ANNENBERG
BROWN UNIVERSITY

EdWorkingPaper No. 24-974

Do later school start times improve adolescents' sleep and substance use? A quasi-experimental study

Jilli Jung
Pennsylvania State University

Andrew Fenelon
University of Minnesota-Twin Cities

A later school start time policy has been recommended as a solution to adolescents' sleep deprivation. We estimated the impacts of later school start times on adolescents' sleep and substance use by leveraging a quasi-experiment in which school start time was delayed in some regions in South Korea. A later school start time policy was implemented in 2014 and 2015, which delayed school start times approximately 30-90 minutes. We applied difference-in-differences and event-study designs to longitudinal data on a nationally-representative cohort of adolescents from 2010 to 2015, which annually tracked sleep and substance use of 1,133 adolescents from grade 7 through grade 12. The adoption of a later school start time policy was initially associated with a 19-minute increase in sleep duration (95% CI, 5.52 to 32.04), driven by a delayed wake time and consistent bedtime. The policy was also associated with statistically significant reductions in monthly smoking and drinking frequencies. However, approximately a year after implementation, the observed increase in sleep duration shrank to 7-minute (95% CI, -12.60 to 25.86) and became statistically nonsignificant. Similarly, the observed reduction in smoking and drinking was attenuated a year after. Our findings suggest that policies that increase sleep in adolescents may have positive effects on health behaviors, but additional efforts may be required to sustain positive impacts over time. Physicians and education and health policymakers should consider the long-term effects of later school start times on adolescent health and well-being.

VERSION: June 2024

Suggested citation: Jung, Jilli, and Andrew Fenelon. (2024). Do later school start times improve adolescents' sleep and substance use? A quasi-experimental study. (EdWorkingPaper: 24-974). Retrieved from Annenberg Institute at Brown University: https://doi.org/10.26300/6czb-af66

Do later school start times improve adolescents' sleep and substance use?

A quasi-experimental study

Jilli Jung[1]
Doctoral Candidate
Department of Education Policy Studies
Pennsylvania State University

Andrew Fenelon
Associate Professor
School of Public Health
University of Minnesota-Twin Cities

[1] Jilli Jung: jilli.jung@psu.edu. We thank Maithreyi Gopalan, Kelly Rosinger, Douglas Teti, and participants of the Pennsylvania State University Population Research Institute (PRI) Education and Inequality working group, the 2023 Society for Research on Educational Effectiveness, the 2023 Association for Public Policy Analysis and Management, the 2024 Association of Education Finance and Policy, and the 2024 American Educational Research Association annual conferences for helpful comments on an earlier draft. All remaining errors are our own.

Abstract

A later school start time policy has been recommended as a solution to adolescents' sleep deprivation. We estimated the impacts of later school start times on adolescents' sleep and substance use by leveraging a quasi-experiment in which school start time was delayed in some regions in South Korea. A later school start time policy was implemented in 2014 and 2015, which delayed school start times by approximately 30-90 minutes. We applied difference-in-differences and event-study designs to longitudinal data on a nationally representative cohort of adolescents from 2010 to 2015, which annually tracked sleep and substance use of 1,133 adolescents from grade 7 through grade 12. The adoption of a later school start time policy was initially associated with a 19-minute increase in sleep duration (95% CI, 5.52 to 32.04), driven by a delayed wake time and consistent bedtime. The policy was also associated with statistically significant reductions in monthly smoking and drinking frequencies. However, approximately a year after implementation, the observed increase in sleep duration shrank to 7 minutes (95% CI, -12.60 to 25.86) and became statistically nonsignificant. Similarly, the observed reduction in smoking and drinking was attenuated a year after. Our findings suggest that policies that increase sleep in adolescents may have positive effects on health behaviors, but additional efforts may be required to sustain positive impacts over time. Physicians and education and health policymakers should consider the long-term effects of later school start times on adolescent health and well-being.

Keywords: later school start time policy, sleep, substance use, alcohol drinking, cigarette smoking, quasi-experimental design, difference-in-differences, event study, South Korea

Introduction

Sleep deprivation in adolescence harms physical and mental health and increases the likelihood of unhealthy behaviors (Beebe, 2011; Shochat et al., 2014). While adolescents are recommended to sleep approximately 9 hours per night (Fuligni et al., 2019; Short et al., 2018), a recent nationally representative survey of Korean youth shows that only 5.9% of high school and 29.4% of middle school students got more than 8 hours of sleep per night (Ministry of Gender Equality and Family, 2022). Compared to children and adults, getting enough sleep is especially challenging for adolescents due to the natural alteration of circadian rhythms during adolescence that lead to biologically unique sleep patterns (Crowley et al., 2018). With the onset of adolescence, the accumulation of homeostatic sleep pressure decelerates, and this slow building of sleep pressure keeps adolescents awake until 11:00 pm or later and leads to naturally late wake times (Crowley et al., 2018). The delayed circadian rhythm does not necessarily lead to sleep deprivation unless late morning sleep is disturbed. Unfortunately, early school start times require adolescents to wake up early in the morning when their body "clocks" point to sleep time. For example, high schools in Seoul, the capital of Korea, started around 8:00 in 2019 (Seo, 2020). To accommodate this early school start time, adolescents must wake up around 7:00 am, which is equivalent to adults waking up at 4:00 am given adolescents' later peak of melatonin level (Carrell et al., 2011).

Because of this early school start time, most adolescents experience chronic sleep deficits, which lead to various negative consequences. An established body of literature has documented that adolescents' insufficient sleep is associated with poor mental and physical health, increased smoking and alcohol consumption, and worse academic performance (Beebe, 2011; Shochat et al., 2014). As its negative impacts have been recognized, professional organizations, including the American Academy of Pediatrics, have issued policy statements advocating secondary schools begin at 8:30 or later (Adolescent Sleep Working Group et al., 2014).

With increased attention to later school start times, many studies have investigated relationships between school start times and various students' outcomes (Bastian & Fuller, 2023; Biller et al., 2022; Bowers & Moyer, 2017; Carrell et al., 2011; Edwards et al., 2015; Heissel & Norris, 2018; Hinrichs, 2011; Minges & Redeker, 2016; Owens et al., 2017). With respect to substance use outcomes, recent reviews suggest a comprehensive picture of the causal impact of school start time policies is missing (Berger et al., 2019; Marx et al., 2017; Morgenthaler et al., 2016; Wahlstrom et al., 2017; Wahlstrom & Owens, 2017; Wheaton et al., 2015). First, with notable exceptions (Groen & Pabilonia, 2019; Yang & Choi, 2021), most previous studies on substance use rely on less rigorous research designs—pre-post and cross-sectional designs—which have difficulty isolating the causal impact of the policy from confounders (Edwards et al., 2015; Lo et al., 2018; Owens et al., 2010; Thacher & Onyper, 2016; Wahlstrom, 2014; Wahlstrom et al., 2017; Whitaker et al., 2019). Second, with a few exceptions (James et al., 2023; Lo et al., 2018; Thacher & Onyper, 2016; Widome et al., 2020), most previous studies have estimated short-term relationships between the later school start and outcomes. Given that longitudinal data suggest that effects of later school start times may not persist (James et al., 2023; Thacher & Onyper, 2016), longer-term analysis is necessary. Third, previous studies relied on small or non-representative samples, which is critical to external validity.

To fill this gap in knowledge, the present study exploited a quasi-experiment in Korea that delayed school start times to between 8:30 and 9:00 am. In 2014, Gyeonggi province implemented a 9:00 am school start policy, which requires all elementary and secondary schools in the region to begin at 9:00 am. Following Gyeonggi province, Jeollabuk-do delayed school start times by 30 minutes in 2014, and Gwangju and Incheon delayed school start times to 8:30-9:00 and 8:40-9:00 am in 2015, respectively. Given that more than half of schools started before 8:00 am prior to policy adoption, school start times have been delayed approximately 30-90 minutes as a result of the policy (Baek et al., 2015). The policy discussion was initiated as some middle schoolers proposed delaying school start times on the online suggestion box of the Gyeonggi Provincial Office of Education (The Hankyoreh, 2014). Because of the centralized education system of Korea, Gyeonggi and some other regions were able to implement the policy within three months, which makes gradual adoption or anticipation effects less likely.

Leveraging spatial and temporal variation in the later school start time policy, this study estimated causal impacts of the policy on sleep and substance use. This study analyzed longitudinal, nationally representative data from 2010 to 2015 using difference-in-differences. Considering prior research on the association between school start time and sleep duration (Minges & Redeker, 2016; Owens et al., 2017; Widome et al., 2020) and associations between shorter sleep duration and increased substance use (Claudatos et al., 2019; Nguyen-Louie et al., 2018; Terry-McElrath et al., 2016), we hypothesized that later school start times will lead to increased sleep duration and reduced substance use.

Methods

Data and participants

This study used the first to sixth waves of the Korean Children and Youth Panel Survey 2010 (KCYPS), a longitudinal, nationally representative survey of 7th graders in 2010. Administered by the National Youth Policy Institute (NYPI), a government-funded youth policy research institute, KCYPS is the most comprehensive longitudinal survey of adolescent development in Korea (National Youth Poicy Institute, 2020). KCYPS tracked a cohort of 2351 7th graders from 2010 to 2016, collecting annual information on demographics, sleep, socio-emotional well-being, health, and developmental environment (See Supplement for detailed information on sampling). Since the final survey (i.e., 7th wave) was conducted after high school graduation when students were no longer exposed to the policy, it was excluded from the analysis. The final analytic sample sizes varied across models because some outcomes were surveyed only at some waves: 6,324 observations from 1,133 students for models with sleep outcomes and 5,279 observations from 1,120 students for substance use outcomes. The Institutional Review Board at the Pennsylvania State University reviewed this study and considered it exempt.

School start time policy

The treatment was a region-level school start time policy that delayed school start times to between 8:30 and 9:00. The treatment group consisted of students in four regions that implemented later school start times: two regions, Gyeonggi and Jeollabuk-do, did it in 2014, and the other two regions, Gwangju and Incheon, did it in 2015. The comparison group consisted of students in five regions that did not delay school start times—Gyeongsangbuk-do, Daegu, Busan, Ulsan, and Jellanam-do. Some regions did not force but recommended schools to delay school start time, and many schools did not follow the guideline. Therefore, these regions are excluded from the analytic sample to have clean treatment and comparison groups. Since the policy was implemented at the school level, treatment was determined by the location of the school rather than residence. Although we do not have information on each school's start time before the policy implementation, we have information on the average of school start time before the policy implementation. As a result of the policy, school start times were delayed approximately 30-90 minutes (Baek et al., 2015).

Outcomes

Our outcomes of interest were sleep and substance use. Using reported sleep time information on school days, we focus on: (1) bedtime, (2) wake time, and (3) sleep duration. In each wave, students reported bedtime and wake time on school days and non-school days in the hour:minute format (e.g., 23:55). Sleep duration is measured as the difference in hours between bedtime and wake time on school days. The unit of sleep outcomes is hours (e.g., 6.3 = 6 hours and 20 minutes). Using reports of substance use, we examine: monthly frequency of (1) smoking and (2) drinking. For smoking, adolescents reported how many times they smoked cigarettes either per day or per year. For drinking, adolescents reported how many times they drank alcohol either per month or per year. We constructed monthly frequency by either multiplying per day reports by 30 or dividing per year reports by 12. These measures were surveyed from the second to sixth waves.

Time-varying covariates

As this study used individual-fixed effects, time-invariant characteristics, such as race and gender, were accounted for in models. To adjust for time-varying characteristics, models included the natural log of annual family income, family structure (having two parents=1, otherwise=0), and parental working status (having at least one parent currently not working=1, otherwise=0).

Analytic Strategy

We employed a difference-in-differences (DD) design to estimate the impacts of later school start times on students' sleep and substance use, leveraging variation in a later school start time policy adoption across regions and years. Specifically, using a two-way fixed effect (TWFE) approach, which is an extension of the canonical 2X2 DD design, we compared changes in outcomes of adolescents before and after a later school start time policy adoption in treatment regions with changes in comparison regions. Our main model includes individual- and year-fixed effects to adjust time-invariant factors and general trends in outcomes (See Supplement for TWFE model specifications). In addition to the main model, we estimated a model with time-varying covariates to account for time-varying individual-level factors and presented the results in Supplement Table 2. We clustered the standard errors at the region-level to account for the nested characteristics of data and the level of treatment (i.e., students within regions) and account for heteroscedasticity. Also, we used sampling weights to make the analysis nationally representative.

Supplemental Analyses

We conducted several supplemental analyses to better understand dynamic effects of later school start times and check the robustness of our findings. First, we conducted event-study analyses of sleep and substance outcomes to identify the dynamic effect of later school start times and to examine the pre-treatment parallel trend assumption (See Supplement for event-study model specifications). Second, we conducted the event-study analysis exclusively using early adopters (2014) to prevent underestimation of initial policy effects. Since students in the late adopter regions were surveyed six months after implementation, including these students in the analysis might dilute the immediate impact of the policy. Moreover, restricting the analysis to students from early adopter regions enables the first-year effect estimate to capture that in grade 11 instead of a combination of effects in grade 11 and 12. Next, we incorporated recent methodological advances to address potential bias in the TWFE approach with staggered treatment timings (Callaway & Sant'Anna, 2021; Goodman-Bacon, 2021; Sun & Abraham, 2021). First, we diagnosed how much weight is being placed for forbidden comparisons, using Goodman-Bacon decomposition (See Supplemental Table 3) (Goodman-Bacon, 2021). We then estimated policy effects using recently developed estimators for staggered timing to examine the robustness of findings from our TWFE model (See Supplemental Figure 1-5) (Callaway & Sant'Anna, 2021; Sun & Abraham, 2021). Lastly, we reported wild cluster bootstrap confidence intervals in addition to standard cluster-robust standard errors, which are more reliable and conservative when the number of clusters is small (Roodman et al., 2019).

Results

Descriptive statistics

Table 1 shows descriptive statistics for adolescents across treatment and comparison regions in 2013—a year before the first policy implementation. Female students accounted for about half of the analytic sample. Average parental education was 14.5 years, approximately 90% of students were from a two-parent household, and 33% of students had at least one parent currently not working. Across all regions, average sleep duration was 6 hours and 18 minutes on school nights, and average monthly smoking and drinking frequency was 5.81 and 0.10, respectively. Since we use a difference-in-differences (DD) design, which allows differences in levels, level differences before policy adoption (e.g., difference in the proportion of female students between treatment and control regions) do not bias estimations as long as a parallel trends assumption holds.

Effects of later school start times on sleep duration, wake time, and bedtime

Difference-in-differences (DD) estimates of the overall impacts of later school start times on adolescents' sleep duration, wake time, and bedtime are reported in Table 2. First, we found that the late school start policy was associated with longer sleep duration. On average, we observed a 20-minute (0.34 hours, 95% CI, 0.017 to 0.500) increase in sleep duration among adolescents in treatment regions as compared to those in comparison regions. Considering the baseline sleep duration, this represents an increase of 5.3%. This increase reflected later wake times and consistent bedtimes in treatment regions. The average wake time of adolescents in treatment regions increased by approximately 20 minutes (0.320 hours, 95% CI, 0.114 to 0.436) after policy adoption, compared to changes in comparison regions, while the policy did not affect bedtimes overall. Taken together, these findings demonstrated that adolescents' sleep duration increased through delayed wake time after the late school start policy on average across all post-period times.

The event-study analysis for sleep duration, plotted in Figure 1, showed policy impacts at each post-policy period as well as pre-trends. 95% confidence intervals of coefficients of pre-policy periods demonstrated that trends in sleep duration were approximately parallel before policy implementation. It is important to note that we have seven time points in the event-study graphs for sleep outcomes although we have six waves when sleep outcomes were surveyed. This is because the policy implemented in different years across regions as described above. Specifically, two regions implemented the policy in 2014, which resulted in four pre-policy periods and two post-policy periods. The other two regions implemented the policy in 2015, which resulted in five pre-policy periods and one post-policy periods (See Supplement for event-study design specification).

Immediately after policy adoption (i.e., years relative to policy adoption = 0), sleep duration increased by 19 minutes (0.313 hours, 95% CI, 0.092 to 0.534) among students in treatment regions, relative to the difference in sleep duration between groups at baseline (i.e., one year before policy adoption). Event-study results of wake time and bedtime (See Supplemental Figure 2-3) showed that this increase in sleep duration was driven by a 22-minute (0.365 hours, 95% CI, 0.227 to 0.502) delayed wake time and consistent bedtime. However, one year after policy implementation, the observed difference in sleep duration between treatment and comparison regions shrank to 7 minutes (0.111 hours, 95% CI, -0.210 to 0.431) and became statistically non-significant at $p<.05$. This is because bedtime was delayed by 21 minutes (0.351 hours, 95% CI, 0.054 to 0.648) among students in treatment regions, although delayed wake time persisted across two periods. Event-study results restricted to early adopters as well as those with alternative estimators(Callaway & Sant'Anna, 2021; Sun & Abraham, 2021) and the wild bootstrap technique show consistent point estimates and confidence intervals (See Supplemental Figure 1-3 and Supplemental Table 5).

Effects of later school start times on smoking and drinking

DD estimates of the overall impacts of later school start times on adolescents' smoking and drinking were reported in Table 2. First, we found that later school start times were associated with reduced monthly frequency of smoking and drinking. We observed a decrease of 5.32 (95% CI, -13.594 to 0.966) in monthly smoking frequency and of 0.207 (95% CI, -0.346 to -0.097) in monthly drinking frequency.

Event-study estimates of substance use outcomes are presented in Figure 2 and Supplemental Figure 4-5. Again, due to a differential treatment timing, we have six time points in the event-study graphs for substance use outcomes although we have five waves when sleep outcomes were surveyed. Event-study plots showed pre-policy parallel trends in monthly smoking and drinking frequencies. Consistent with sleep outcomes, effects on smoking and drinking outcomes were short-lived. In the year of policy adoption, we observed substantial decreases in monthly smoking (b=-7.422, 95% CI, -13.949 to -0.894) and drinking (b=-0.195, 95% CI, -0.279 to -0.110) frequencies among adolescents in treatment regions vis-à-vis comparison regions. However, the effects shrank a year after the policy implementation. Event-study results restricted to early adopters as well as those with alternative estimators(Callaway & Sant'Anna, 2021; Sun & Abraham, 2021) and wild bootstrap technique show reasonably consistent point estimates and confidence intervals (See Supplemental Figure 4-5 and Supplemental Table 5).

Discussion

This study explored whether a policy that delayed school start times affected adolescents' sleep and substance use, leveraging school start time policy variations and longitudinal, nationally-representative data in Korea. First, we found a 19-minute increase in sleep duration driven by delayed wake time and stable bedtime right after start times were delayed. However, follow-up a year after the policy showed that the initial 19-minute increase shrank to 7 minutes due to delayed bedtime. Our findings are consistent with previous longitudinal studies that reported a gradual shift in bedtimes as students slowly adjust their bedtime to delayed school start times, despite differences with our sleep duration findings (Thacher & Onyper, 2016; Widome et al., 2020). A longitudinal study in New York showed that the initial increase in sleep duration returned to baseline sleep duration after 8 months, reflecting delayed bedtime (Thacher & Onyper, 2016). However, a longitudinal study in Minnesota reported a 16-minute delay in bedtime after a year but found that increased sleep duration persisted due to further delayed wake time (Widome et al., 2020).

This subsequent shift to later bedtime is a warning that delayed school start times alone might not be sufficient to guarantee enough sleep over time. One potential factor reducing the persistence of the sleep effect might be pressure for academic success among high school students. According to the 2020 National Survey of Adolescents' Health and Lifestyle, 62.9% of students chose "study" as a cause of their sleep deficit (National Youth Poicy Institute, 2020) In this case, students might allocate the additional time afforded by the later start time to studying rather than sleep, which offsets the benefits of later school start times. To address the fade-out, policymakers may use informational interventions, including lessons on sleep hygiene for students and a letter to parents to explain impacts of sleep on health and well-being. Communicating scientific reasons behind the policy could alter sleep cultures and behaviors. That being said, it is important to note that this fade-out effect could be specific to the experience of students in 12th grade and might not manifest among younger students. Future research should investigate differential impacts across grades, the interaction between the policy and other social factors such as academic pressure, sleep culture, electronic device use, and caffeine consumption, and effects of the later time policy combined with informative interventions, such as sleep education. In addition, heterogeneity in immediate and fade-out effects should be explored in future research. For example, the fade-out effect might be prominent among high-achieving students under enormous academic pressure. Supplement Table 6 shows that all students experienced fade-outs regardless of reported academic performance, though we interpret this result with caution given the limited measure and small sub-sample size.

Second, we found that later school start times were associated with reduced monthly smoking and drinking frequencies, which provides causal evidence consistent with prior correlational studies (Owens et al., 2010; Stea et al., 2014; Wahlstrom et al., 2017; Wahlstrom & Owens, 2017; Winsler et al., 2015). Our study expanded the literature on later school start times and substance use in two ways. First, while most prior research focused on a binary measure of substance use (Stea et al., 2014; Wahlstrom et al., 2017; Winsler et al., 2015), our study documented the impacts on the frequency of use. Second, we examined longer-term effects on substance use. Third, our study suggests that sleep duration may be a mechanism that links school start time policies and substance use. Importantly, we found the same fade-out effect for substance use as was seen for sleep duration; monthly frequency of substance use rebounded to the usual baseline a year after policy implementation. This finding is consistent with prior research demonstrating a close link between sleep and substance use (Claudatos et al., 2019; McKnight-Eily et al., 2011; Nguyen-Louie et al., 2018; Terry-McElrath et al., 2016; Winsler et al., 2015).

Lastly, our study expanded the literature on later school start times in two additional ways. First, we provide more assurance that effects are not due to individual factors or temporal trends, using a quasi-experimental research design. In addition, the generalizability of prior findings across a range of locations was restricted due to the use of small, non-representative, single-site (mostly western) samples. By using a nationally representative sample of Korean adolescents, which represents a geographically and demographically diverse population, this study highlights the benefits of delayed school start times and various outcomes in a non-western society.

Limitations

Our study had several limitations. First, findings were based on self-reported measures, and thus future research should corroborate these findings using objective measures of sleep and substance use. However, our findings are consistent with literature using objective measures of sleep (Widome et al., 2020). Second, because we did not have information on each school's start time before the policy implementation and only have regional-level summary information, our estimations did not account for treatment “dosage”. Relatedly, if there were voluntary start time delays in comparison regions, it might introduce downward bias into our estimates. However, we minimized this concern by excluding comparison regions that allowed schools to independently decide on later start times. Moreover, given the highly centralized education system in Korea (Bodovski et al., 2017), a lack of fidelity (e.g., schools not adhering to mandated start time delays) is less of a concern. Fourth, due to the nature of longitudinal data and the specific timing of policy and survey implementation, our analysis does not isolate the fade-out effect from “12th grader” effects. In other words, the disappearance of policy benefits observed a year after implementation may partly stem from the lack of policy effect on 12th graders. Lastly, although we leveraged a difference-in-differences design to isolate the effects of later school start time from unobserved confounders, we acknowledge that unobserved time-varying variables, such as changes in sleep culture only in treated regions at the policy implementation period, which are associated with both outcomes and policy adoption, could bias the estimates.

Conclusion

This study provides quasi-experimental evidence for impacts of a later school start time policy on adolescents’ sleep, smoking, and drinking, using nationally representative data from Korea. We demonstrated that policies that delay school start times can have health benefits beyond increases in sleep duration—getting more sleep can reduce unhealthy behaviors for adolescents. However, our results also serve as a warning that, without measures to ensure the persistence of increased sleep, the health benefits of delayed start times may be relatively short-lived. Our study has insights on how adolescents’ health behaviors are shaped within the context of a school start time so that physicians and policymakers across Global regions can better understand the underlying processes of sleep deficit and substance use.

References

Adolescent Sleep Working Group, Committee on Adolescence, & Council on School Health. (2014). School start times for adolescents. *Pediatrics*, *134*(3), 642–649. https://doi.org/10.1542/peds.2014-1697

Baek, B., Cho, Y., & Woo, O. (2015). *An analysis of school start at 9*. Gyeonggi Institute of Education.

Bastian, K. C., & 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. https://doi.org/10.1086/723063

Beebe, D. W. (2011). Cognitive, Behavioral, and Functional Consequences of Inadequate Sleep in Children and Adolescents. *Pediatric Clinics of North America*, *58*(3), 649–665. https://doi.org/10.1016/j.pcl.2011.03.002

Berger, A. T., Widome, R., & Troxel, W. M. (2019). Chapter 33—Delayed school start times and adolescent health. In M. A. Grandner (Ed.), *Sleep and Health* (pp. 447–454). Academic Press. https://doi.org/10.1016/B978-0-12-815373-4.00033-2

Biller, A. M., Meissner, K., Winnebeck, E. C., & Zerbini, G. (2022). School start times and academic achievement—A systematic review on grades and test scores. *Sleep Medicine Reviews*, *61*, 101582. https://doi.org/10.1016/j.smrv.2021.101582

Bodovski, K., Byun, S., Chykina, V., & Chung, H. J. (2017). Searching for the golden model of education: Cross-national analysis of math achievement. *Compare: A Journal of Comparative and International Education*, *47*(5), 722–741. https://doi.org/10.1080/03057925.2016.1274881

Bowers, J. M., & 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. https://doi.org/10.1016/j.sleh.2017.08.004

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

Carrell, S. E., Maghakian, T., & 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. https://doi.org/10.1257/pol.3.3.62

Claudatos, S., Baker, F. C., & Hasler, B. P. (2019). Relevance of Sleep and Circadian Rhythms to Adolescent Substance Use. *Current Addiction Reports*, *6*(4), 504–513. https://doi.org/10.1007/s40429-019-00277-9

Crowley, S. J., Wolfson, A. R., Tarokh, L., & Carskadon, M. A. (2018). An update on adolescent sleep: New evidence informing the perfect storm model. *Journal of Adolescence*, *67*, 55–65. https://doi.org/10.1016/j.adolescence.2018.06.001

Edwards, S., Reeves, G. M., & Fishbein, D. (2015). Integrative Model of the Relationship Between Sleep Problems and Risk for Youth Substance Use. *Current Addiction Reports*, *2*(2), 130–140. https://doi.org/10.1007/s40429-015-0052-0

Fuligni, A. J., Bai, S., Krull, J. L., & Gonzales, N. A. (2019). Individual Differences in Optimum Sleep for Daily Mood During Adolescence. *Journal of Clinical Child & Adolescent Psychology*, *48*(3), 469–479. https://doi.org/10.1080/15374416.2017.1357126

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Groen, J. A., & Pabilonia, S. W. (2019). Snooze or lose: High school start times and academic achievement. *Economics of Education Review*, *72*, 204–218. https://doi.org/10.1016/j.econedurev.2019.05.011

Heissel, J. A., & 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. https://doi.org/10.3368/jhr.53.4.0815-7346R1

Hinrichs, P. (2011). When the Bell Tolls: The Effects of School Starting Times on Academic Achievement. *Education Finance and Policy*, *6*(4), 486–507. https://doi.org/10.1162/EDFP_a_00045

James, S. A., Erickson, D. J., Lammert, S., & Widome, R. (2023). School start time delays and high school educational outcomes: Evidence from the START/LEARN study. *Journal of Adolescence*, *n/a*(n/a). https://doi.org/10.1002/jad.12151

Lo, J. C., Lee, S. M., Lee, X. K., Sasmita, K., Chee, N. I. Y. N., Tandi, J., Cher, W. S., Gooley, J. J., & Chee, M. W. L. (2018). Sustained benefits of delaying school start time on adolescent sleep and well-being. *Sleep*, *41*(6), zsy052. https://doi.org/10.1093/sleep/zsy052

Marx, R., Tanner-Smith, E. E., Davison, C. M., Ufholz, L.-A., Freeman, J., Shankar, R., Newton, L., Brown, R. S., Parpia, A. S., Cozma, I., & Hendrikx, S. (2017). Later school start times for supporting the education, health, and well-being of high school students: A systematic review. *Campbell Systematic Reviews*, *13*(1), 1–99. https://doi.org/10.4073/csr.2017.15

McKnight-Eily, L. R., Eaton, D. K., Lowry, R., Croft, J. B., Presley-Cantrell, L., & Perry, G. S. (2011). Relationships between hours of sleep and health-risk behaviors in US adolescent students. *Preventive Medicine*, *53*(4–5), 271–273. https://doi.org/10.1016/j.ypmed.2011.06.020

Minges, K. E., & Redeker, N. S. (2016). Delayed school start times and adolescent sleep: A systematic review of the experimental evidence. *Sleep Medicine Reviews*, *28*, 86–95. https://doi.org/10.1016/j.smrv.2015.06.002

Ministry of Gender Equality and Family. (2022). *2022 adolescent statistics*. Ministry of Gender Equality and Family.

Morgenthaler, T. I., Hashmi, S., Croft, J. B., Dort, L., Heald, J. L., & Mullington, J. (2016). High School Start Times and the Impact on High School Students: What We Know, and What We Hope to Learn. *Journal of Clinical Sleep Medicine : JCSM : Official Publication of the American Academy of Sleep Medicine*, *12*(12), 1681–1689. https://doi.org/10.5664/jcsm.6358

National Youth Poicy Institute. (2020). *National survey of Adolescents' health and lifestyle*. National Youth Poicy Intitute.

Nguyen-Louie, T. T., Brumback, T., Worley, M. J., Colrain, I. M., Matt, G. E., Squeglia, L. M., & Tapert, S. F. (2018). Effects of sleep on substance use in adolescents: A longitudinal perspective. *Addiction Biology*, *23*(2), 750–760. https://doi.org/10.1111/adb.12519

Owens, J. A., Belon, K., & Moss, P. (2010). Impact of Delaying School Start Time on Adolescent Sleep, Mood, and Behavior. *Archives of Pediatrics & Adolescent Medicine*, *164*(7). https://doi.org/10.1001/archpediatrics.2010.96

Owens, J. A., Dearth-Wesley, T., Herman, A. N., Oakes, J. M., & 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. https://doi.org/10.1016/j.sleh.2017.09.001

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal: Promoting Communications on Statistics and Stata*, *19*(1), 4–60. https://doi.org/10.1177/1536867X19830877

Seo, D. (2020). *Report on adolecent sleep, breakfast, and school start time*. Office of legislator Dongyong Seo.

Shochat, T., Cohen-Zion, M., & Tzischinsky, O. (2014). Functional consequences of inadequate sleep in adolescents: A systematic review. *Sleep Medicine Reviews*, *18*(1), 75–87. https://doi.org/10.1016/j.smrv.2013.03.005

Short, M. A., Weber, N., Reynolds, C., Coussens, S., & Carskadon, M. A. (2018). Estimating adolescent sleep need using dose-response modeling. *Sleep*, *41*(4). https://doi.org/10.1093/sleep/zsy011

Stea, T. H., Knutsen, T., & Torstveit, M. K. (2014). Association between short time in bed, health-risk behaviors and poor academic achievement among Norwegian adolescents. *Sleep Medicine*, *15*(6), 666–671. https://doi.org/10.1016/j.sleep.2014.01.019

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

Terry-McElrath, Y. M., Maslowsky, J., O'Malley, P. M., Schulenberg, J. E., & Johnston, L. D. (2016). Sleep and Substance Use among US Adolescents, 1991–2014. *American Journal of Health Behavior*, *40*(1), 77–91. https://doi.org/10.5993/AJHB.40.1.9

Thacher, P. V., & Onyper, S. V. (2016). Longitudinal Outcomes of Start Time Delay on Sleep, Behavior, and Achievement in High School. *Sleep*, *39*(2), 271–281. https://doi.org/10.5665/sleep.5426

The Hankyoreh. (2014, August). "I had breakfast". The first scene of Uijeongbu Girls' Middle School at 9 o'clock. *The Hankyoreh*. https://www.hani.co.kr/arti/society/society_general/652582.html

Wahlstrom, K. L. (2014). *Examining the Impact of Later High School Start Times on the Health and Academic Performance of High School Students: A Multi-Site Study*. https://conservancy.umn.edu/bitstream/handle/11299/162769/Impact%20of%20Later%20Start%20Time%20Final%20Report.pdf?sequence=1&isAllowed=y.pdf

Wahlstrom, K. L., Berger, A. T., & Widome, R. (2017). Relationships between school start time, sleep duration, and adolescent behaviors. *Sleep Health*, *3*(3), 216–221. https://doi.org/10.1016/j.sleh.2017.03.002

Wahlstrom, K. L., & Owens, J. A. (2017). School start time effects on adolescent learning and academic performance, emotional health and behaviour. *Current Opinion in Psychiatry*, *30*(6), 485–490. https://doi.org/10.1097/YCO.0000000000000368

Wheaton, A. G., Ferro, G. A., & Croft, J. B. (2015). School Start Times for Middle School and High School Students—United States, 2011–12 School Year. *Morbidity and Mortality Weekly Report*, *64*(30), 809–813.

Whitaker, R. C., Dearth-Wesley, T., Herman, A. N., Oakes, J. M., & Owens, J. A. (2019). A quasi-experimental study of the impact of school start time changes on adolescents' mood, self-regulation, safety, and health. *Sleep Health*, *5*(5), 466–469. https://doi.org/10.1016/j.sleh.2019.06.011

Widome, R., Berger, A. T., Iber, C., Wahlstrom, K., Laska, M. N., Kilian, G., Redline, S., & Erickson, D. J. (2020). Association of Delaying School Start Time With Sleep Duration, Timing, and Quality Among Adolescents. *JAMA Pediatrics*, *174*(7), 697. https://doi.org/10.1001/jamapediatrics.2020.0344

Winsler, A., Deutsch, A., Vorona, R. D., Payne, P. A., & Szklo-Coxe, M. (2015). Sleepless in Fairfax: The Difference One More Hour of Sleep Can Make for Teen Hopelessness, Suicidal Ideation, and Substance Use. *Journal of Youth and Adolescence*, *44*(2), 362–378. https://doi.org/10.1007/s10964-014-0170-3

Yang, D., & 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), 7–29. https://doi.org/10.1080/17441730.2020.1794312

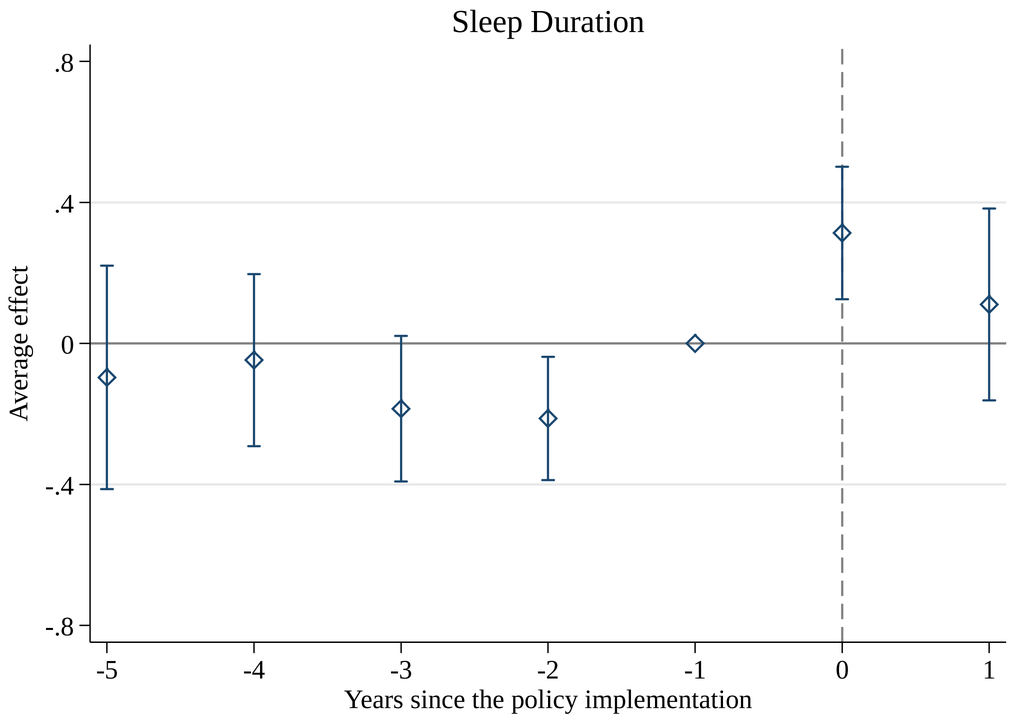


Figure 1. Event-study estimates of the effects of a later school start policy on sleep duration among adolescents (7th-12th grades) in South Korea, 2010-2015.
NOTE. Each point represents the point estimate from TWFE event-study specification, and the coefficient for the one time period before the policy is normalized to zero. Each bar represents the 95% confidence interval, calculated with standard errors clustered at the region level. Point estimates from TWFE and alternative estimators (Callaway & SantAnna, 2021; Sun and Abraham 2021) and cluster-robust and wild bootstrap confidence intervals are reported in Supplemental Figure 1.

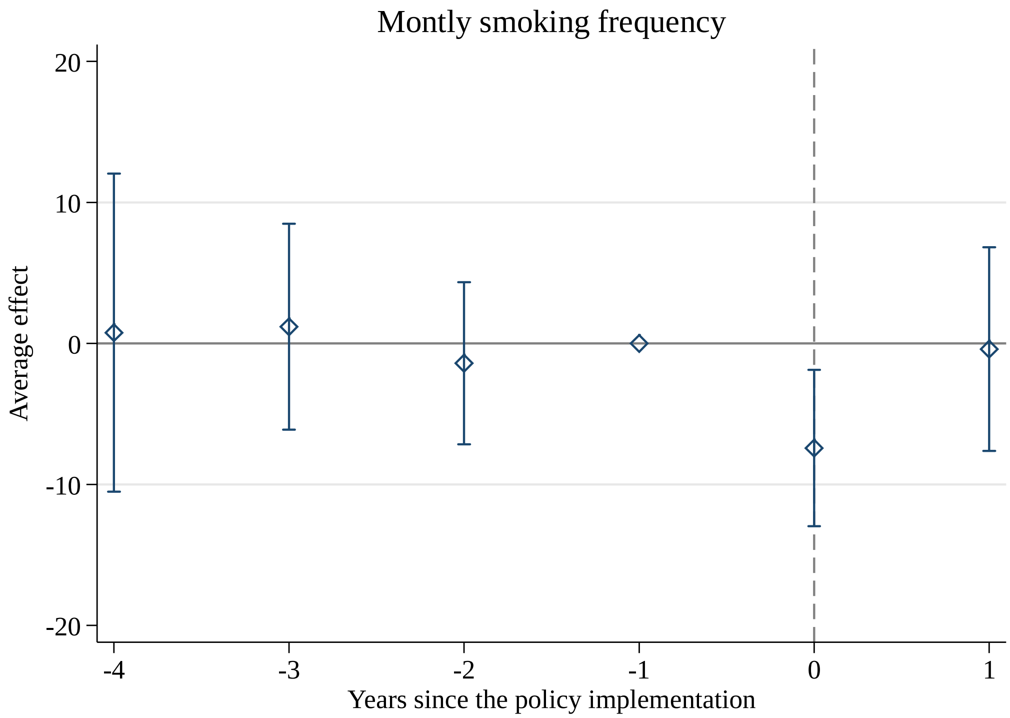


Figure 2. Event-study estimates of the effects of a later school start policy on monthly smoking frequency among adolescents (7th-12th grades) in South Korea, 2010-2015.
NOTE. Each point represents the point estimate from TWFE event-study specification, and the coefficient for the one time period before the policy is normalized to zero. Each bar represents the 95% confidence interval, calculated with standard errors clustered at the region level. Point estimates from TWFE and alternative estimators (Callaway & SantAnna, 2021; Sun and Abraham 2021) and cluster-robust and wild bootstrap confidence intervals are reported in Supplemental Figure 4.

Table 1. Descriptive statistics of sleep and substance use outcomes and sociodemographic characteristics by treatment condition for 10th graders in South Korea, 2013

	All regions		Treatment regions		Comparison regions	
	Mean	SD	Mean	SD	Mean	SD
Outcomes						
Sleep duration (hour:minute)	6:18	1:05	6:23	1:08	6:13	1:01
Wake time (hour:minute)	6:36	0:31	6:32	0:32	6:40	0:29
Bedtime (hour:minute)	0:18	1:02	0:09	1:04	0:26	0:59
Number of times smoking per month	5.805	35.347	5.859	33.696	5.757	36.806
Number of times drinking per month	0.102	0.536	0.086	0.422	0.116	0.622
Individual Characteristics						
Female	0.506	—	0.561	—	0.456	—
Highest parental education	14.451	2.257	14.414	2.281	14.486	2.236
Annual income	4553.003	2611.216	4380.471	2798.711	4709.571	2420.468
Two-parent household	0.897	—	0.878	—	0.915	—
Parental working status	0.326	—	0.308	—	0.343	—
N	1070~1072		508~510		562	

NOTE. Unweighted descriptive statistics of the analytic sample for each outcome in 2013 (i.e., the year before the first policy implementation) are reported. For time-varying covariates, the analytic sample for sleep duration outcome is used. The unit of sleep duration, wake time, and bedtime is hours. The highest parental education is measured as years of schooling, which is a continuous variable. Parental working status shows the percentage of students with at least one parent working currently. Weighted descriptive statistics are reported in Supplemental Table 1.

Table 2. Estimated effects of a later school start policy on sleep and substance use among adolescents (11th and 12th grades) in South Korea, 2014-2015

	Sleep duration (hours)	Wake time (hours)	Bedtime (hours)	Monthly smoking frequency	Monthly drinking frequency
Later school start time policy	0.340	0.320	-0.021	-5.317	-0.207
Cluster robust 95% CI	[0.149,0.532]**	[0.182,0.457]***	[-0.135,0.093]	[-11.521,0.887]~	[-0.314,-0.100]**
Wild bootstrap 95% CI	[0.017,0.500]*	[0.114,0.436]*	[-0.154,0.144]	[-13.594,0.966]~	[-0.346,-0.097]**
N	6324	6324	6324	5279	5279
R^2 *(Within-student)*	0.448	0.184	0.366	0.023	0.043

NOTE. Robust standard errors are clustered at the region level. Wild bootstrap 95% confidence intervals are also presented. Individual fixed effects and year fixed effects are included in all models. The unit of sleep duration, wake time, and bedtime is hours. ~ p<.1, * p < .05, ** p < .01, *** p < .001

Supplemental Online Content

This Supplemental Online Content provides additional details on the methods described in the main paper.

Method

Data and Sampling

Based on the thorough sampling frame, which was a full list of schools and students provided by the Department of Education, KCYPS used a stratified multi-stage cluster sampling design. Specifically, (1) the population was stratified into 27 areas, (2) middle schools were randomly sampled within a region using probability proportional to size (PPS) sampling, and (3) 7th-grade classes were randomly sampled within selected schools. Students were interviewed in school in the baseline survey, and each student was traced separately from the second year survey. The initial response rate was over 80 percent, and 87.5 percent of the original respondents participated in the first to sixth waves. To account for the complex survey design and non-random attrition across waves, we used sampling weights which were computed by NYPI. The analytic sample was limited to students in either treatment regions or comparison regions, and students who moved across regions across waves were excluded.

Difference-in-Differences Analysis

We employed a difference-in-differences (DD) design leveraging variation in a delayed school start time policy adoption across regions (Angrist & Pischke, 2009). To be specific, a two-way fixed effects model with the following specification was used:

$$Y_{ist} = \beta_0 + \beta_1 Late\ School\ Start_{st} + \alpha_i + \lambda_t + \eta X_{ist} + \varepsilon_{ist}$$

Where Y_{ist} is the outcome—sleep duration, wake time, bedtime, irritability, the number of smoking, and the number of drinking—of student *i* who attends a school in region *s* at each period *t*. $Late\_School\_Start_{st}$ is a binary variable equal to 1 if a student *i* attends a school in a region *s* that adopted the late school start policy at period *t*, and zero otherwise. Therefore, β_1, the parameter of interest, represents the effect of late school start policy implementation. α_i is a set of individual fixed effects, and λ_t is a set of year fixed effects. Since students who moved across regions were removed, there is no region-level time invariant variation beyond and above the individual-level variation. Therefore, region fixed effects are omitted in the model. X_{ist} is a vector of time-varying individual-level covariates. We presented both models with and without time-varying individual-level covariates.

Event-study Analysis and staggered treatment timing

We conducted event-study analyses of sleep and substance use outcomes to test whether the effect of delayed school changes over time (Angrist & Pischke, 2009). In this model specification, we include pre-policy years and distinct post-policy years instead of one treatment indicator to allow the effect to change over time:

$$Y_{ist} = \beta_0 + \sum_{r \neq -1} 1[t = G_s + r]\beta_r + \alpha_i + \lambda_t + \varepsilon_{ist}$$

Where Y_{ist} is the outcome—sleep duration, wake time, bedtime, irritability, monthly smoking frequency, and monthly drinking frequency—of student *i* who attends a school in region *s* at each period *t*. The term $1[t = G_s + r]$ represents a set of indicators for each time period relative to treatment, with G_s denoting the year in which region s implemented the later school start time policy and $r \in [-4,1]$ for sleep outcomes and $r \in [-3,1]$ for substance use outcomes. Therefore, the parameters of interest are the β_r's, which represent the effects of a late school start policy relative to a year before the policy adoption (r = -1). α_i is a set of individual fixed effects, and λ_t is a set of year fixed effects. Since students who moved across regions were removed, there is no region-level time invariant variation beyond and above the individual-level variation. Therefore, region fixed effects are omitted in the model. In addition, we estimated a model with a set of school-fixed effects for wake time and bedtime outcomes, concerning potential violation of the parallel trend assumption. After including school-fixed effects, in support of the parallel trends assumption, 95% confidence intervals of coefficients of all implementation periods except the -4 period for wake time outcome include zero.

As we discussed in the main paper, we incorporated recent advances in the econometrics literature to address potential bias in the two-way fixed effect approach (Callaway & Sant'Anna, 2021; de Chaisemartin & D'Haultfoeuille, 2022; de Chaisemartin & D'Haultfœuille, 2020; Gardner, 2022; Goodman-Bacon, 2021; Roth et al., 2022; Sun & Abraham, 2021; Wooldridge, 2021). Among recently developed estimators for staggered timing, we conducted Callaway and Sant'Anna (2021)'s estimation and Sun and Abraham (2021)'s estimation to examine the robustness of findings from our TWFE model. For these analyses, we used *csdid* and *eventstudyinteract* commands in STATA (Rios-Avila et al., 2023; Sun, 2022). Since *csdid* has a wild bootstrap cluster option, we used the wild cluster bootstrap technique to compute our standard errors, which enables us to address the issue of underestimated cluster-robust standard errors with a small number of clusters. Specifically, we reported wild cluster bootstrap confidence intervals with 999 replications and Webb weights, which are proper when the number of replications exceeds two raised to the number of clusters (in this case, 2^9 = 512) (Roodman, 2023; Roodman et al., 2019).

Goodman-bacon decomposition

We conducted the Goodman-Bacon decomposition for sleep and substance use outcomes that have heterogeneity in treatment timing (Goodman-Bacon, 2021; Goodman-Bacon et al., 2022). Since TWFE estimator within the DD framework is a weighted average of all possible 2X2 DD estimates, understanding how coefficients from each comparison are weighted is important when there is differential treatment timing. Goodman-Bacon decomposition computes weights and coefficients for each potential 2X2 comparison and reports coefficients and weights for these three comparisons: (1) the early-treated group as a comparison group for the late-treated group (i.e., forbidden comparison), (2) the late-treated group as a comparison group for the early-treated group, and (3) never-treated group as a comparison group. We presented Goodman-bacon decomposition results in Supplemental Table 3. The decomposition shows that more than 85 percent of our DD estimates come from the comparison between never-treated and treatment for all outcomes. Approximately 11 percent of our DD estimates were from forbidden comparison, which is early group control vs. late group treatment.

A List of Online-only Tables and Figures

13. Supplemental Figure 7. Event-study estimates of the effects of a later school start policy on bedtime with school-fixed effects

Supplemental Table 1. Weighted Descriptive statistics in 2013 by a treatment condition

	All regions		Treatment regions		Comparison regions	
	Mean	SD	Mean	SD	Mean	SD
Outcomes						
Sleep duration (hour:minute)	6:21	1:05	6:25	1:05	6:15	1:03
Wake time (hour:minute)	6:35	0:30	6:32	0:32	6:40	0:29
Bedtime (hour:minute)	0:14	1:02	0:07	1:02	0:25	1:01
Number of times smoking per month	6.102	36.017	5.901	33.349	6.394	39.599
Number of times drinking per month	0.115	0.615	0.084	0.408	0.159	0.825
Individual Characteristics						
Female	0.489	—	0.499	—	0.476	—
Highest parental education	14.516	2.240	14.504	2.267	14.532	2.203
Annual income	4535.524	2742.746	4490.442	2994.253	4601.287	2329.942
Two-parent household	0.891	—	0.880	—	0.906	—
Parental working status	0.403	—	0.383	—	0.431	—
N	1070~1072		508~510		562	

NOTE. Weighted descriptive statistics of the analytic sample for each outcome in 2013 (i.e., the year before the first policy implementation) are reported. For time-varying covariates, the analytic sample for sleep duration outcome is used. The unit of sleep duration, wake time, and bedtime is hours. The highest parental education is measured as years of schooling, which is a continuous variable. Parental working status shows the percentage of students with at least one parent working currently. Unweighted descriptive statistics are reported in Table 1.

Supplemental Table 2. Estimated effects of a later school start policy on sleep and substance use

	Sleep duration (hours)	Wake time (hours)	Bedtime (hours)	Monthly smoking frequency	Monthly drinking frequency
Later school start time policy	0.337	0.319	-0.018	-5.282	-0.209
Cluster robust 95% CI	[0.149,0.525]**	[0.182,0.455]***	[-0.133,0.096]	[-11.498,0.933]~	[-0.315,-0.102]**
Wild bootstrap 95% CI	[0.020, 0.493]*	[0.114, 0.434]*	[-0.152, 0.141]	[-13.576, 1.003]~	[-0.346, -0.098]**
N	6324	6324	6324	5279	5279
R^2 *(Within-student)*	0.449	0.185	0.366	0.024	0.044

NOTE. Robust standard errors are clustered at the region level. Wild bootstrap 95% confidence intervals are also presented. Individual fixed effects and year fixed effects are included in all models. Time-varying individual covariates include annual income, parental working status, and family structure. The unit of sleep duration, wake time, and bedtime is hours. ~ p<.1, * p < .05, ** p < .01, *** p < .001

Supplemental Table 3. Goodman-Bacon Decomposition

	Sleep duration (hours)		Wake time (hours)		Bedtime (hours)		Monthly smoking frequency		Monthly drinking frequency	
	Beta	Weight	Beta	Weight	Beta	Weight	Beta	Weight	Beta	Weight
Early group control vs. late group treatment	0.490	0.111	0.424	0.111	-0.065	0.111	-7.358	0.110	-0.162	0.110
Late group control vs. Early group treatment	0.340	0.028	0.216	0.028	-0.124	0.028	-15.118	0.036	-0.283	0.036
Never treated vs. treatment	0.317	0.861	0.309	0.861	-0.008	0.861	-4.569	0.854	-0.209	0.854
Weighted DD	0.340		0.320		-0.021		-5.710		-0.187	

NOTE. Individual fixed effects and year fixed effects are included in all models, and time-varying individual covariates are not included in all models. Therefore, it provides decompositions of our main DD models presented in Table 2. The unit of sleep duration, wake time, and bedtime is hours.

Supplemental Table 4. Full models for event-study estimates of the effects of a later school start policy on sleep and substance use

	Two-way fixed effect		Callaway and Sant'Anna (2020)		Sun and Abraham (2020)	
	Sleep duration (hours)					
5 years before	-0.096	[-0.469,0.277]	-0.217	[-0.771,0.336]	-0.231	[-0.804,0.342]
4 years before	-0.047	[-0.334,0.240]	-0.079	[-0.318,0.160]	-0.119	[-0.423,0.185]
3 years before	-0.185	[-0.428,0.058]	-0.224*	[-0.397,-0.051]	-0.245*	[-0.485,-0.006]
2 years before	-0.213*	[-0.418,-0.007]	-0.278***	[-0.441,-0.114]	-0.260**	[-0.421,-0.098]
1 year before	Omitted					
Year of policy adoption	0.313*	[0.092,0.534]	0.321*	[0.019,0.624]	0.317*	[0.096,0.538]
1 year after	0.111	[-0.210,0.431]	-0.027	[-0.303,0.248]	-0.015	[-0.357,0.327]
N	6324		6130		6324	
R^2	0.451		—		0.535	
	Wake time (hours)					
5 years before	0.184*	[0.045,0.323]	0.168**	[0.045,0.290]	0.153	[0.037,0.270]
4 years before	0.175**	[0.057,0.294]	0.170***	[0.071,0.269]	0.176**	[0.057,0.296]
3 years before	0.131*	[0.037,0.226]	0.135***	[0.061,0.208]	0.132	[0.019,0.244]
2 years before	0.069*	[0.004,0.134]	0.087*	[0.014,0.160]	0.084*	[0.018,0.151]
1 year before	Omitted					
Year of policy adoption	0.365***	[0.227,0.502]	0.376***	[0.257,0.496]	0.378***	[0.260,0.497]
1 year after	0.461***	[0.352,0.571]	0.456***	[0.373,0.539]	0.465***	[0.368,0.562]
N	6324		6130		6324	
R^2	0.191		—		0.427	
	Bedtime (hours)					
5 years before	0.281	[-0.148,0.709]	0.385	[-0.117,0.886]	0.384	[-0.203,0.972]
4 years before	0.223	[-0.122,0.567]	0.249	[-0.048,0.545]	0.296	[-0.074,0.665]
3 years before	0.317*	[0.034,0.599]	0.358**	[0.131,0.586]	0.377	[0.083,0.671]
2 years before	0.282*	[0.037,0.527]	0.365**	[0.136,0.593]	0.344**	[0.150,0.538]
1 year before	Omitted					
Year of policy adoption	0.051	[-0.147,0.249]	0.055	[-0.239,0.348]	0.061	[-0.212,0.334]
1 year after	0.351*	[0.054,0.648]	0.484**	[0.172,0.796]	0.480	[0.144,0.817]
N	6324		6130		6324	
R^2	0.372		—		0.499	
	Monthly smoking frequency					
4 years before	0.763	[-12.509,14.035]	-0.016	[-9.257,9.224]	-0.790	[-19.345,17.764]
3 years before	1.181	[-7.414,9.776]	0.147	[-4.646,4.940]	0.251	[-6.018,6.520]
2 years before	-1.407	[-8.166,5.352]	-2.256	[-6.823,2.311]	-3.027	[-8.707,2.652]
1 year before	Omitted					
Year of policy adoption	-7.422*	[-13.949,-0.894]	-7.575**	[-12.133,-3.016]	-7.593**	[-11.788,-3.399]
1 year after	-0.406	[-8.903,8.092]	-2.743	[-9.475,3.989]	-1.805	[-8.842,5.232]
N	5279		5153		5279	
R^2	0.025		—		0.285	
	Monthly drinking frequency					

4 years before	0.119	[-0.136,0.373]	0.102	[-0.123,0.327]	0.097	[-0.186,0.379]
3 years before	0.063	[-0.111,0.236]	0.085	[-0.049,0.219]	0.063	[-0.120,0.246]
2 years before	0.079	[-0.039,0.198]	0.081	[-0.035,0.198]	0.086	[-0.061,0.233]
1 year before			Omitted			
Year of policy adoption	-0.195***	[-0.279,-0.110]	-0.187***	[-0.259,-0.115]	-0.187***	[-0.263,-0.111]
1 year after	-0.112~	[-0.232,0.008]	-0.126*	[-0.232,-0.020]	-0.109	[-0.223,0.005]
N	5279		5153		5279	
R^2	0.044		—		0.186	

NOTE. Point estimates and 95% confidence intervals, in brackets, from three event-study estimations are reported: standard two-way fixed effect, Callaway and Anna (2020), and Sun and Abraham (2020). Robust standard errors are clustered at the region level. For Callaway and Sant'Anna (2020) estimations, wild bootstrap 95% confidence intervals are presented. Models for Callaway and Sant'Anna (2020) estimations have different analytic sample sizes because it only uses observations with pair balanced, and it does not provide R-squared. The unit of sleep duration, wake time, and bedtime is hours. ~ p<.1, * p < .05, ** p < .01, *** p < .001

Supplemental Table 5. Event-study estimates of the effects of a later school start policy on sleep and substance use among the early and late adopters, separately

	Sleep duration (hours)		Wake time (hours)		Bedtime (hours)		Monthly smoking frequency		Monthly drinking frequency	
Early adopters										
4 years before	-0.124	[-0.484,0.237]	0.189*	[0.060,0.317]	0.312	[-0.120,0.745]	—		—	
3 years before	-0.257$^{\sim}$	[-0.548,0.034]	0.125$^{\sim}$	[-0.006,0.257]	0.382*	[0.029,0.736]	1.195	[-3.715,6.106]	0.061	[-0.142,0.264]
2 years before	-0.282*	[-0.527,-0.037]	0.107*	[0.010,0.205]	0.389*	[0.116,0.662]	-4.048$^{\sim}$	[-8.642,0.546]	0.099	[-0.110,0.307]
1 year before						Omitted				
Year of policy adoption	0.483*	[0.063,0.904]	0.408***	[0.255,0.561]	-0.076	[-0.562,0.411]	-5.840*	[-10.802,-0.877]	-0.172*	[-0.295,-0.049]
1 year after	-0.015	[-0.385,0.355]	0.465***	[0.360,0.570]	0.480*	[0.117,0.844]	-1.805	[-9.408,5.798]	-0.109$^{\sim}$	[-0.232,0.015]
N	5316		5316		5316		4438		4438	
R^2	0.449		0.181		0.379		0.027		0.043	
Late adopters										
5 years before	-0.231	[-0.851,0.389]	0.153*	[0.027,0.279]	0.384	[-0.251,1.019]	—		—	
4 years before	-0.105	[-0.708,0.498]	0.137	[-0.060,0.335]	0.242	[-0.314,0.798]	-0.790	[-20.839,19.259]	0.097	[-0.209,0.402]
3 years before	-0.209	[-0.695,0.277]	0.151*	[0.009,0.293]	0.360$^{\sim}$	[-0.073,0.794]	-2.699	[-24.183,18.784]	0.070	[-0.238,0.378]
2 years before	-0.190	[-0.584,0.203]	0.012	[-0.036,0.060]	0.202	[-0.184,0.589]	0.123	[-19.977,20.223]	0.048	[-0.217,0.313]
1 year before						Omitted				
Year of policy adoption	-0.188	[-0.827,0.451]	0.288*	[0.031,0.545]	0.476$^{\sim}$	[-0.108,1.060]	-12.920$^{\sim}$	[-27.028,1.188]	-0.233*	[-0.421,-0.045]
N	4337		4337		4337		3642		3642	
R^2	0.506		0.202		0.41		0.024		0.05	

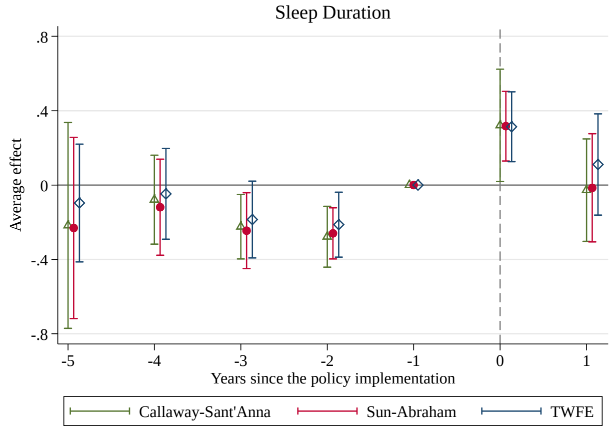
NOTE. Point estimates and 95% confidence intervals, in brackets, from two-way fixed effect estimations were reported. Robust standard errors are clustered at the region level. The unit of sleep duration, wake time, and bedtime is hours. ~ p<.1, * p < .05, ** p < .01, *** p < .001

Supplemental Table 6. Event-study estimates of the effects of a later school start policy on sleep duration among high-achievers and non-high-achievers, separately

	High-achievers		Non high-achievers	
5 years before	0.084	[-0.857,1.024]	-0.110	[-0.523,0.302]
4 years before	0.282	[-0.585,1.150]	-0.076	[-0.404,0.251]
3 years before	0.068	[-0.476,0.612]	-0.215	[-0.494,0.065]
2 years before	-0.230	[-0.574,0.113]	-0.209	[-0.433,0.014]
1 year before	Omitted		Omitted	
Year of policy adoption	0.785	[0.176,1.395]	0.280	[0.058,0.501]
1 year after	0.452	[-0.206,1.111]	0.087	[-0.239,0.413]
N	371		5889	
R^2	0.510		0.451	

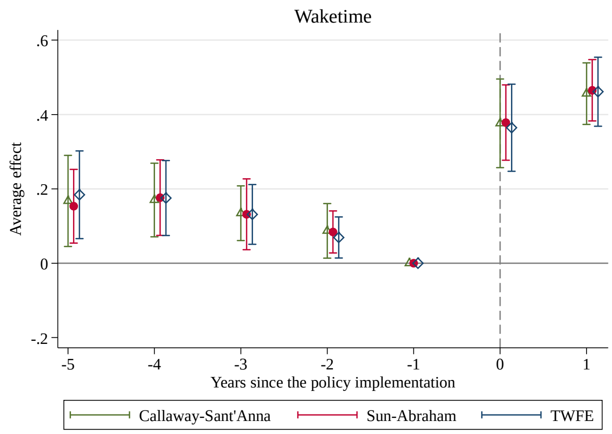
NOTE. Point estimates and 95% confidence intervals, in brackets, from two-way fixed effect estimations were reported. Robust standard errors are clustered at the region level. The unit of sleep duration is hours. ~ p<.1, * p < .05, ** p < .01, *** p < .001

Supplemental Figure 1. Event-study estimates of the effects of a later school start policy on sleep duration using three different estimators



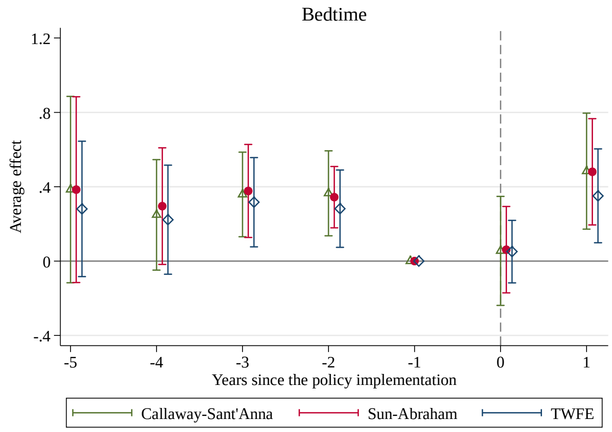
NOTE. Each point represents the point estimate from each event-study estimator, and the coefficient for the one time period before the policy is normalized to zero. Each bar represents the 95% confidence interval, calculated with standard errors clustered at the region level. For Callaway and Sant'Anna (2020) estimations, wild bootstrap 95% confidence intervals are presented. The unit of sleep duration is hours.

Supplemental Figure 2. Event-study estimates of the effects of a later school start policy on wake time using three different estimators



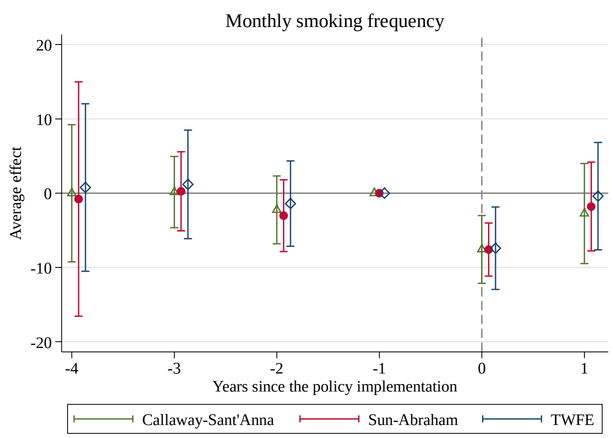
NOTE. Each point represents the point estimate from each event-study estimator, and the coefficient for the one time period before the policy is normalized to zero. Each bar represents the 95% confidence interval, calculated with standard errors clustered at the region level. For Callaway and Sant'Anna (2020) estimations, wild bootstrap 95% confidence intervals are presented. The unit of wake time is hours.

Supplemental Figure 3. Event-study estimates of the effects of a later school start policy on bedtime using three different estimators



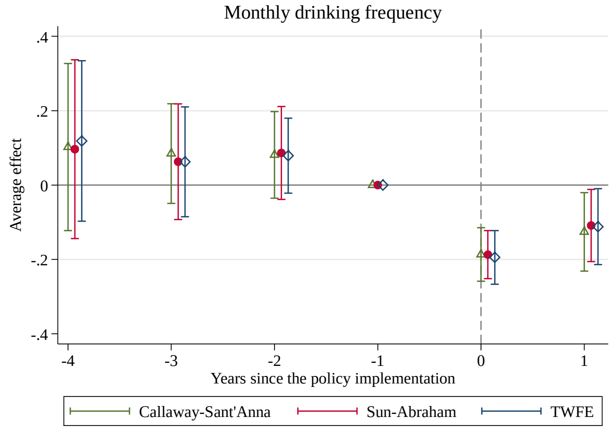
NOTE. Each point represents the point estimate from each event-study estimator, and the coefficient for the one time period before the policy is normalized to zero. Each bar represents the 95% confidence interval, calculated with standard errors clustered at the region level. For Callaway and Sant'Anna (2020) estimations, wild bootstrap 95% confidence intervals are presented. The unit of bedtime is hours.

Supplemental Figure 4. Event-study estimates of the effects of a later school start policy on the number of smoking using three different estimators



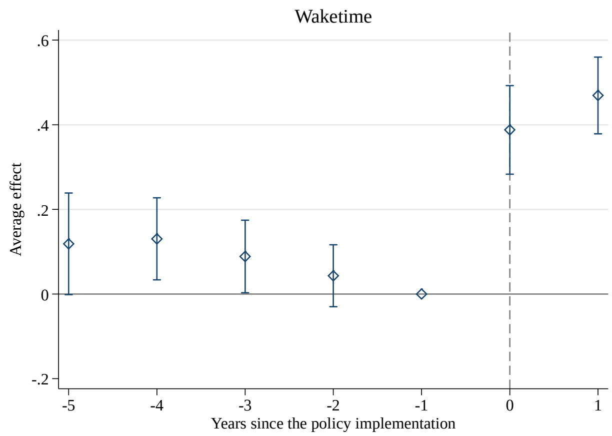
NOTE. Each point represents the point estimate from each event-study estimator, and the coefficient for the one time period before the policy is normalized to zero. Each bar represents the 95% confidence interval, calculated with standard errors clustered at the region level. For Callaway and Sant'Anna (2020) estimations, wild bootstrap 95% confidence intervals are presented.

Supplemental Figure 5. Event-study estimates of the effects of a later school start policy on the number of drinking using three different estimators



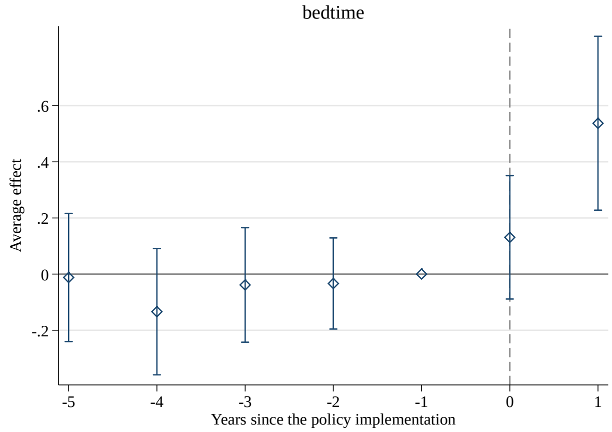
NOTE. Each point represents the point estimate from each event-study estimator, and the coefficient for the one time period before the policy is normalized to zero. Each bar represents the 95% confidence interval, calculated with standard errors clustered at the region level. For Callaway and Sant'Anna (2020) estimations, wild bootstrap 95% confidence intervals are presented.

Supplemental Figure 6. Event-study estimates of the effects of a later school start policy on waketime with school-fixed effects



NOTE. Each point represents the point estimate from each event-study estimator, and the coefficient for the one time period before the policy is normalized to zero. Each bar represents the 95% confidence interval, calculated with standard errors clustered at the region level. School-fixed effects are included in the model in addition to individual- and year-fixed effects. The unit of wake time is hours.

Supplemental Figure 7. Event-study estimates of the effects of a later school start policy on bedtime with school-fixed effects



NOTE. Each point represents the point estimate from each event-study estimator, and the coefficient for the one time period before the policy is normalized to zero. Each bar represents the 95% confidence interval, calculated with standard errors clustered at the region level. School-fixed effects are included in the model in addition to individual- and year-fixed effects. The unit of bedtime is hours.

References

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly harmless econometrics: An empiricist's companion*. Princeton University Press.

Callaway, B., & Sant'Anna, P. H. C. (2021). Difference-in-Differences with multiple time periods. *Journal of Econometrics*, *225*(2), 200–230. https://doi.org/10.1016/j.jeconom.2020.12.001

de Chaisemartin, C., & D'Haultfœuille, X. (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, *110*(9), 2964–2996. https://doi.org/10.1257/aer.20181169

de Chaisemartin, C., & D'Haultfoeuille, X. (2022). *Difference-in-Differences Estimators of Intertemporal Treatment Effects* (Working Paper 29873). National Bureau of Economic Research. https://doi.org/10.3386/w29873

Gardner, J. (2022). *Two-stage differences in differences* (arXiv:2207.05943). arXiv. http://arxiv.org/abs/2207.05943

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, *225*(2), 254–277. https://doi.org/10.1016/j.jeconom.2021.03.014

Goodman-Bacon, A., Goldring, T., & Nichols, A. (2022). BACONDECOMP: Stata module to perform a Bacon decomposition of difference-in-differences estimation. *Statistical Software Components*. https://ideas.repec.org//c/boc/bocode/s458676.html

Rios-Avila, F., Sant'Anna, P. H. C., & Callaway, B. (2023). CSDID: Stata module for the estimation of Difference-in-Difference models with multiple time periods. *Statistical Software Components*. https://ideas.repec.org//c/boc/bocode/s458976.html

Roodman, D. (2023). BOOTTEST: Stata module to provide fast execution of the wild bootstrap with null imposed. *Statistical Software Components*. https://ideas.repec.org//c/boc/bocode/s458121.html

Roodman, D., Nielsen, M. Ø., MacKinnon, J. G., & Webb, M. D. (2019). Fast and wild: Bootstrap inference in Stata using boottest. *The Stata Journal: Promoting Communications on Statistics and Stata*, *19*(1), 4–60. https://doi.org/10.1177/1536867X19830877

Roth, J., Sant'Anna, P. H. C., Bilinski, A., & Poe, J. (2022). *What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature*. 54.

Sun, L. (2022). EVENTSTUDYINTERACT: Stata module to implement the interaction weighted estimator for an event study. *Statistical Software Components*. https://ideas.repec.org//c/boc/bocode/s458978.html

Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, *225*(2), 175–199. https://doi.org/10.1016/j.jeconom.2020.09.006

Wooldridge, J. M. (2021). Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Difference-in-Differences Estimators. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3906345