 through 2024–25 (cols. 5-6). In all regressions, academic years 2019-2020 and 2020-2021 are excluded due to COVID-19 disruptions. All specifications employ elementary/middle school teacher-arrival treatment intensity (IPUMS occupation 2310; bottom two quartiles vs. top two quartiles plus rural counties). All regressions further include school fixed effects, year fixed effects, baseline-enrollment-quartile-by-year linear trends, and control for total enrollment, the shares of White, Hispanic, female students, and the share of socioeconomically disadvantaged students. Standard errors in parentheses, clustered at the county level. *** Significant at the 1% level, ** Significant at the 5% level, * Significant at the 10% level.

Evidence Taxonomy

Table A.23 summarizes principal studies discussed in the text, organized by evidence type. The pool used to construct Figures A.2 and A.10 includes additional within-paper rows from Bastian & Fuller (2018) and Fuller & Bastian (2024), and several Panel A sleep studies (Owens 2010; Boergers; Temkin; Yang & Choi) not separately listed here.

Table A.23: Evidence Taxonomy: School Start Times, Sleep, and Student Outcomes

Study	Evidence	Design	Sleep	Academic / Other	Verdict
Panel A: Correlational and descriptive					
Wahlstrom et al. (2001, 2014)	Descriptive	Pre/post; no control	+	Slight GPA gains; ↓ tardiness	Mixed
Dewald et al. (2010)	Correlational	Meta-analysis; 26 studies	—	$r = 0.07$–0.13	Suggestive
Dunster et al. (2018)	Pre/post	Actigraphy; 2 HS; no control	+34 min	+4.5% grades	Mixed
Bowers & Moyer (2017)	Meta-analysis	20 studies	+	No academic pooling	Sleep: +
Yip et al. (2022)	Meta-analysis	28 studies; $N \approx$ 1.8M	+	No significant pooled effect[a]	Sleep: +
Panel B: Quasi-experimental — United States					
Hinrichs (2011)[§]	QE	Cross-district DiD	—	0.00 (ACT)[b]	Null
Carrell et al. (2011)[§]	QE	Randomized; USAFA	—	+0.15 SD[k]	Positive
Edwards (2012)[§]	QE	Transportation IV; NC	—	+0.04–0.06 SD[c]	Positive

Continued on next page

Table A.23 continued

Study	Evidence	Design	Sleep	Academic / Other	Verdict
Lusher & Yasenov (2016)	QE	Double-shift; Eastern Europe	—	0.00 (aggregate)[d]	Null (aggregate)
Lusher & Yasenov (2018)	QE	Same; gender-disaggregated	—	+ boys; 0.00 girls	Positive (boys)
Luong et al. (2019)	Experimental	Random assignment; Vietnam	+	+ (smaller than obs.)	Positive
Groen & Pabilonia (2019)[§]	Observational	OLS + Oster bounds; PSID-CDS	+ (F)	+0.16–0.28 SD (F; suggestive)[e]	Suggestive
Heissel & Norris (2018)[§]	QE	Time-zone IV; FL	—	+0.06–0.08 SD[f]	Positive
Owens et al. (2017)[§]	Pre/post	Pre/post; Fairfax County VA	+30 min	—	Sleep: +
Widome et al. (2020)[§]	QE	DiD + actigraphy; MN	+41–43 min	—	Sleep: +
Lenard et al. (2020)[§]	QE	DiD (advance); Wake County	—	+0.04 SD $[-0.19, 0.28]^{\dagger}$	Null (test)
James et al. (2023)	QE	DiD; MN (START cohort)	—[g]	+0.17 GPA pts; ↓ tardiness	Mixed[g]
Panel C: Quasi-experimental — international and time-of-day					
Cortes et al. (2012)	QE	Admin data; first-period; Chicago	—	↓ grades, first period	Positive
Pope (2016)	QE	Admin data; LAUSD	—	Mixed (teacher learning)	Mixed
Jung & Fenelon (2024)[§]	QE	DiD; South Korea mandate	+19–7 min[h]	↓ substance use	Sleep: + (fades)

Continued on next page

Table A.23 continued

Study	Evidence	Design	Sleep	Academic / Other	Verdict
Lo et al. (2018)[§]	Pre/post	Pre/post; Singapore	+23 min (TIB)	—	Sleep: +
Chan et al. (2018)	Pre/post	Pre/post boarding; HK (1-mo follow-up)	+48 min[l]	—	Excluded from pool[l]
Chan et al. (2024)[§]	Pre/post	Pre/post boarding; HK (7-mo follow-up)	+35 min	↑ life satisfaction	Positive[i]
Sadikova et al. (2024)	QE	IV (SST→sleep); MN	—	↓ depression	Positive[j]
Bastian & Fuller (2023)[§]	QE	Panel DiD; NC middle	—	+0.08 SD math; +0.04 reading	Positive
Panel D: Experimental evidence on sleep					
Giuntella et al. (2026)	RCT	Field experiment; US university	+19 min	+0.10–0.11 SD	Positive

Continued on next page

Table A.23 continued

Study	Evidence	Design	Sleep	Academic / Other	Verdict

Notes: QE = quasi-experimental; RCT = randomized controlled trial; DiD = difference-in-differences; F = female students only; HS = high school; MN = Minnesota. Sleep effects in minutes of additional school-night sleep; academic effects in SD units where available. Dashes indicate outcome not primary focus. [§] Study enters the pooled forest plots (Appendix Figures A.2 and A.10); for sleep duration the pool includes pre/post designs with repeated measurement of the same students, as stated in inclusion criterion (iii). Rows without this mark are background studies and are not pooled. [a] The pooled academic effect aggregates estimates across designs of varying quality. [b] The ACT captures cumulative knowledge and is likely to respond slowly to short-run schedule changes. This caveat applies to every ACT-outcome row in the meta-analysis, including Hinrichs (2011), Lenard et al. (2020), Bastian & Fuller (2018, ACT), and Fuller & Bastian (2024, ACT). [c] Standardized-test (Z-score) coefficients from Edwards (2012), Table 7, student-school FE specification (math +0.057, SE 0.010; reading +0.035, SE 0.013). [d] Aggregate null reflects offsetting gender-specific effects. [e] OLS with Oster (2019) bounds; not causally identified. [f] Effects concentrated among pubescent students; near zero for pre-pubescent children. [g] The sleep result of +40 minutes sometimes attributed to James et al. (2023) is a citation of Widome et al. (2020) inside their introduction, not an independent estimate. James's GPA effect is reported in grade-point units (not SD), so the study contributes to the qualitative discussion only and is excluded from the pooled academic estimate. [†] Studies a start time advance rather than a delay; the per-hour-delay equivalent is computed by dividing the raw ACT-points coefficient by the in-sample ACT SD (5.727) and flipping sign to express on a delay scale. [h] Initial gain of 19 min. fades to 7 min. within one year as bedtimes shift later. [i] Seven-month follow-up; significant reductions in depression, anxiety, and insomnia ($d = 0.29$–0.67). A separate $d = 0.72$ effect-size in the same paper refers to total sleep time (paired-sample), not mental health. [j] IV-estimated: 30-min sleep gain $\rightarrow -0.78$ depression points (SE 0.27); -1.36 fatigue points (SE 0.38). [k] Per-hour-equivalent of the 7:00 vs. 7:50 a.m. first-period contrast at USAFA, rescaled from the source paper's 50-min contrast (raw 0.13 SD over 50 min $\rightarrow$ 0.15 SD per hour). [l] Same Hong Kong residential boarding-school intervention as Chan et al. (2024), reported at 1-month follow-up rather than 7-month. Excluded from the sleep pool to avoid double-counting the shared Cohort 1 baseline sample.

Meta-analytic synthesis. To situate our estimates within the broader literature, we conduct a PRISMA-style systematic review of quasi-experimental evidence on school start times, which also provides a basis for revisiting the mixed conclusions of prior meta-analyses. We anchor the search on Bowers and Moyer (2017) and Yip et al. (2022) and on the most frequently cited causally-identified studies in the field (Carrell et al., 2011; Heissel and Norris, 2018), conducting backward and forward citation searches via OpenAlex supplemented with keyword searches of the NBER and IZA working paper series. After three rounds of citation expansion and deduplication, 320 candidate records were reduced to 136 after title/abstract screening and to 42 after applying five inclusion criteria (Appendix Figure A.1): (i) school start time as the independent variable; (ii) sleep, academic, or mental health outcomes; (iii) a design that provides a counterfactual: for academic and mental-health outcomes, a comparison-group design (pre/post with a control group, DiD, IV, or RCT); for sleep duration, also pre/post designs that measure the same students before and after a schedule change, the standard design of actigraphy studies; or a selection-on-observables design accompanied by a formal omitted-variable sensitivity analysis that bounds the estimate (e.g., Oster (2019) coefficient-stability bounds), in all cases excluding unadjusted cross-sectional correlations; (iv) peer-reviewed publication in English; and (v) extractable effect sizes. Of the 42 included studies, 33 report parametric standard errors and form the quantitative synthesis; 18 of these provide effect sizes in a common metric and enter the pooled forest plots (10 for sleep duration and 9 for test scores, with one study in both), while the other 15 are tabulated but not pooled. The remaining 9 studies, which report no parametric standard errors, contribute to the qualitative synthesis. Exactly one quantitative study, Groen and Pabilonia (2019), enters under the sensitivity-bounds provision of criterion (iii) rather than a design-based contrast; it is reported throughout as corroborating rather than design-based evidence, and its inclusion does not affect the screening counts above. Because the included estimates span heterogeneous treatment contrasts (delays versus advances; sub-hour versus multi-hour changes) and populations (U.S. military cadets, middle schoolers, Korean

high schoolers under a national mandate), and because several studies contribute correlated math and reading coefficients from a common specification, we present the full distribution of effects rather than a single inverse-variance-pooled point estimate, which would convey artificial precision about a parameter whose interpretation is not invariant across the underlying studies.[23] Across the quasi-experimental estimates of academic effects (Figure A.10), point estimates range from −0.04 to +0.16 SD per one-hour delay, with the bulk between 0.00 and 0.08 SD. Four estimates are non-positive (Hinrichs (2011), the biology and English rows of Bastian and Fuller (2018), and the Math I row of Fuller and Bastian (2024)); Lenard et al. (2020) separately studies a start time advance rather than a delay. For sleep duration (Appendix Figure A.2), non-boarding studies in the U.S., Korea, and Singapore cluster around 20–45 additional minutes per one-hour delay; the largest effects come from boarding-school settings, where U.S. boarding schools (Owens et al. 2010, Boergers et al. 2014) yield 70–90 minutes and Hong Kong boarding schools (Chan et al. 2018, Chan et al. 2024) yield 35–48 minutes per one-hour delay. Appendix Table A.1 reports representative study-level estimates across sleep, academic, and mental health outcomes; Appendix Table A.23 summarizes the principal studies discussed in the text, and Figure A.1 reports the PRISMA flow.

[23]For completeness, inverse-variance-weighted random-effects (Restricted Maximum Likelihood, REML) estimation yields a one-hour delay increasing sleep duration by approximately 39 minutes (95% CI: 26–51; $k = 10$; $I^2 = 90\%$) and raising standardized test scores by 0.03 SD (95% CI: 0.01–0.05; $k = 18$ study–outcome estimates from 9 independent studies; $I^2 = 72.8\%$). These pooled estimates remain larger and more precisely estimated than the null effect reported by Yip et al. (2022) ($r = 0.026$, $p = .70$, $k = 8$), a difference we attribute to our stricter inclusion criteria, which admit design-based estimates, pre/post designs that follow the same students, and selection-on-observables estimates carrying formal omitted-variable bounds, and so exclude the unadjusted cross-sectional correlations that enter broader pools such as that of Yip et al. (2022). We do not provide a forest plot for mental-health outcomes because the included studies use highly heterogeneous instruments (DASS-21, Kutcher Adolescent Depression Scale, PANAS, single-item depressed-mood scales, life-satisfaction Likert scales, and substance-use frequency), which preclude a common effect-size metric without auxiliary scale-conversion assumptions.