How Does a Four-Day School Week Affect High School Students? Evidence from Colorado

Meng Song[*] Ruinan Zhao[†]

Working Draft Please Do Not Circulate

Abstract

We examine the impact of the four-day school week on high school students' outcomes using comprehensive administrative data from Colorado, the state with the highest prevalence of this schedule in the United States. The dataset spans nearly two decades of entry cohorts and more than one million unique student records, providing an exceptional opportunity to study the policy's long-run effects. Leveraging quasi-random variation in adoption across districts within a difference-in-differences framework, we estimate how a shortened school week shapes students' educational trajectories. In contrast to much of the existing literature, we find modest improvements in standardized test performance, lower dropout rates, and higher on-time and overall graduation. The duration of exposure plays a key role: students with longer and consistent exposure experience the largest improvements in persistence and attainment. At the same time, heterogeneous effects indicate that disadvantaged groups may benefit less, raising concerns about equity. Overall, the findings challenge the perception that a shorter school week harms student learning and suggest that, when designed and implemented carefully, it can enhance both performance and completion.

JEL Classification: H75 I21 I28

Keywords: four-day school week, student outcomes, teacher outcomes

[*]Department of Economics, University of Connecticut. Email: meng.song@uconn.edu.
[†]Department of Economics, University of California, Riverside. Email: ruinan.zhao@email.ucr.edu.

1. Introduction

Across the United States, the four-day school week (4DSW) has emerged as one of the fastest-growing educational reforms, rapidly reshaping classroom routines, family schedules, and teacher labor markets. Since 1999, the number of districts adopting this schedule has increased six-fold, reaching over 850 districts in 24 states today (National Conference of State Legislatures, 2023). Under this model, students attend school four days per week instead of five, with longer daily hours intended to maintain instructional time. While proponents highlight potential benefits such as cost savings, improved teacher morale, and greater student engagement, critics warn that reducing the school week could harm learning outcomes and exacerbate inequities. The rapid spread of the 4DSW raises urgent questions: how does this schedule affect high school achievement, persistence, and completion, and under what conditions can students adapt without loss of learning?

Existing research has concentrated on elementary and middle school students, with mixed results. Some studies report declines in math achievement when instructional time is reduced, while reading effects are more variable (Thompson, 2021; Morton, 2023). Other work finds modest gains in attendance but possible declines in engagement, along with only limited cost savings, mostly from transportation and utilities (Ward, 2019; Morton, 2021). Fischer and Argyle (2018) also show that the 4DSW can increase juvenile property crime, particularly in urban and low-income settings. Evidence on high school students is scarce and comes mainly from two studies. In Oregon, Thompson et al. (2022) find reductions in test scores, lower on-time graduation, and higher chronic absenteeism, while in Oklahoma, Morton (2023) reports modest declines in disciplinary incidents but no effects on ACT scores or attendance.

Our study examines the impact of a shortened school week on high school students, a group largely overlooked in prior research despite the high stakes involved. Unlike elementary school, which focuses on building foundational skills, high school consolidates academic, cognitive, and socio-emotional abilities that directly shape postsecondary opportunities, labor market outcomes, and long-term life trajectories. High school completion thus strongly predicts further education, stable employment, and higher lifetime earnings, while non-completion imposes enduring economic penalties (Heckman, 2006; Oreopoulos & Salvanes, 2011; Lochner & Moretti, 2004). Educational attainment at this stage also affects broader social outcomes, including crime, health, and civic participation (Lochner & Moretti, 2004; Dee, 2004). Because high school outcomes have lasting consequences for both individuals and society, understanding how reductions in instructional time affect this stage is crucial for evaluating the potential benefits and risks of this reform.

This paper makes four central contributions to the literature on school scheduling, education policy, and human capital formation.

First, we leverage newly available administrative data from Colorado covering the full population of public school students and teachers from 2004 to 2023, comprising over one million unique student records. To our knowledge, this is the first study to exploit these comprehensive data, which combine complete longitudinal coverage with the nation's highest concentration of four-day school week adoption. The dataset's unprecedented scope, granularity, and population-level scale enable a rigorous, large-scale causal evaluation of the policy and allow us to examine its long-term effects on high school outcomes.

Second, we provide new evidence that challenges prevailing conclusions in the literature. Prior studies from Oregon and Oklahoma document declines or no meaningful changes in student achievement and persistence, often linked to reductions in instructional hours. In Colorado, however, 4DSW adoption does not reduce instructional time, and we observe modest but statistically significant gains: English and math scores improve, dropout rates fall by roughly 2 percentage points, and both overall and on-time graduation rates rise. These findings demonstrate that the 4DSW does not necessarily harm student outcomes and underscore the importance of investigating the mechanisms that drive observed effects. Understanding why similar reforms produce contrasting results across contexts is essential for determining the conditions under which a shorter school week can support, rather than hinder, learning and completion.

Third, we uncover substantial heterogeneity in treatment effects. Students with longer and consistent exposure before high school experience the largest improvements, suggesting that sustained engagement allows adolescents to adapt to the schedule and accumulate gains over time. While most students benefit, some disadvantaged groups, such as English learners and homeless students, appear to gain less, indicating that equity may remain an important consideration.

Fourth, we investigate the mechanisms underlying these effects. Analysis of implementation practices, teacher mobility, and district spending reveals two key mechanisms: higher teacher retention and greater classroom investment. Both contribute to a more stable teaching workforce and improved learning environments. Additionally, we find similar performance and persistence outcomes across districts regardless of whether they take Monday or Friday off, but more sustained gains in attainment among Monday-off adopters, suggesting that the organization of instructional time may reinforce the broader institutional mechanisms driving these effects. Together, these findings help reconcile conflicting evidence across states and highlight how local design and investment decisions shape the success of four-day school schedules.

Overall, this paper provides policy-relevant evidence on one of the most rapidly expanding school reforms in the United States. By documenting both the overall gains and the variation in impacts across student groups, the study offers valuable guidance for states and districts considering its adoption.

The remainder of the paper is organized as follows. Section 2 provides background on the 4DSW nationwide and in Colorado. Section 3 outlines the conceptual framework. Section 4 describes the data and empirical strategy. Section 5 presents main results and robustness checks. Section 6 examines mechanisms, and Section 7 concludes.

2. Background of Four-Day School Week

The 4DSW is an alternative scheduling model in which students attend school four days instead of the traditional five. To satisfy state instructional-hour requirements, districts typically extend each school day by 60 to 90 minutes (CDE, 2024). The model was first introduced during the Great Depression as a cost-saving measure and resurged during the 1970s energy crisis as districts sought to limit fuel and transportation costs (Donis-Keller & Silvernail, 2009; Tharp, 2014). Since then, adoption has been concentrated in rural areas, where districts cite not only fiscal savings but also advantages for teacher recruitment and work-life balance (Tharp, 2014; Walker, 2020). Interest rose again during the 2008 financial crisis, when many states faced significant budget shortfalls (Anderson & Walker, 2015; Walker, 2020).

More than 1,600 districts nationwide now operate on a 4DSW, with Colorado, New Mexico, Montana, and Oregon leading adoption (Thompson, 2021; CDE, 2024). Supporters argue that the model improves teacher morale, reduces absenteeism, and provides more time for lesson planning (Anderson & Walker, 2015; Thompson, 2021). Critics raise concerns about declines in achievement, additional childcare burdens for working families, and greater inequities for low-income students who rely on school meals and structured environments (Tharp, 2014; Anderson & Walker, 2015). Empirical evidence remains mixed and context-specific, with results differing by grade level, subject, and state policy setting (Thompson, 2021; Morton, 2022).

Colorado stands out as the most extensive adopter of the 4DSW, both in absolute numbers and in the share of districts participating. The state legislature formally authorized alternative calendars in 1980. Although initial uptake was modest, the Colorado Department of Education (CDE) gradually formalized the approval process, requiring districts operating fewer than 160 school days to seek approval (CDE, 2024). By the early 2000s, adoption accelerated, and as of the 2023–2024 school year, 132 of the state's 187 districts, or more than 70 percent, operate on a four-day schedule, more than tripling adoption in two decades (CDE, 2024). The pattern is mostly rural: more than 80 percent of students in Colorado's rural districts attend schools on a four-day calendar, compared with about 11 percent in non-rural districts. Figure 1 shows the statewide distribution.

The operational structure of Colorado's 4DSW districts differs substantially from that of five-day districts. Traditional schedules typically consist of 160 school days at six instructional hours per day. In contrast, 4DSW districts average 144 days at approximately 7.5 hours per day, usually

Monday through Thursday or Tuesday through Friday (CDE Office of Field Services, 2016; CDE, 2024). This structure ensures compliance with the statewide requirement of 1,080 instructional hours per year.

Flexibility on the fifth day is a defining feature of the policy. Teachers often use the non-instructional day for lesson planning, grading, or professional development (Walker, 2020). Students may engage in enrichment programs, extracurricular activities, community service, or family time. Some districts also provide academic supports such as tutoring, remedial classes, or supervised childcare to mitigate potential learning gaps, although the availability and quality of these services vary considerably across communities (Clear Creek School District RE-1, n.d.).

Colorado's long history with the 4DSW, its widespread adoption, and its binding instructional-hour requirements make it an especially valuable setting for evaluating the policy. The combination of high adoption rates, strict regulatory oversight, and varied local practices creates conditions that differ from those in other states, allowing us to assess the effects of the 4DSW under a distinctive institutional environment.

3. Conceptual Framework

While Section 2 describes the origins and implementation of the 4DSW, we now turn to a conceptual framework that formalizes the channels through which the policy may influence high school outcomes. In particular, we model how the 4DSW affects high school outcomes using a production-function and potential-outcomes framework. Let Y_{ist} denote an outcome for student i in school s and year t. The potential outcome under schedule $D_{st} \in \{0,1\}$ (0 = five days, 1 = 4DSW) is

$$Y_{ist}(D_{st}) = \mathscr{F}\big(T_{st}(D_{st}), \phi_{st}(D_{st}), Q_{st}(D_{st}), C_{ist}(D_{st}); X_{ist}, Z_{st}\big) + \varepsilon_{ist},$$

where T_{st} represents total instructional hours, ϕ_{st} captures instructional productivity (balancing pace and fatigue), Q_{st} denotes teacher quality and productivity, C_{ist} reflects student inputs outside of school (e.g., how the fifth day is used), X_{ist} are student characteristics, Z_{st} captures contextual factors such as rurality and budgets, and ε_{ist} is an idiosyncratic shock.

To clarify notation, let d_{st} be the number of school days per week and h_{st} the length of each day, so instructional time is $T_{st} = d_{st}h_{st}$. Moving to 4DSW reduces d_{st} from 5 to 4, often with longer daily hours. Some districts maintain roughly constant total time ($T_{st}(1) \approx T_{st}(0)$), while others increase or reduce weekly hours. Learning per hour depends on both productivity and day length, $g(\phi_{st}, h_{st})$, with $\partial g/\partial\phi > 0$ and $\partial^2 g/\partial h^2 < 0$, implying diminishing returns as days get longer. Weekly in-school learning is thus

$$\mathscr{L}_{\text{school}}(D) \;=\; d_{st}(D) \cdot h_{st}(D) \cdot g\big(\phi_{st}(D), h_{st}(D)\big).$$

Other channels may shift as well. Teacher quality can improve if shorter weeks reduce burnout or attract staff, or decline if compressed days make teaching harder:

$$Q_{st}(1) \;=\; Q_{st}(0) + \Delta Q_{st}, \qquad \Delta Q_{st} \gtrless 0.$$

Similarly, student out-of-school inputs can rise if the extra day is used for rest or enrichment, but may fall if students face work or caregiving constraints:

$$C_{ist}(1) \;=\; C_{ist}(0) + \Delta C_{ist}, \qquad \Delta C_{ist} \gtrless 0.$$

Because 4DSW adoption is concentrated in rural districts, contextual factors Z_{st} may also differ systematically across adopting and non-adopting schools. Thus, the average treatment effect can be decomposed into contributions from each channel:

$$\tau \;\equiv\; \mathbb{E}\big[Y_{ist}(1) - Y_{ist}(0)\big] \;=\; \frac{\partial \mathscr{F}}{\partial T}\Delta T \;+\; \frac{\partial \mathscr{F}}{\partial \phi}\Delta\phi \;+\; \frac{\partial \mathscr{F}}{\partial Q}\Delta Q \;+\; \frac{\partial \mathscr{F}}{\partial C}\Delta C \;+\; \Delta(\text{interactions}).$$

Conceptually, the sign of τ is indeterminate. Total instructional time may remain constant or change; productivity may improve with focus or deteriorate with fatigue; teacher quality could rise or fall depending on labor-market responses; and student inputs outside of school vary with household circumstances. These offsetting channels imply no clear theoretical prediction regarding the overall effect of the 4DSW. Accordingly, empirical analysis is crucial for establishing whether, in practice, the 4DSW fosters or hinders student achievement.

4. Data and Method

4.1 Data

Our analysis draws on comprehensive administrative records from the CDE, covering the universe of public school students and teachers from the 2003–2004 through 2022–2023 academic years. The student-level data include detailed demographic information (e.g., race, gender), academic attributes (gifted status, language proficiency, migrant status), and key educational outcomes (test participation, standardized test scores, grade repetition, dropout, and graduation). Unique student identifiers enable the construction of longitudinal records from elementary through high school, allowing us to track educational trajectories with precision. The teacher-level data provide demographic information (race, gender), professional characteristics (years of experience, subjects taught, grade levels), and employment details.

Information on 4DSW adoption is also provided by CDE and includes the district name and year of first implementation. Although a handful of districts adopted the 4DSW prior to 2000, our records begin in 2003-2004 academic year, with the bulk of adoptions occurring after the 2008 financial crisis. As of 2024, more than 70% of Colorado public schools operate on a four-day schedule. While districts granted reduced-calendar waivers are not formally required to verify implementation, discussions with CDE officials confirm that nearly all waiver recipients transition to a four-day schedule, substantially alleviating concerns about misclassification.

We supplement these administrative records with school- and district-level characteristics, including student composition, enrollment, and rural designation. These measures serve primarily as controls to account for compositional differences across schools.

Although the dataset includes all Colorado public school students in grades 1 through 12, our baseline sample focuses on first-time ninth-grade entrants from the 2004 to 2005 through the 2019 to 2020 cohorts. The 2004 to 2005 cohort is the first with standardized test scores, and restricting follow-up through 2019 to 2020 ensures that each cohort can be observed through the end of high school. This design allows us to measure complete long-run outcomes, including dropout, on-time graduation, and overall graduation. To avoid treatment that turns on and off, we exclude students in districts that adopted 4DSW and later returned to a five day week. These cases account for fewer than one percent of the full sample.[1] We also examine heterogeneity in exposure timing by distinguishing two subgroups: (1) students first exposed to the 4DSW in high school, and (2) students first exposed in middle school.[2]

4.2 Outcomes

We examine outcomes in several domains: performance, progression, and attainment. Performance is measured by standardized English and math test scores, drawn from the ACT (administered statewide prior to 2016–2017) and the SAT (administered thereafter). To ensure comparability across years and formats, all scores are standardized within administration year across all districts to have mean zero and unit standard deviation. Progression is captured by grade repetition, defined as whether a student repeated a grade in high school, and dropout, defined by CDE records indicating whether a student exited school before earning a diploma. Attainment is measured by on-time graduation, defined as completing high school within four years of entering ninth grade, and overall graduation, which includes students completing in more

[1] Ninth grade marks the transition into high school, a stage when families typically have limited information about school-specific rigor, teacher quality, or extracurricular opportunities. As a result, initial enrollment choices are less likely to reflect detailed knowledge of school effectiveness or schedule design, reducing concerns about self-selection into schools. Consistent student identifiers allow us to link records across elementary, middle, and high school, making it possible to capture cumulative exposure to the four day school week. Additionally, four-day schools are disproportionately located in rural areas, and the overall share of students enrolled in such schools remains relatively small. To improve covariate balance between treated and control schools, our robustness checks employ propensity propensity score matching and a random forest classifier to exclude control schools that are observationally dissimilar from treated schools.

[2] We exclude students first exposed in elementary school because evaluating their high school outcomes requires a substantially longer observation window, yielding a smaller and more selective sample.

than four years.

Teacher outcomes include:[3] (1) retention, defined as whether a teacher remains in the same school from one year to the next; (2) transfer, defined as leaving the current school but taking a position in another Colorado public school; and (3) turnover, defined as leaving the public school system entirely, with no subsequent employment in any Colorado public school. These outcomes are mutually exclusive. We also observe rich teacher characteristics, including gender, race, grade level, subjects taught, degree attainment, and years of experience.

Following prior research on Oregon's four-day week (Thompson, 2021), we examine if total instructional hours affect student outcomes. In Oregon, declines in English and math performance were largely attributed to reduced instructional time. Colorado, however, imposes strict minimum instructional-hour requirements on districts adopting a four-day schedule, which could account for the more favorable results we observe. To test this directly, we analyze whether the policy changes instructional hours in Colorado, providing evidence to reconcile our findings with prior work. We also examine school district expenditures, since a shorter week may alter how resources are allocated across instructional and non-instructional uses. Together, these outcomes allow us to assess whether Colorado's experience reflects distinct implementation choices and resource patterns, helping to explain why its results diverge from those reported elsewhere.

4.3 Summary Statistics

Table I presents means for four-day school districts before and after adoption of the 4DSW, as well as for traditional five-day districts.

Panel A describes student composition and academic outcomes. The demographic profile of students in four-day schools is broadly comparable to that in five-day districts, though modest differences remain. The share of male students is nearly identical across groups at about 51 percent, while racial and ethnic composition diverges slightly. Four-day schools serve fewer White

[3] We use teacher outcomes primarily for mechanism analysis. Details are in Section 6.

and Asian students but a larger share of Hispanic students, with smaller Black representation relative to five-day districts. Other indicators such as the shares of English learners, migrant students, gifted students, and homeless students are closely aligned.

Academic outcomes indicate that students in four-day districts begin from a somewhat lower baseline. Average English and mathematics test scores are roughly 1.5 points lower than those in five-day districts. Graduation outcomes also lag modestly, with lower on-time and overall graduation rates and slightly higher dropout rates. These descriptive differences are consistent with the fact that many four-day districts tend to be smaller and more rural.

Panel B presents teacher characteristics and labor market outcomes. Teacher demographics are overall similar across the two groups, with only minor differences in the shares of male, White, Black, Hispanic, and Asian teachers. Differences emerge more clearly in educational attainment: teachers in four-day schools are more likely to hold a bachelor's degree and less likely to hold a master's degree than those in five-day districts. Teachers in four-day districts also have slightly more experience on average, around 9.7 years compared with 9.1 years in five-day districts. Labor market outcomes are likewise comparable. Retention rates are somewhat higher in four-day schools at about 76 percent, while exit rates are lower at around 18 percent and transfer rates differ little between the two groups.

4.4 Econometric Method

We begin by examining the impacts of 4DSW on student outcomes using variation in when and where schools adopted 4DSW in a generalized difference-in-differences framework. Specifically, we estimate the following equation:

$$Y_{icsd} = \beta_1 G1_{cd} + \beta_2 G2_{cd} + X^{'}_{icsd} + \theta_s + \lambda_c + \epsilon_{icsd} \tag{1}$$

where Y_{icsd} is the outcome of interest, such as an indicator equal to one if student i in school s, district d, and high school entry cohort (9th grade) c drops out during high school. The indicator $G2_{cd}$ takes the value one for cohorts that entered ninth grade in the year of, or after, the adoption of the 4DSW in district d, ensuring full exposure throughout high school. Our main outcomes, including English and math test scores, dropout, and graduation, are typically realized several years after students first enter high school (most often around 12th grade).[4] The coefficient β_2 therefore captures the average effect for cohorts that were fully exposed to the 4DSW from the start of high school. By contrast, earlier cohorts may have been only partially exposed beginning in 10th, 11th, or 12th grade. To account for this, we include a second indicator, $G1_{cd}$, which equals one for cohorts that entered high school one to three years prior to the 4DSW adoption in their districts. Thus, $G1_{cd}$ identifies students who were not exposed in ninth grade but encountered the 4DSW later in high school, and β_1 measures the effect of this partial exposure.

The model also includes a vector of controls, X'_{icsd}, to account for observed student, teacher, school, and district characteristics. θ_s denotes school fixed effects to absorb any school-level time-invariant confounding factors, while λ_c denotes entry-cohort fixed effects. ϵ_{icsd} is the error term. Robust standard errors are clustered at the district level.

Furthermore, for mechanism analysis purpose, we explore how 4DSW affects teacher outcomes, such as retention, transfer, and turnover. To do so, we estimate a similar equation to equation (1) but exclude the partial treatment indicator, as teacher outcomes can be observed at the school-year level. The model is specified as follows:

$$Y_{jsdt} = \beta_1 G_{dt} + X'_{jsdt} + \theta_s + \lambda_t + \epsilon_{jsdt} \tag{2}$$

where Y_{jsdt} denotes the outcome for teacher j in school s and district d in school year t; X'_{jsdt} is a vector of controls; θ_s and λ_t are school and school-year fixed effects, respectively. Robust

[4] A potential concern is that some outcomes are observed only for students who remain enrolled through the relevant grade level, which could introduce sample selection if the four day school week affects persistence. In Colorado, however, the minimum legal dropout age is seventeen, so nearly all students remain enrolled through the end of eleventh grade. Consequently, standardized test taking and course repetition, which occur before twelfth grade, are measured prior to the earliest dropout age and observed for the full student population. Only graduation, realized at the end of twelfth grade, is mechanically conditional on persistence. To address this, we estimate effects on both dropout and graduation, capturing the extensive and intensive margins of high school completion and mitigating concerns about endogenous sample selection.

standard errors are clustered at the district level.

The baseline DiD approach summarizes impacts over the outcome time horizon. To examine how effects vary by the timing of exposure, we further estimate the following event-study equation:

$$Y_{icsd} = \sum_{k=-7,k\neq -4}^{k=3} \beta_k G_{cd}^k + \beta_2 X_{icsd}^{'} + \theta_s + \lambda_c + \epsilon_{icsd} \tag{3}$$

where the indicator variable G_{cd} is replaced by an event-year indicator G^k. All other variables are defined as same as equation (1). The event windows are seven years prior to and three years after 4DSW.[5] The student cohorts where $k = \{-1, -2, -3\}$ years from the fully treated cohorts ($k = 0$) are considered partially treated cohorts. The omitted category is the last fully untreated cohort, $k = -4$.[6]. Coefficients corresponding to the years after 4DSW (β_k for $k \geq 0$) captures the post-adoption treatment effects that occur in response to 4DSW over time.

4.5 Threats to Identification

A major threat to our DiD framework is the violation of the parallel trends assumption. While the event study model in equation (3) estimates the post-treatment effect of the 4DSW, it can also be used to test for parallel trends. If lagged effects exist, our identification strategy is likely invalidated. Figure 2 depicts the dynamics of the 4DSW effect on various outcomes. In all cases, the pre-adoption coefficients are small in magnitude and statistically insignificant, providing little evidence of pre-trends.[7]

[5]The event time window extends three cohorts beyond the adoption year because majority of the treated districts can be consistently observed for at least three subsequent ninth grade entry cohorts. In addition, the plus one to plus three cohorts likely include students who were first exposed to the 4DSW during middle school, assuming they remained in the same district, thereby capturing early treated students with prior exposure to the policy before entering high school.

[6]As for teacher outcomes, there is no partial treatment issue, so the omitted category is $k = -1$. Specifically, we estimate the event-study equation for teacher as:

$$Y_{jsdt} = \sum_{k=-7,k\neq -1}^{k=3} \beta_k G_{dt}^k + \beta_2 X_{jsdt}^{'} + \theta_s + \lambda_t + \epsilon_{jsdt} \tag{4}$$

All variables are defined as same as in equation (2).

[7]The coefficients for periods prior to 4DSW adoption are jointly insignificant at conventional significance levels.

Although the lack of large and significant pre-treatment effects supports the causal interpretation under the DiD framework, the conventional two-way fixed effects (TWFE) model is not well suited for testing pre-trends, particularly when treatment is staggered and effects are heterogeneous (Sun & Abraham, 2021; Goodman-Bacon, 2021; Borusyak et al., 2022). In such settings, unless treatment effects are homogeneous, β_k is a linear combination of group-specific effects from both its own and other relative periods. As a result, treatment effects from other periods may contaminate the estimate of β_k. To address this concern, we re-estimate the event study model using the "interaction-weighted" estimator proposed by Sun and Abraham (2021). Figure 3 presents the coefficient plots. The alternative method also shows little evidence of pre-trends, strengthening the credibility of our DiD design.

The difference in differences and event study models require that adoption of the 4DSW be unrelated to any preexisting trend in long run outcomes across school districts. We present two empirical checks of this assumption. We begin by assessing whether adoption was preceded by systematic movements in district characteristics. To investigate this, we estimate a modified event study that replaces the pre 4DSW indicators with a single linear trend:

$$Y_{dt} = \alpha_1 \,\textbf{year\_before\_4DSW} + \sum_{k=1}^{3} \gamma_k (\textbf{year\_after\_4DSW} = k) + \beta_2 X_{dt} + \theta_d + \lambda_t + \varepsilon_{dt}, \qquad (5)$$

where Y_{dt} denotes district-level characteristics, **year_before_4DSW** is a linear time trend in the years prior to a district's adoption, and θ_d and λ_t are district and year fixed effects. The coefficient α_1 captures the slope of outcomes prior to 4DSW adoption. Table A.1 reports the estimates. Across the characteristics considered, only one coefficient is statistically significant. Second, we examine the association between the year of 4DSW adoption and baseline district characteristics.[8] Table A.1, Column (2), provides no evidence of such a relationship, indicating that adoption was not systematically driven by initial district conditions.

We then assess whether the timing of 4DSW adoption is systematically related to pre-policy values of the outcome variables among adopting districts. Specifically, we regress the year of

[8] The baseline year is the initial data year 2005.

adoption on the two-, three-, and four-year averages of each outcome variable prior to implementation. As indicated by Table A.2, across all specifications, the coefficients are small and statistically indistinguishable from zero. These findings indicate that the sequence of adoption is not driven by prior differences in student or teacher outcomes, supporting the identifying assumption that variation in adoption timing is plausibly exogenous to underlying outcome dynamics.

Next, we examine whether adoption of the 4DSW induces student mobility that could bias the estimated effects. Two forms of movement are relevant. The first is within-state transfers, where families relocate between districts to seek or avoid a four-day schedule.[9] Such sorting could alter district composition and confound causal interpretation. The second is attrition from the public system, in which students leave Colorado's administrative records entirely through private schooling, homeschooling, or out-of-state migration. Table A.3 shows that adoption of the four-day week does not significantly affect district demographics or total enrollment, while Table A.4 directly tests both forms of mobility using the full student panel.[10] The estimates reveal no change in the likelihood of transferring to another district or exiting the state public system prior to graduation. Overall, these results indicate that four-day adoption does not induce selective migration or attrition, supporting the credibility of the identification strategy.

[9]For example, a student may transfer from a five-day district to a four-day district, or vice versa, due to preferences related to scheduling, commuting, or child-care logistics.

[10]Specifically, Table A.4 examines (i) whether students in any grade are more likely to switch school districts after adoption of the 4DSW, and (ii) whether students previously observed in the state's public data later fail to reappear — either before entering high school (entry margin) or before graduation (exit margin) — without being classified as dropouts.

5. Results

5.1 Student-Level Outcomes

Table II reports difference-in-differences estimates of the effect of adopting the 4DSW on high school outcomes. All models include school and entry-cohort fixed effects along with a set of controls.

For achievement, we find no evidence that the 4DSW alters standardized test participation, with estimates very close to zero. Test performance, however, shows modest but meaningful improvements under full exposure. English scores increase by 0.027 standard deviations, significant at the 5 percent level, and mathematics scores by 0.025, significant at the 10 percent level. While modest in absolute terms, these gains translate into roughly one to one-and-a-half months of additional learning, using the common benchmark that 0.25 standard deviation corresponds to one grade-level year of progress (Bloom et al., 200). Moreover, the magnitudes are comparable to other widely studied school-based reforms. For example, accountability reforms under No Child Left Behind produced improvements of a similar order (Hanushek & Raymond, 2005; Dee & Jacob, 2011). By contrast, the estimates for students first exposed later in high school are smaller and imprecisely estimated.

Turning to progression outcomes, grade repetition is unaffected. Dropout, however, declines by a statistically significant 2.1 percentage points under full exposure, a 18 percent reduction relative to the control mean of 11.8 percent. This is a notable result given that Colorado law permits students to leave school at age 17, underscoring the policy's role in keeping marginal students enrolled at the critical stage when early exit becomes possible. Moreover, partial exposure produces smaller, statistically insignificant estimates, reinforcing the importance of sustained implementation.

The largest impacts appear in attainment. Under full exposure, graduation increases by 3.2 percentage points and on-time graduation by 2.9 percentage points, both statistically significant at the one percent level. These correspond to improvements of approximately 4.3 and 4.0 percent relative to baseline means of 0.747 and 0.732, respectively. Students first exposed after ninth grade also experience gains, with graduation and on-time graduation rising by 1.3 and 1.4 percentage points, both significant at conventional levels. Although smaller in magnitude, these effects indicate that the 4DSW improves completion even for cohorts with partial exposure, with the largest and most consistent benefits accruing to those continuously exposed from ninth grade onward.

To put these estimates in perspective, we benchmark them against findings from related studies. For example, Jackson et al. (2020) show that attending a school one standard deviation higher in predicted test score value-added increases high school graduation by about 1.3 percent for ninth-grade students in Chicago public schools. By comparison, the 4.3 percent increase in graduation that we estimate under full exposure to the 4DSW is roughly equivalent to attending a school with 3.3 standard deviations higher test score value-added. Furthermore, scaling the estimated effects of the 4DSW on high school students underscores their magnitude. With roughly 68,000 students in each annual cohort in Colorado (CDE, 2025), full exposure to the policy would prevent about 1,428 dropouts, yield 2,176 additional graduates, and produce 1,972 more on-time graduates. Even modest percentage gains at the student level therefore translate into substantial aggregate improvements in human capital, consistent with evidence on the high returns to high school completion (Lochner & Moretti, 2004; Oreopoulos & Salvanes, 2011) and the considerable long-term costs of dropping out (Heckman, 2006).

The positive effects observed in Colorado stand in contrast to prior evidence from Oregon, where Thompson et al. (2021) report declines of roughly 0.09 standard deviations in test scores and 4 to 5 percentage points in graduation rates, with no change in dropout. In Colorado, by comparison, both achievement and attainment rise. This difference underscores that the consequences of a 4DSW depend crucially on implementation rather than on the schedule itself. In Oregon, instructional hours fell substantially, reducing learning time. In Colorado, districts largely maintained required hours and often redirected savings toward classroom resources and teacher support, effectively ruling out lost instructional time as the primary channel (see Section 6). These institutional contrasts highlight that school schedules interact with broader resource and personnel decisions. As Jackson et al. (2020) emphasize, schools influence students through multiple pathways, including academic, behavioral, and organizational dimensions, and the relative importance of these channels varies across contexts. Understanding these mechanisms is therefore central to explaining why the same policy can yield divergent outcomes across settings and to identifying the design features that make alternative schedules most effective.

5.2 Duration of Exposure

The results in Table II show that students fully exposed to the 4DSW throughout high school experienced larger gains in persistence and attainment than those only partially exposed, suggesting that the duration of exposure may be an important determinant of impact. This raises a natural question: do the effects of the 4DSW depend on when students are first introduced to the policy and how long they remain exposed? To address this, we estimate the effects of 4DSW for cohorts first exposed in middle school and for those whose exposure began in high school.[11] Table III summarizes the results.

The evidence reveals consistent but heterogeneous benefits. For students first exposed only in high school, the estimates show moderate gains in both performance and attainment. Test participation remains unchanged, while English and math scores increase by about 0.02 to 0.03 standard deviations. Dropout falls slightly but not significantly, and both graduation and on-time graduation rates rise by roughly 1.6 to 1.7 percentage points. These findings indicate that even limited exposure to the four-day schedule can improve academic outcomes during high school, though the magnitude of the effects remains modest.

Students first exposed in middle school, by contrast, experience larger and more precisely estimated effects. English and math scores rise by roughly 0.02 standard deviations, while dropout declines by about 1.9 percentage points. Graduation and on-time graduation rates increase by more than 4 percentage points, both statistically significant. These results suggest that earlier

[11] Students first exposed in elementary school are excluded because requiring continuous records from elementary through high school substantially reduces the sample size and statistical power.

and sustained exposure magnifies the policy's benefits, producing more durable improvements in persistence and completion.

Panel C provides further evidence of a dose–response relationship. Each additional year of exposure to the 4DSW is associated with 0.6 to 0.7 percentage point increase in graduation and on-time graduation and a 0.4 percentage point decline in dropout, all statistically significant. The cumulative pattern indicates that the gains from the 4DSW accrue gradually: the longer students remain under the policy, the greater the improvements in high-school outcomes.

Taken together, these results highlight that both the timing and duration of exposure are central to understanding the effects of the 4DSW. Early adoption, beginning before high school, appears most conducive to sustained improvements in attainment because it coincides with a critical stage in students' academic and behavioral development. Adolescence is a formative period when students begin to internalize learning routines, self-discipline, and time management skills that strongly influence later outcomes (Cunha & Heckman, 2007). Longer and more consistent exposure allows these behavioral and organizational adjustments to take hold and reinforce one another over time. It also provides teachers and schools with greater opportunity to adapt instructional practices and schedules. Districts that transition earlier often pair the reform with complementary investments such as professional development, curriculum redesign, and improved planning processes that enhance the overall effectiveness of the four-day schedule (Jackson, 2018). In contrast, students encountering the policy for the first time in high school may have already established routines that are less responsive to changes in the school structure. Policymakers should therefore expect smaller short-term effects among partially exposed cohorts and consider additional supports such as tutoring, transition assistance, or family coordination to facilitate adjustment. The evidence suggests that sustained and consistently supported implementation, rather than temporary experimentation, is essential for realizing the longer-term educational benefits of the 4DSW.

5.3 Do All Students Benefit Equally

Table IV examines whether the effects of the 4DSW vary across students. The analysis considers whether improvements in achievement, progression, and attainment are broadly shared or concentrated within particular populations.

By gender, the estimated effects are broadly similar across most outcomes. Both male and female students show significant improvements in progression and attainment measures, while gains in test performance are somewhat larger and more precisely estimated for males. Overall, the results indicate that the benefits of the 4DSW are comparable across genders, with only differences in test performance.

Across racial and ethnic groups, graduation outcomes improve overall, though the strength and precision of the estimates vary. White and Hispanic students exhibit consistent gains in both achievement and attainment, while effects for Black students are positive but less precisely estimated due to smaller sample sizes. Taken together, the findings suggest that the 4DSW's positive impacts are not confined to a particular demographic group but are present across major racial and ethnic populations, with some variation in magnitude.

Differences are more pronounced when examining subgroups defined by student background. English learners and homeless students, who represent smaller shares of the population, display weaker or even negative estimates in some outcomes, while their non-English learner and non-homeless peers experience more consistent gains. Because these groups are relatively small, the results should be interpreted with caution. Nonetheless, the pattern suggests that students facing greater socioeconomic barriers or relying more heavily on school-based services may benefit less from a shortened week. Reduced instructional time and fewer in-school resources could possibly present greater challenges for these populations, even if the overall school environment becomes more flexible for most students.

5.4 Robustness Checks

To reinforce the credibility of our baseline estimates, we conduct an extensive set of robustness and complementary analyses.

We begin by addressing potential imbalance between treated and control units. Because adoption of the 4DSW is concentrated in smaller, rural districts, the number of treated students is relatively limited compared to the statewide pool of non-adopting students, even though adoption spans a substantial share of districts. To enhance comparability, we first implement propensity-score matching based on pre-treatment school characteristics predicting adoption. Each treated school is matched to its nearest control schools in terms of the estimated propensity score, without replacement, and the least comparable control schools are excluded.[12] The resulting estimates (Table A.5, Panel A) are nearly identical to the baseline, indicating that differences in sample composition are not driving the main results.

To further examine balance, we re-estimate adoption propensities using a random-forest classifier, which flexibly captures nonlinearities and higher-order interactions among covariates without imposing a parametric form. Applying the same matching procedure based on these nonparametric predictions,[13] the effects (Table A.5, Panel B) remain close to the baseline, suggesting that the findings are not sensitive to functional-form assumptions in the propensity-score model. As an alternative reweighting strategy, we also estimate inverse probability weighted (IPW) models that retain the full sample but reweight observations so that the control group mirrors the covariate distribution of treated schools. The results (Table A.5, Panel C) again align closely with the baseline, reinforcing the conclusion that the estimated impacts are robust to different approaches for achieving covariate balance.

Next, we assess the sensitivity of the results to model specification. Estimating the baseline without any control variables yields coefficients that are very similar to the main specification, showing that the inclusion of covariates does not materially affect the estimates. Adding district-specific linear pre-trends to account for potential differences in underlying trajectories across districts (Table A.5, Panel E) produces nearly identical coefficients. Moreover, to test

[12]Specifically, the analysis is restricted to schools whose estimated propensity scores fall between the 5th and 95th percentiles of the combined treated–control distribution, using 1:3 nearest-neighbor matching without replacement.

[13]The same 1:3 nearest-neighbor matching procedure is used based on the random-forest estimates.

whether any single adopter drives the results, we sequentially exclude each adopting district and re-estimate the model. As illustrated in Figure A.3, the estimated effects are stable, suggesting that no individual district drives the results.

We then examine the treatment definition. The baseline specification includes an indicator for students who begin high school on a five-day schedule but whose district adopts the 4DSW while they are still enrolled, thereby accounting for partial exposure. As a robustness check, we restrict the sample to cohorts fully exposed since ninth grade and those never exposed, excluding all partially treated cohorts. This cleaner comparison avoids functional-form assumptions about partial exposure. The results, presented in Table A.6, remain stable, confirming that the findings do not hinge on how partial exposure is modeled.

To ensure results are not driven by specific subsamples, we perform several additional checks. Limiting attention to students entering high school in 2015 or earlier, thus graduating before the Covid pandemic, produces estimates that closely mirror the baseline. Excluding charter schools and Boards of Cooperative Educational Services (BOCES), which operate under distinct governance structures, yields similar results. As Colorado's 4DSW adopters include several urban districts, we also re-estimate the models restricting attention to small rural systems.[14] The estimates remain consistent with the baseline, suggesting that the main conclusions are not an artifact of urban–rural composition.

We further assess the robustness of the achievement outcomes to test-format changes. Colorado administered the ACT before 2016–2017 and the SAT thereafter. Although our baseline standardizes scores across assessments, we re-estimate the models excluding all SAT-based cohorts (and, separately, all ACT-based cohorts). In both cases, the results remain virtually unchanged,[15] indicating that format transitions are not driving the findings.

Finally, we conduct an inference test by randomly reassigning adoption years across districts 1,000 times. Figure 4 shows that the actual estimates fall in the extreme tails of the placebo

[14]Rural districts are defined based on size and distance from the largest urbanized area, with an enrollment of 6,500 or fewer students. Small rural districts meet the same criteria but have fewer than 1,000 students (CDE, 2025).

[15]Excluding all ACT-based cohorts yields similar results.

distribution, with fewer than 10 percent of simulated draws exceeding the observed effects in absolute value. This pattern provides additional evidence that the results are unlikely to arise by chance.

Taken together, these exercises demonstrate that the estimated impacts of the 4DSW are stable across a wide range of specifications, samples, and inference procedures.

6. Potential Mechanisms and Discussion

6.1 Potential Mechanism

Guided by the conceptual framework in Section 3, we next explore the primary mechanisms through which 4DSW may influence student outcomes. We focus on five potential channels: (i) changes in instructional hours, (ii) variation in how districts implement the 4DSW, (iii) adjustments in the teacher labor market, (iv) shifts in instructional and supplies expenditures, and (v) opportunities for student enrichment on the additional day off.

Among the most direct pathways, instructional time has received the greatest attention in prior research. Evidence from Oregon suggests that reductions in instructional hours are the dominant channel: fewer hours account for much of the negative academic impact of the 4DSW there (Thompson, 2020), consistent with broader work emphasizing the centrality of instructional time (Rivkin & Schiman, 2015; Lavy, 2015). Colorado, however, operates under a distinct institutional environment. The CDE enforces a binding minimum of 1,080 instructional hours per year across all districts (CDE, 2025). Moreover, adoption of the 4DSW in Colorado is associated with improved, rather than worsened outcomes, raising the question of which mechanisms may be driving these effects.

To assess the role of instructional time, we first document average annual hours in adopting districts (Figure A.4). All treated districts meet or exceed the 1,080-hour threshold, and many provide more. When we re-estimate our baseline specification using instructional hours as the dependent variable, as shown in Table A.9, we find that four-day districts provide, on average, 23 additional hours annually, equivalent to roughly 0.56 more hours per week.[16] Although this difference is not statistically significant, it confirms that, unlike in Oregon, fewer school days do not translate into fewer instructional hours in Colorado.

We next examine whether the choice of the non-instructional day influences the impacts of the 4DSW. In Colorado, districts have discretion to determine which day of the week becomes non-instructional, with nearly all opting for either Monday or Friday.[17] These alternatives plausibly carry different implications for student behavior and learning. A Monday-off schedule may allow students to recover from weekend activities and begin the week more rested, potentially reducing absenteeism and fatigue (Carrell et al., 2011; Lo et al., 2016). In contrast, a Friday-off schedule extends the weekend, offering greater opportunities for work, travel, and family obligations that could affect engagement (King, 2002). More broadly, research shows that the organization of instructional time is consequential for learning (Patall et al., 2010), and cognitive studies suggest that maintaining midweek continuity, as under a Monday-off schedule, may better support retention and routine (Pashler et al., 2007).

When we disaggregate the effects by the designated off-day, test performance and dropout outcomes are remarkably consistent across the two schedules: both Monday- and Friday-off districts exhibit nearly identical effects on performance and persistence, with no meaningful differences in test scores or dropout rates. Friday-off districts show a modest increase in grade repetition, whereas Monday-off districts display a slight decline, although the difference between the two estimates is not statistically significant. The more substantive divergence arises in longer-run attainment. Districts that take Monday off experience significantly larger gains in graduation and on-time graduation — about 2 to 3 percentage points higher than comparable Friday-off districts — while Friday-off adopters show smaller and statistically insignificant improvements. This pattern indicates that although both variants of the 4DSW yield similar short-term academic outcomes, the Monday-off model generates more sustained benefits for

[16] Assuming 160 instructional days under a four-day schedule.

[17] Information on the designated off-day is provided by the CDE and supplemented by manual searches of school and district websites.

persistence and completion, highlighting the importance of how instructional time is structured across the week.

The results by instructional hours and day-off scheduling suggest that these dimensions cannot fully account for the observed student improvements under the 4DSW. A natural next step is to consider the role of teachers, who are central to the educational production process, as outlined in Section 3. If the 4DSW alters teachers' career decisions or working conditions, these changes may in turn shape student outcomes. In particular, prior research links teacher stability to student achievement and persistence, making the teacher labor market a plausible channel through which the 4DSW could operate.

Table V shows that adoption of the 4DSW increases high school teacher retention by 2.7 percentage points, a result that is statistically significant. With roughly 55,000 public high school teachers employed statewide (CDE, 2025),[18] this effect corresponds to nearly 1,500 additional teachers remaining in their schools each year. By contrast, the estimates for teacher transfers and exits from the profession are small and statistically indistinguishable from zero. This pattern suggests that the main impact of the 4DSW is to enhance stability within schools rather than alter broader labor-market mobility. One plausible explanation is that the 4DSW improves job satisfaction by offering teachers a more flexible schedule without reducing compensation. As shown in Appendix Table A.11, the 4DSW does not significantly affect teacher salaries or benefits, meaning that educators receive the same pay and benefits while teaching one fewer day per week. This effectively increases their non-pecuniary compensation by lowering commuting costs and providing an additional day for lesson planning, professional development, or personal responsibilities. Such improvements in work–life balance likely reduce stress and burnout, consistent with evidence that schedule flexibility is a key determinant of teacher retention (Anderson, 2015; Morton, 2020).

As teacher retention emerges as the most consistently significant teacher outcome, we next assess whether the magnitude of this change could reasonably account for part of the student

[18] Results for the full sample of teachers across all grades are similar; see Appendix Table A.8.

gains associated with the 4DSW. Prior research shows that teacher turnover reduces student achievement and persistence, with school level losses on the order of several hundredths of a standard deviation for moderate increases in turnover (Ronfeldt et al., 2013; Sorensen et al., 2021). Using these estimates to benchmark the size of our retention effect suggests that improved workforce stability could explain roughly 10 to 20 percent of the overall treatment effects on graduation and dropout.[19] Although this is only a rough estimate, it indicates that enhanced teacher stability is an important channel through which the 4DSW strengthens student persistence.

Beyond teacher outcomes, we also consider whether districts shifted resources within the classroom. We focus on instructional and supplies expenditures, which provide a direct measure of the resources available for day-to-day teaching and learning.[20] A large body of evidence shows that increases in instructional spending can improve student achievement and long-run outcomes (Jackson et al., 2016). More recent work also highlights the importance of classroom materials, finding that access to adequate supplies is associated with stronger academic performance (Harris, 2023). Parallel evidence from higher education demonstrates that lowering the cost of course materials improves completion rates and grades (Publishers Association, 2025; Open Praxis, 2023). Results reported in Appendix A.10 indicate that both instructional and supplies expenditures rose under the 4DSW. While the increase in instructional spending is small and imprecisely estimated, supplies expenditures grew by roughly 18 percent, equivalent to about $74 more per pupil each year.[21] This represents a meaningful increase in classroom re-

[19]A back of the envelope calculation based on published elasticities helps contextualize this magnitude. Ronfeldt et al.(2013) and Sorensen et al. (2021) estimate that a 10 percentage point increase in teacher turnover lowers average student achievement by about 0.02 to 0.05 standard deviations at the grade level. Inverting these estimates implies that a comparable improvement in retention would raise achievement by roughly 0.02 to 0.05 standard deviations. Given that the 4DSW increases teacher retention by 2.7 percentage points, a linear approximation suggests expected gains of about 0.005 to 0.014 standard deviations in student achievement. Relative to the observed achievement effects of approximately 0.025 standard deviations in our data, this range corresponds to roughly 10 to 20 percent of the total effect, consistent with the interpretation that improved teacher stability contributes meaningfully to the overall student gains.

[20]Instructional expenditures include classroom-related costs tied to instruction. Supplies expenditures capture items such as textbooks, classroom consumables, technology, and materials used in instruction or enrichment activities. All expenditure measures are constructed at the district-year level and expressed on a per-pupil basis.

[21]The average district supplies expenditure per pupil is $412 per year.

sources that could plausibly strengthen the learning environment.[22]

We next turn to how students may have used their additional day off under the 4DSW. Although the schedule reduces the number of formal instructional days, some districts program the fifth day with enrichment, tutoring, athletics, and extracurricular activities. Research on out-of-school time consistently finds that structured activities can foster persistence and socio-emotional development by providing safe environments, mentoring, and opportunities for skill building (Lauer et al., 2006; Fredricks & Eccles, 2006). Districts may also use the day for targeted academic support, such as credit recovery or individualized tutoring, which has been shown to improve high school completion and test performance (Dynarski & Gleason, 2002; Nickow et al., 2020). To approximate these opportunities, we examine district spending on purchased services and community services.[23] These budget categories provide an indirect measure of whether resources are reallocated toward enrichment on the fifth day. As reported in Appendix A.11, however, we do not find systematic increases in either category under the 4DSW. While these measures are imperfect proxies and cannot capture participation or program quality, the available evidence does not suggest a major shift in spending toward extracurricular activities.

Taken together, the evidence indicates that Colorado's positive experience with the 4DSW is not explained by reductions in instructional hours or by whether districts adopt a Monday or Friday schedule. Instead, the results point to a combination of factors: improvements in teacher retention and modest increases in classroom resources such as supplies. These mechanisms appear to contribute to student persistence and attainment, though to varying degrees and with important caveats. Other plausible channels — such as changes in student fatigue, mental health, or parental labor supply — may also play a role. Overall, the findings suggest that the benefits of the 4DSW arise from a mix of school-level adjustments and student-level responses, while highlighting the need for further research to fully understand how the shorter

[22]At the same time, existing studies provide little direct evidence on how increases of this magnitude translate into achievement gains, so the results here should be viewed as suggestive rather than conclusive.

[23]Purchased services typically cover payments to external providers of educational or extracurricular programs; community services encompass activities directed toward student and family engagement outside of regular instruction. Expenditures are measured at the district-year level and expressed on a per-pupil basis.

week reshapes the educational production process.

6.2 Discussion

As a back-of-the-envelope exercise, we translate the estimated impacts on high school graduation into monetary terms. Following Heller et al.(2017) and Cohodes et al.(2023), we treat each additional graduate as equivalent to one more year of schooling relative to a dropout. We monetize benefits through two channels: (i) higher lifetime earnings and (ii) health improvements associated with greater educational attainment.

For earnings, we draw on synthetic work-life profiles in Julian & Kominski (2011) and apply the conventional 12 percent return to an additional year of schooling. Prior work indicates that high school completion yields a present value on the order of $200,000 per graduate in lifetime earnings (Lochner & Moretti, 2004; Oreopoulos, 2007; Rouse, 2007). For health, we follow Cutler & Lleras-Muney (2006) and assign $28,750 per additional graduate, the midpoint of their reported range. On the cost side, consistent with Cohodes et al. (2023), who proxy the resource cost of an additional year of high school using per-pupil spending, we use average operating expenditures in Colorado of about $11,000 per student (CDE, 2025).

Applying this framework, the 4DSW increases the number of graduates by roughly 2,176 per entering cohort of 68,000 students. This implies aggregate benefits of about $435 million in lifetime earnings and $63 million in health, offset by roughly $24 million in additional schooling costs. The resulting net graduation benefit is approximately $474 million per cohort, which equates to about $7,000 per entering student or $218,000 per additional graduate.

While illustrative, these magnitudes indicate that modest increases in graduation can generate sizable social returns. The calculation abstracts from other channels through which the 4DSW may affect welfare, including reduced criminal activity, higher civic engagement, and intergenerational spillovers, as well as potential costs such as increased childcare needs or foregone parental income. Even so, the graduation channel alone suggests that the 4DSW yields hundreds of millions of dollars in added lifetime value per cohort, underscoring the policy's economic significance.

7. Conclusion

This paper provides the large-scale causal evidence on the effects of the 4DSW on high school students in Colorado, drawing on nearly two decades of linked administrative records covering more than one million unique students across successive cohorts. The breadth and depth of these data allow for comprehensive analysis of both academic performance and educational attainment. We find that the 4DSW leaves test participation unchanged but modestly improves achievement and meaningfully raises persistence and completion. Scaled to the size of a typical statewide cohort, these gains represent substantial improvements in educational attainment and human capital formation.

The benefits, however, are not evenly distributed. Students with longer and more consistent exposure to the 4DSW experience the largest and most reliable gains, while those exposed later show more limited improvements. Subgroup analysis further reveals that male, non ESL, and non homeless students benefit most consistently, whereas English learners and homeless students exhibit weaker or even negative effects. These disparities suggest that although the 4DSW raises average outcomes, it may also widen inequalities without targeted supports for vulnerable populations.

Analysis of mechanisms helps explain some of the observed effects in Colorado. The state's binding instructional hour requirements ensure that total classroom time is preserved, ruling out changes in instructional time as a key channel. Instead, teacher retention emerges as a central channel, as districts adopting the 4DSW experience significantly higher retention that strengthens student and teacher continuity and contributes to improved persistence and completion. We also find modest increases in supplies expenditures, which may have enhanced classroom environments, although the literature offers limited guidance on whether changes of this magnitude yield measurable achievement effects. In contrast, we find little evidence of systematic changes in purchased or community service expenditures that would indicate expanded extracurricular programming on the fifth day. The choice of non instructional day also appears consequential, with Monday-off districts recording more sustained improvements in attainment than Friday-off districts. While the data do not permit direct observation of how students and families use the additional day, potential mechanisms such as changes in student fatigue, mental health, parental labor supply, and enrichment opportunities remain important areas for future research.

Overall, the evidence from Colorado indicates that the 4DSW can modestly enhance high school achievement and meaningfully increase persistence and completion when implemented under supportive institutional conditions. The central challenge for policymakers is to sustain these benefits while addressing equity concerns to ensure that vulnerable students are not left behind. More broadly, the findings highlight that the consequences of calendar reforms depend critically on institutional design and local context. As districts and states continue to weigh the trade offs of shortened weeks, understanding not only average effects but also distributional impacts and underlying mechanisms will be essential for designing schedules that promote both performance and equity.

References

[1] Anderson, D. M., & Walker, M. B. (2015). Does shortening the school week impact student performance? Evidence from the four-day school week. *Education Finance and Policy, 10*(3), 314–349.

[2] Bloom, H. S., Hill, C. J., Black, A. R., & Lipsey, M. W. (2008). Performance trajectories and performance gaps as achievement effect-size benchmarks for educational interventions. *Journal of Research on Educational Effectiveness, 1*(4), 289–328.

[3] Borusyak, K., Jaravel, X., & Spiess, J. (2022). Revisiting event study designs: Robust and efficient estimation. *arXiv preprint* arXiv:2108.12419.

[4] Bound, J., Lovenheim, M. F., & Turner, S. (2010). Why have college completion rates declined? *American Economic Journal: Applied Economics, 2*(3), 129–157.

[5] 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.

[6] Cohodes, S., Eren, O., & Ozturk, O. (2023). The long run effects of a comprehensive teacher performance pay program on student outcomes. *NBER Working Paper No. 31056*. National Bureau of Economic Research. https://doi.org/10.3386/w31056

[7] Colorado Department of Education (CDE), Office of Field Services. (2016). *Instructional time guidance for Colorado school districts*. Denver, CO: Colorado Department of Education.

[8] Colorado Department of Education (CDE). (2024). *Four-day school week implementation and requirements*. Denver, CO: Colorado Department of Education.

[9] Colorado Department of Education (CDE). (2025). *Instructional time requirements for Colorado public schools*. Denver, CO: Colorado Department of Education.

[10] Chetty, R., Hendren, J., Kline, P., & Saez, E. (2014). Where is the land of opportunity? *Quarterly Journal of Economics, 129*(4), 1553–1623.

[11] Clear Creek School District RE-1. (n.d.). *Friday enrichment programs in Clear Creek.* Idaho Springs, CO: Clear Creek School District.

[12] Cunha, F., & Heckman, J. J. (2007). The technology of skill formation. *American Economic Review, 97*(2), 31–47.

[13] Dee, T. S., & Jacob, B. (2011). The impact of No Child Left Behind on student achievement. *Journal of Policy Analysis and Management, 30*(3), 418–446.

[14] Dynarski, M., & Gleason, P. (2002). How can we help? What we have learned from recent federal dropout prevention evaluations. *Journal of Education for Students Placed at Risk, 7*(1), 43–69. https://doi.org/10.1207/S15327671ESPR0701.

[15] Dynarski, S., Hyman, J., & Schanzenbach, D. W. (2013). Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion. *Journal of Policy Analysis and Management, 32*(4), 692–717.

[16] Fischer, S., & Argyle, D. (2018). Juvenile crime and the four-day school week. *Economics of Education Review, 64,* 31–39.

[17] Fitzpatrick, M. D., Grissmer, D., & Hastedt, S. (2011). What a difference a day makes: Estimating daily learning gains during kindergarten and first grade using a natural experiment. *Economics of Education Review, 30*(2), 269–279. https://doi.org/10.1016/j.econedurev.2010.08.004

[18] Fredricks, J. A., & Eccles, J. S. (2006). Is extracurricular participation associated with beneficial outcomes? Concurrent and longitudinal relations. *Developmental Psychology, 42*(4), 698–713. https://doi.org/10.1037/0012-1649.42.4.698

[19] Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics, 225*(2), 254–277.

[20] Hanushek, E. A., Kain, J. F., & Rivkin, S. G. (2004). Why public schools lose teachers. *Journal of Human Resources, 39*(2), 326–354. https://doi.org/10.2307/3559017

[21] Hanushek, E. A., & Raymond, M. E. (2005). Does school accountability lead to improved student performance? *Journal of Policy Analysis and Management, 24*(2), 297–327.

[22] Heckman, J. J. (2006). Skill formation and the economics of investing in disadvantaged children. *Science, 312*(5782), 1900–1902.

[23] Hill, C. J., Bloom, H. S., Black, A. R., & Lipsey, M. W. (2008). Empirical benchmarks for interpreting effect sizes in research. *Child Development Perspectives, 2*(3), 172–177.

[24] Ingersoll, R. M. (2001). Teacher turnover and teacher shortages: An organizational analysis. *American Educational Research Journal, 38*(3), 499–534. https://doi.org/10.3102/00028312038003499

[25] Jackson, C. K., Johnson, R. C., & Persico, C. (2016). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *Quarterly Journal of Economics, 131*(1), 157–218. https://doi.org/10.1093/qje/qjv036

[26] Jackson, C. K. (2018). What do test scores miss? The importance of teacher effects on non-test score outcomes. *Journal of Political Economy, 126*(5), 2072–2107.

[27] Jackson, C. K., Porter, S. C., Easton, J. Q., Blanchard, A., & Kiguel, S. (2020). School effects on socioemotional development, school-based arrests, and educational attainment. *American Economic Review: Insights, 2*(4), 491–508.

[28] Kim, B., Kim, M., & Park, G. (2024). The opioid crisis and firm skill demand: Evidence from job posting data. *Available at SSRN 4825126.*

[29] King, A. C. (2002). Time spent in extracurricular activities and adolescent academic achievement. *Journal of Youth and Adolescence, 31*(6), 437–447.

[30] Kini, T., & Podolsky, T. (2016). Does teaching experience increase teacher effectiveness? A review of the research. *Learning Policy Institute*.

[31] Kraft, M. A., Blazar, D., & Hogan, D. (2020). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research, 88*(4), 547–588.

[32] Kraft, M. A., Marinell, W. H., & Yee, D. S. (2020). School organizational contexts, teacher turnover, and student achievement: Evidence from panel data. *American Educational Research Journal, 57*(2), 273–313.

[33] Krueger, A. B. (1999). Experimental estimates of education production functions. *Quarterly Journal of Economics, 114*(2), 497–532.

[34] Lauer, P. A., Akiba, M., Wilkerson, S. B., Apthorp, H. S., Snow, D., & Martin-Glenn, M. L. (2006). Out-of-school-time programs: A meta-analysis of effects for at-risk students. *Review of Educational Research, 76*(2), 275–313. https://doi.org/10.3102/00346543076002275

[35] Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. *Economic Journal, 125*(588), 397–424.

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

[37] Lochner, L., & Moretti, E. (2004). The effect of education on crime: Evidence from prison inmates, arrests, and self-reports. *American Economic Review, 94*(1), 155–189.

[38] Morton, E. (2020). Four-day school weeks and adolescent outcomes: Evidence from Oklahoma. *Economics of Education Review, 77*, 101998.

[39] Morton, E. (2021). Effects of four-day school weeks on school finance and achievement: Evidence from Oklahoma. *Educational Researcher, 50*(1), 30–40.

[40] Morton, E. (2023). Effects of 4-day school weeks on older adolescents: Examining impacts of the schedule on academic achievement, attendance, and behavior in high school. *Educational Evaluation and Policy Analysis, 45*(1), 52–78.

[41] Murnane, R. J., Willett, J. B., & Boudett, K. P. (1995). Do high school dropouts benefit from obtaining a GED? *Educational Evaluation and Policy Analysis, 17*(2), 133–147.

[42] Murnane, R. J. (2013). U.S. high school graduation rates: Patterns and explanations. *Journal of Economic Literature, 51*(2), 370–422.

[43] Nickow, A., Oreopoulos, P., & Quan, V. (2020). The impressive effects of tutoring on preK–12 learning: A systematic review and meta-analysis of the experimental evidence. *NBER Working Paper No. 27476*. https://doi.org/10.3386/w27476

[44] Oreopoulos, P. (2007). Do dropouts drop out too soon? Wealth, health and happiness from compulsory schooling. *Journal of Public Economics, 91*(11–12), 2213–2229.

[45] Oreopoulos, P., & Salvanes, K. G. (2011). Priceless: The nonpecuniary benefits of schooling. *Journal of Economic Perspectives, 25*(1), 159–184.

[46] Patall, E. A., Cooper, H., & Allen, A. B. (2010). Extending the school day or school year: A systematic review of research (1985–2009). *Review of Educational Research, 80*(3), 401–436.

[47] Pashler, H., McDaniel, M., Rohrer, D., & Bjork, R. (2007). Learning styles: Concepts and evidence. *Psychological Science in the Public Interest, 9*(3), 105–119.

[48] Pashler, H., Rohrer, D., Cepeda, N. J., & Carpenter, S. K. (2007). Enhancing learning and retarding forgetting: Choices and consequences. *Psychonomic Bulletin & Review, 14*(2), 187–193. https://doi.org/10.3758/BF03194050

[49] Rivkin, S. G., & Schiman, J. C. (2015). Instruction time, classroom quality, and academic achievement. *Economic Journal, 125*(588), F425–F448.

[50] Ronfeldt, M., Loeb, S., & Wyckoff, J. (2013). How teacher turnover harms student achievement. *American Educational Research Journal, 50*(1), 4–36.

[51] Ross, S. L., Jiang, S., Young, M. F., Meyer, J. L., & Strambler, M. J. (2024). In-person learning during the pandemic: Student take-up and school-level effects of remote and hybrid instruction on student outcomes. *Proceedings of the National Academy of Sciences, 121*(33), e2407322121.

[52] Rouse, C. E. (2007). Consequences for the labor market. In C. R. Belfield & H. M. Levin (Eds.), *The price we pay: Economic and social consequences of inadequate education* (pp. 99–124). Brookings Institution Press.

[53] Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics, 225*(2), 175–199.

[54] Tharp, R. G. (2014). *Teaching transformed: Achieving excellence, fairness, inclusion, and harmony*. Routledge.

[55] Thompson, P. N. (2020). Does more time in school matter? The impact of four-day school weeks on student achievement. *Economics of Education Review, 76*, 101982.

[56] Thompson, P. N. (2021). Is four less than five? Effects of four-day school weeks on student achievement in Oregon. *Journal of Public Economics, 193*, 104308.

[57] Thompson, P. N., Tomayko, E. J., Gunter, K. B., & Schuna Jr, J. (2022). Impacts of the four-day school week on high school achievement and educational engagement. *Education Economics, 30*(5), 527–539.

[58] Thompson, P. N., & Ward, J. (2022). Only a matter of time? The role of time in school on four-day school week achievement impacts. *Economics of Education Review, 86*, 102198.

[59] Walker, T. J. (2020). The four-day school week: Research, implementation, and teacher perspectives. *Journal of Education Policy, 35*(7), 987–1005.

[60] Ward, J. (2019). The four-day school week and parental labor supply. *Available at SSRN 3301406.*

[61] Harris, M. (2023). The relationship between academic supplies and student academic achievement. *Undergraduate Research Journal for the Human Sciences, 22*(1). Retrieved from https://newprairiepress.org/urjhs/vol22/iss1/12

[62] Hilton, J., & Robinson, T. (2023). Affordable learning and student success: Evidence from open educational resources. *Open Praxis, 15*(2), 123–135. https://openpraxis.org/articles/10.55982/open

Table I. Summary Statistics: Student and Teacher Baseline Samples

	Four-Day Schools			Five-Day Schools
	All Time	Pre-4DSW	Post-4DSW	All Time
Panel A: Students				
Characteristics				
Male (%)	0.515	0.516	0.514	0.513
White (%)	0.532	0.546	0.520	0.580
Black (%)	0.018	0.019	0.015	0.059
Hispanic (%)	0.401	0.384	0.410	0.288
Asian (%)	0.015	0.017	0.011	0.035
English Learner (%)	0.095	0.096	0.092	0.100
Migrant Student (%)	0.013	0.015	0.017	0.004
Gifted Student (%)	0.036	0.035	0.038	0.055
Homeless Student (%)	0.019	0.021	0.014	0.014
Total Students	323,003	178,231	144,773	902,728
Academic Outcomes				
Test Taking Rate	0.832	0.840	0.825	0.851
English Test Score	17.42	17.30	17.60	18.99
Math Test Score	18.16	18.05	18.32	19.62
Grade Repetition (%)	0.148	0.156	0.135	0.135
Dropout (%)	0.121	0.140	0.109	0.118
On-Time Graduation (%)	0.720	0.716	0.736	0.747
Overall Graduation (%)	0.734	0.731	0.745	0.732
Panel B: Teachers				
Characteristics				
Male (%)	0.271	0.269	0.273	0.261
White (%)	0.928	0.929	0.926	0.922
Black (%)	0.002	0.002	0.002	0.004
Hispanic (%)	0.055	0.054	0.057	0.060
Asian (%)	0.002	0.002	0.002	0.006
Bachelor's Degree (%)	0.652	0.655	0.647	0.579
Master's Degree (%)	0.332	0.329	0.336	0.403
Years of Experience	9.74	9.72	9.77	9.13
Total Teachers	15,266	8,229	7,037	101,776
Labor Market Outcomes				
Retention (%)	0.755	0.750	0.760	0.743
Transfer (%)	0.061	0.062	0.059	0.053
Exit (%)	0.181	0.185	0.176	0.206

Notes: This table reports means for the baseline student and teacher samples. "Pre-4DSW" refers to cohorts or years before a district's adoption of the four-day school week, and "Post-4DSW" refers to cohorts following adoption. The five-day school column provides corresponding means for traditional districts.

Table II. Effect of 4DSW on High School Students' Academic Outcomes

	Participation	Achievement		Progression		Attainment	
	Taking Tests	English Score	Math Score	Repetition	Dropout	Graduation	On-Time Graduation
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Full Exposure	0.009	0.027**	0.025**	-0.001	-0.021*	0.032***	0.029***
	(0.010)	(0.009)	(0.007)	(0.004)	(0.011)	(0.011)	(0.011)
Partial Exposure	-0.002	0.006	0.002	-0.001	-0.010	0.013*	0.014***
	(0.009)	(0.009)	(0.007)	(0.004)	(0.009)	(0.007)	(0.011)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,224,731	720,489	720,473	1,224,731	1,224,731	1,224,731	1,224,731
Control Mean	0.851	0.006	0.006	0.135	0.118	0.747	0.732

Notes: This table reports difference-in-differences estimates of the impact of adopting 4DSW on high school student outcomes. Outcomes are grouped into four categories: *Participation* (an indicator for taking any standardized test during high school); *Achievement* (standardized ACT/SAT English and math scores); *Progression* (indicators for course repetition in grades 9–12 and for dropout); and *Attainment* (graduation and on-time graduation). "Full exposure" refers to cohorts attending 4DSW schools continuously from grade 9 onward, while "partial exposure" refers to cohorts first exposed in grades 10–12. All specifications include school and entry-cohort fixed effects and the full set of controls. Robust standard errors, clustered at the school-district level, are reported in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table III. Effect of the Four-Day School Week on Student Outcomes by Timing and Duration of Exposure

	Taking Tests	**English Score**	**Math Score**	**Grade Repetition**	**Dropout**	**Graduation**	**On-Time Graduation**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Students Exposed Since Middle School							
4DSW Indicator	0.015	0.019*	0.023*	-0.006	-0.019**	0.043**	0.041**
	(0.013)	(0.016)	(0.010)	(0.010)	(0.006)	(0.008)	(0.009)
Panel B: Students First Exposed in High School							
4DSW Indicator	0.007	0.023*	0.026*	-0.001	-0.009	0.017**	0.016**
	(0.014)	(0.015)	(0.005)	(0.005)	(0.009)	(0.004)	(0.005)
Panel C: Incremental Effect of an Additional Year of Exposure							
4DSW (Years of Exposure)	0.003	0.004	0.003	-0.002	-0.004*	0.007***	0.006***
	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,224,731	720,489	720,473	1,224,731	1,224,731	1,224,731	1,224,731

Notes: Each panel reports results from difference-in-differences specifications estimating the impact of the 4DSW on high school student outcomes. Panel A includes cohorts whose districts adopted the 4DSW while they were in middle school (grades 6–8). Panel B includes cohorts first exposed to the 4DSW upon entering ninth grade. Panel C estimates a specification relating outcomes to the cumulative number of years of exposure to the 4DSW, thereby capturing potential dose–response effects. All models include school and entry-cohort fixed effects and control for student demographics and baseline characteristics. Robust standard errors, clustered at the school district level, are shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table IV. Heterogeneous Effects of 4DSW on Student Outcomes by Subgroup

	Gender		**Race/Ethnicity**			**English Learner Status**		**Homeless Status**	
	Male	Female	White	Black	Hispanic	ESL	Non-ESL	Homeless	Non-Homeless
Panel A: Participation									
Test Taking	0.011	0.006	0.004	0.018	0.018	0.028	0.006	0.005	0.009
	(0.011)	(0.011)	(0.009)	(0.023)	(0.015)	(0.023)	(0.009)	(0.041)	(0.010)
Panel B: Achievement									
English (z)	0.045***	0.009	0.009	0.029	0.062**	-0.059	0.021*	-0.039	0.027**
	(0.014)	(0.018)	(0.018)	(0.054)	(0.028)	(0.044)	(0.013)	(0.064)	(0.014)
Math (z)	0.036**	0.014	0.009	0.010	0.007**	-0.093**	0.020*	-0.010	0.026**
	(0.015)	(0.014)	(0.018)	(0.088)	(0.031)	(0.039)	(0.012)	(0.064)	(0.012)
Panel C: Progression									
Grade Repetition	-0.010	-0.019	-0.008	0.003	0.000	0.034	-0.020	0.025	-0.015
	(0.014)	(0.013)	(0.007)	(0.017)	(0.019)	(0.022)	(0.011)	(0.062)	(0.013)
Dropout Rate	-0.025**	-0.016	-0.009	-0.028*	-0.017	0.001	-0.023**	0.015	-0.021*
	(0.013)	(0.011)	(0.007)	(0.018)	(0.013)	(0.018)	(0.011)	(0.027)	(0.012)
Panel D: Attainment									
Graduation Rate	0.042***	0.023*	0.026***	0.043	0.025**	-0.017	0.036***	-0.084	0.033***
	(0.011)	(0.013)	(0.008)	(0.039)	(0.013)	(0.034)	(0.010)	(0.083)	(0.011)
On-Time Graduation	0.036***	0.024*	0.024***	0.046	0.021	-0.028	0.026**	-0.121*	0.031***
	(0.011)	(0.013)	(0.008)	(0.033)	(0.013)	(0.034)	(0.034)	(0.076)	(0.011)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry-Cohort Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	628,368	596,344	703,884	66,181	367,717	121,659	1,103,030	18,530	1,206,134

Notes: Each cell reports difference-in-differences estimates of the effect of adopting the four-day school week (4DSW) for the indicated subgroup. Robust standard errors clustered at the school-district level are shown in parentheses. Outcomes are grouped into *Participation* (test taking), *Achievement* (English and math test scores), *Progression* (course repetition and dropout), and *Attainment* (graduation and on-time graduation). All specifications include school and entry-cohort fixed effects and a full set of control variables.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table V. Effect of 4DSW on Teacher Outcomes

	Retention (1)	**Transfer** (2)	**Exit** (3)
4DSW	0.027***	0.002	-0.008
	(0.007)	(0.005)	(0.008)
School Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
N	191,096	191,096	191,096
Control Mean	0.743	0.054	0.206

Notes: The table reports difference-in-differences estimates of the impact of 4DSW on high school teacher outcomes. The dependent variables are: an indicator for teachers retained at the same school (column 1), an indicator for teachers transferring to another school (column 2), and an indicator for teachers exiting the education sector (column 3). Robust standard errors are clustered at the school district level and shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

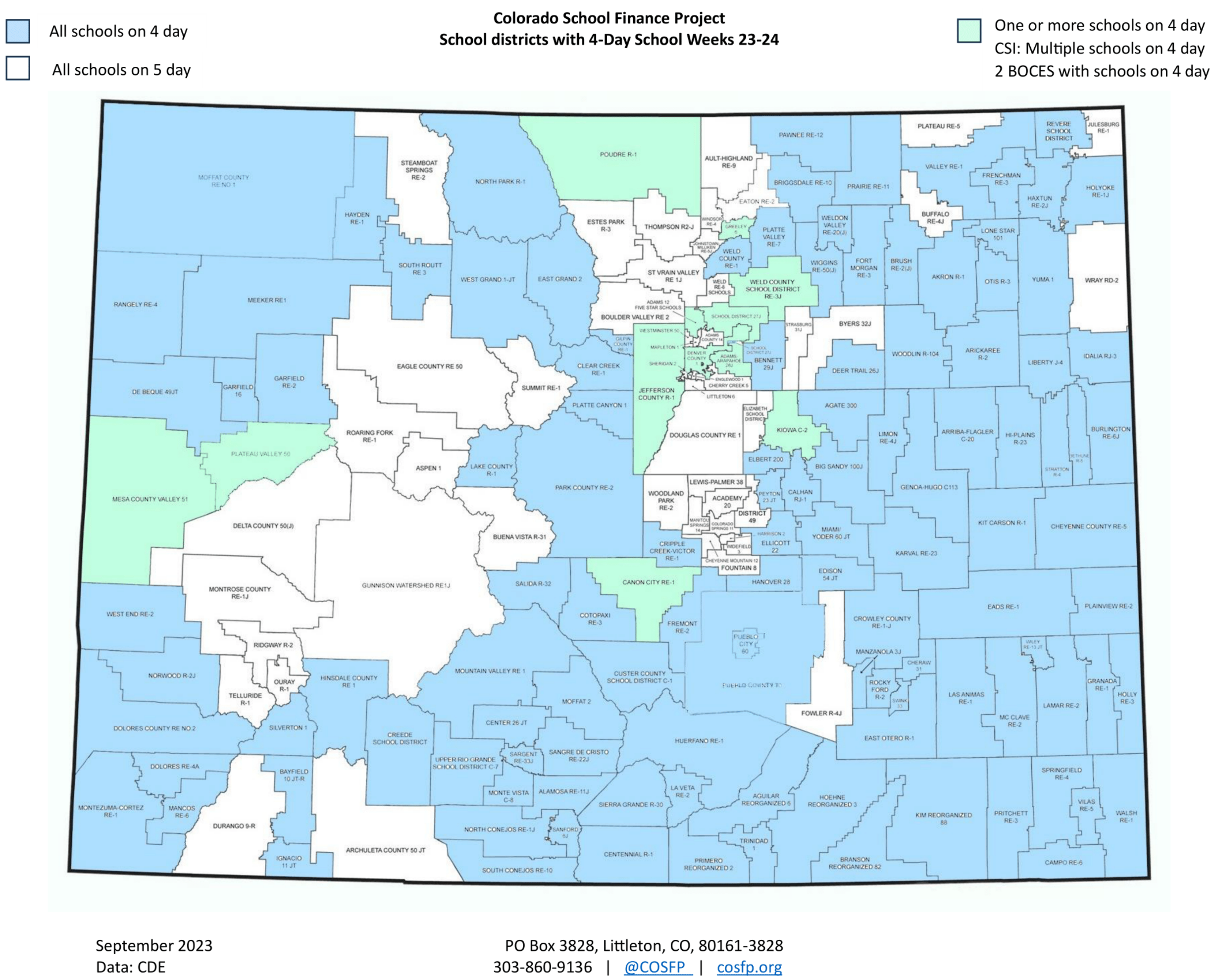

Figure 1. Map of 4DSW Districts in Colorado.

Notes: The figure shows the map of 4DSW districts in Colorado for the 2023–2024 academic year, as provided by the Colorado School Finance Project.

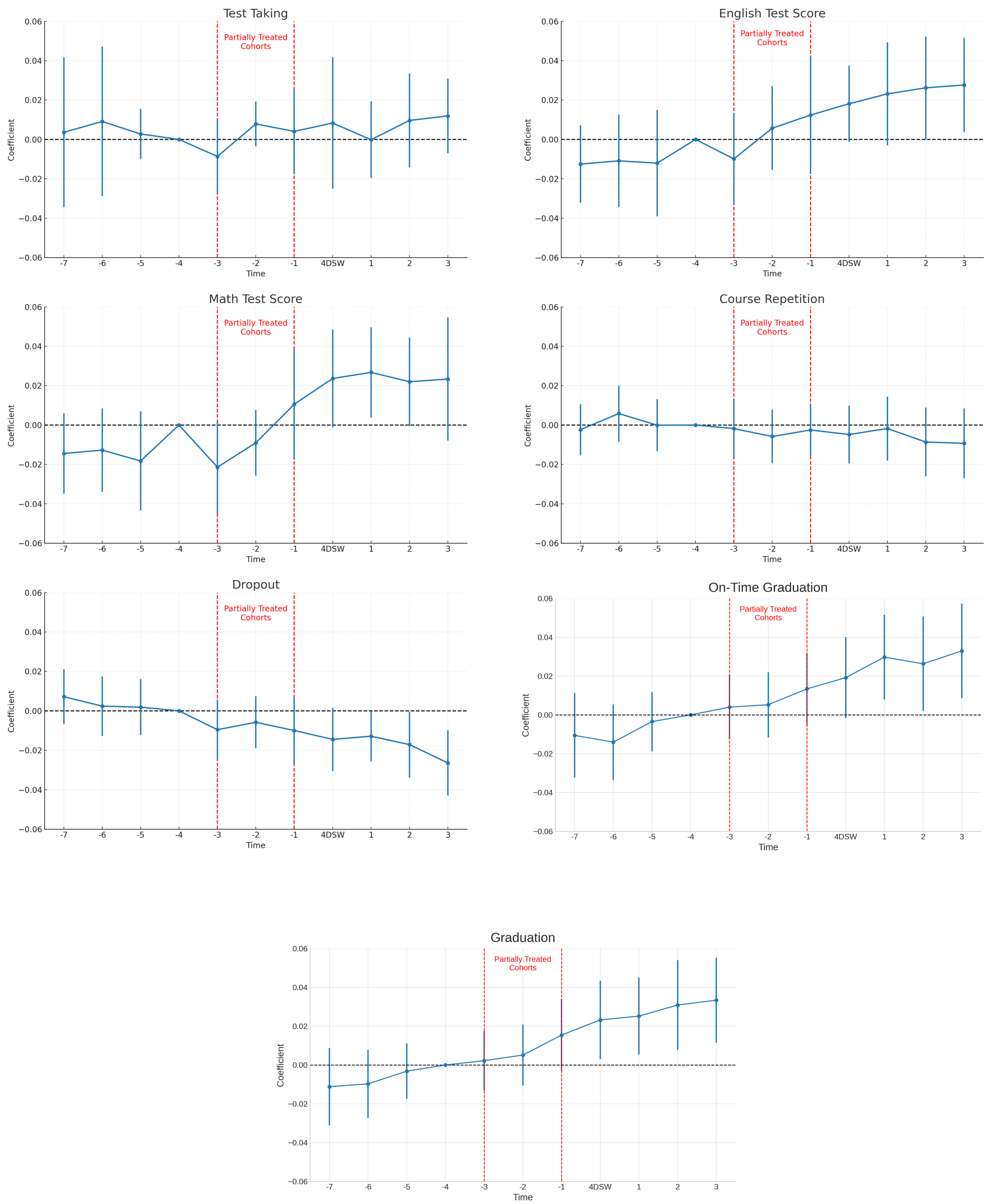

Figure 2. Event study of the effect of 4DSW on student outcomes.

Notes: The figure graphs the regression coefficient estimates and their 90% confidence intervals (vertical lines). Cohorts entered 9th grade 4 years before 4DSW implementation is the omitted category.

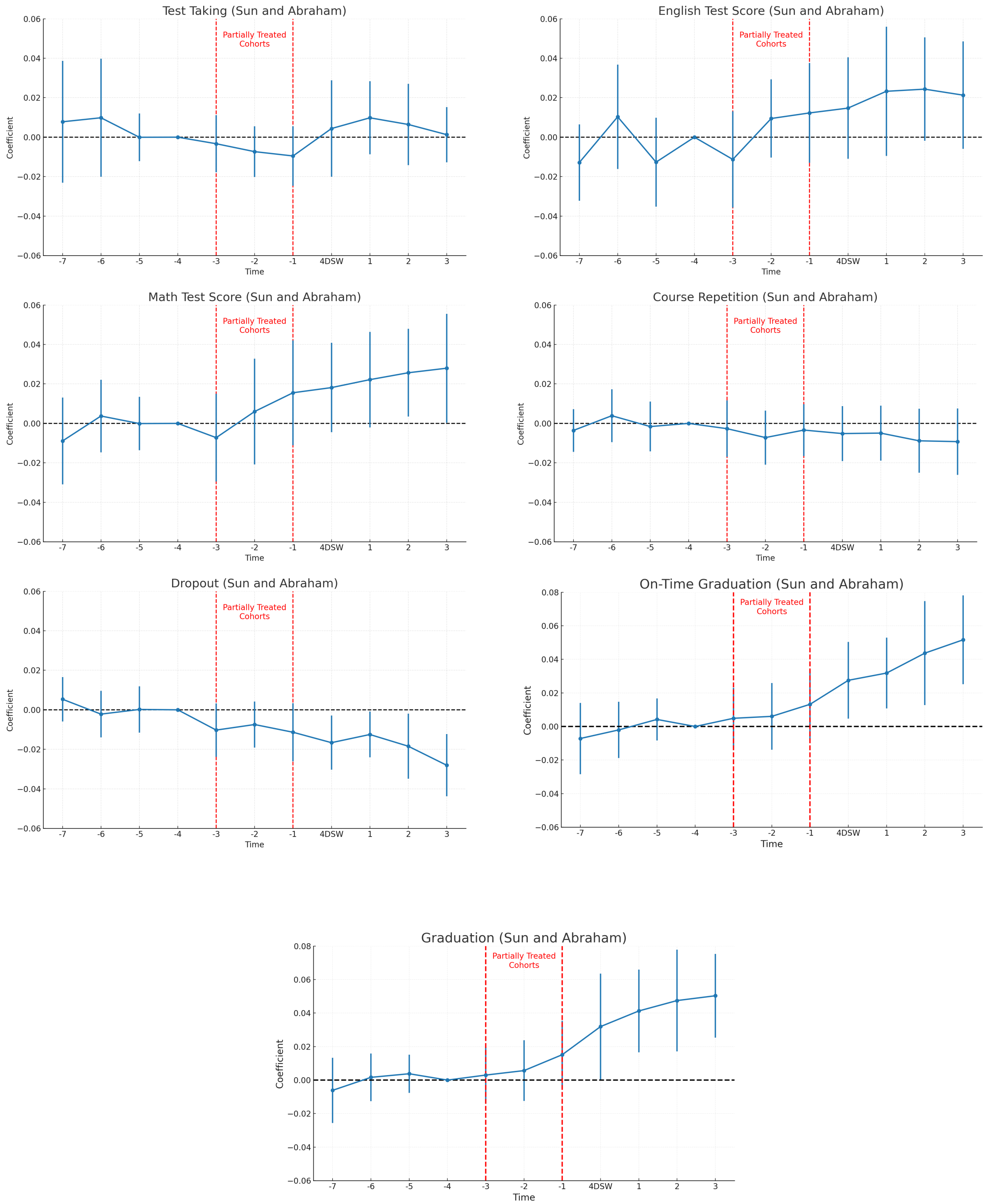

Figure 3. Event study of the effect of 4DSW on student outcomes: Sun and Abraham (2021).

Notes: The figure graphs the regression coefficients obtained by using interaction weighted estimator from Sun & Abraham (2021) and their 90% confidence intervals (vertical lines). Cohorts entered 9th grade 4 years before 4DSW implementation is the omitted category.

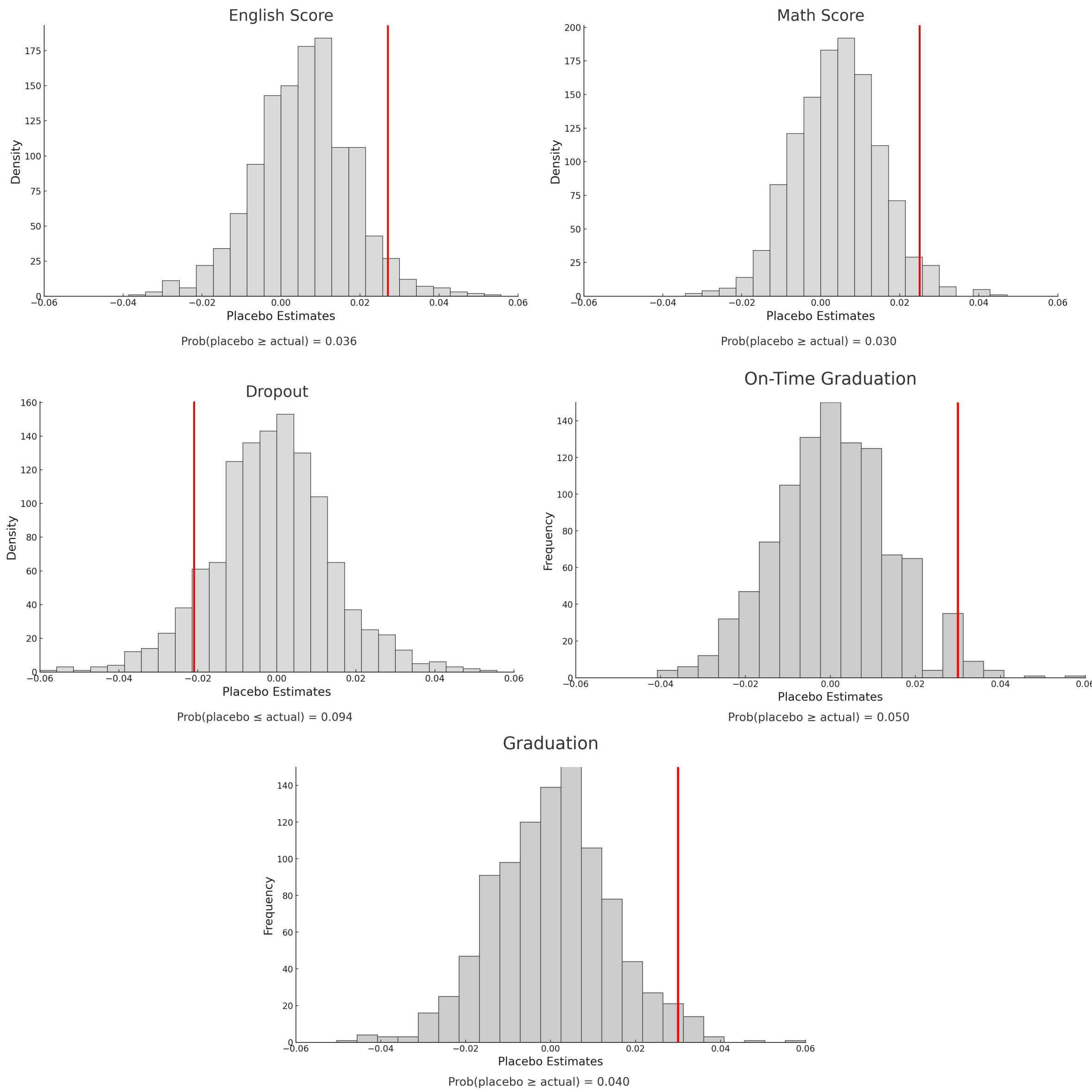

Figure 4. Placebo estimates.

Notes: The figure graphs the distribution of placebo coefficient estimates resulting from 1000 sets of random assignments of districts to 4DSW adoption. The red vertical lines represent the true estimates from Table 2.

Appendix Tables and Figures

Table A.1. Trends in Characteristics Before 4DSW Adoption and Predicting 4DSW Adoption Year

	Trend	**4DSW Adoption Year**
	(1)	(2)
Panel A: Student-level characteristics		
Total Enrollment	274.17	0.001
	(184.59)	(0.001)
Percent of Male Students	0.006	-0.012
	(0.005)	(0.022)
Percent of Minor Students	0.014	0.009
	(0.014)	(0.019)
Percent of ESL Students	0.001	-0.004
	(0.002)	(0.011)
Percent of Migrant Students	0.002	0.003
	(0.001)	(0.008)
Percent of Special Education	-0.001	0.003
	(0.002)	(0.006)
Percent of Gifted Students	-0.005	-0.008
	(0.009)	(0.025)
Percent of Homeless Students	0.001	-0.011
	(0.001)	(0.020)
Panel B: Teacher-level characteristics		
Total Teachers	8.39	-0.000
	(6.48)	(0.001)
Percent of Male Teachers	-0.017**	-0.014
	(0.009)	(0.028)
Percent of Minor Teachers	0.014	0.007
	(0.012)	(0.016)
% Teachers: High School Degree	0.002	-0.003
	(0.003)	(0.009)
% Teachers: Bachelor's Degree	0.019	0.005
	(0.014)	(0.012)
% Teachers: Master's Degree	-0.001	0.002
	(0.013)	(0.010)
N	3,674	119

Notes: Column (1) reports estimates of α_1 from equation (5). Each variable is obtained from a separate regression including district and year fixed effects; robust standard errors clustered at the district level are in parentheses. Column (2) tests whether the timing of 4DSW adoption is correlated with baseline (2005) district characteristics. Heteroskedasticity-robust standard errors are reported in parentheses for Column (2).
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.2. Estimations of 4DSW Entry Decisions by Districts

	English Score	**Math Score**	**Dropout**	**Graduation**	**Teacher Retention**	**Teacher Transfer**	**Teacher Exit**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
2-Year Trend	0.001	-0.002	-0.008	0.002	0.005	-0.003	0.001
	(0.012)	(0.012)	(0.010)	(0.002)	(0.004)	(0.004)	(0.004)
3-Year Trend	0.007	0.006	-0.008	0.002	0.006	-0.004	0.001
	(0.011)	(0.013)	(0.009)	(0.002)	(0.004)	(0.004)	(0.004)
4-Year Trend	0.009	0.011	-0.009	0.002	0.007	-0.005	0.002
	(0.010)	(0.012)	(0.008)	(0.002)	(0.004)	(0.004)	(0.003)
Observations	115	115	115	115	115	115	115

Notes: This table reports estimates of whether the year of 4DSW implementation is correlated with outcomes prior to adoption. The “2-, 3-, and 4-Year Trends” correspond to the average student and teacher outcomes in the two, three, and four years preceding 4DSW adoption. Standard errors are shown in parentheses.

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

Table A.3. Student and Teacher Sorting Following 4DSW Implementation

Student Characteristic	**Student Sorting**	**Teacher Characteristic**	**Teacher Sorting**
	(1)		(2)
Total Enrollment	-272.80	Total Teachers	-7.53
	(170.89)		(6.57)
Percent of Male Students	-0.006	Percent of Male Teachers	0.024**
	(0.005)		(0.010)
Percent of Minor Students	-0.018	Percent of Minor Teachers	0.009
	(0.012)		(0.015)
Percent of ESL Students	-0.001	Percent of High School Degree	0.005
	(0.002)		(0.004)
Percent of Migrant Students	-0.001	Percent of Bachelor's Degree	-0.008
	(0.001)		(0.015)
Percent of Special Education	0.001	Percent of Master's Degree	-0.044
	(0.002)		(0.015)
Percent of Gifted Students	-0.005		
	(0.006)		
Percent of Homeless Students	-0.001		
	(0.001)		
District Fixed Effects	Yes	District Fixed Effects	Yes
Year Fixed Effects	Yes	Year Fixed Effects	Yes
N	3,674	Observations	3,674

Notes: The table reports how adoption of the 4DSW affects district-level student and teacher composition. Student characteristics and their sorting estimates are presented in the left two columns, while teacher characteristics and their sorting estimates are shown in the right two columns.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.4. Effects of 4DSW on Student Attrition and Switching

	Attrition (Entry)	**Attrition (Exit)**	**Switching Schools**
	(1)	(2)	(3)
4DSW	-0.001	0.001	-0.013
	(0.002)	(0.002)	(0.012)
School Fixed Effects	Yes	Yes	Yes
Entry Cohort Fixed Effects	Yes	Yes	Yes
N	1,224,731	1,224,731	1,224,731

Notes: This table reports estimated effects of 4DSW adoption on student attrition and school switching. *Attrition (Entry)* equals one if a student is observed in any grade but does not appear in any Colorado public high school in ninth grade, capturing attrition prior to high-school entry. *Attrition (Exit)* equals one if a student is observed in ninth grade but never reappears in grades ten through twelve and is not classified as a dropout or graduate, capturing attrition after entering high school. *Switching Schools* equals one if a student changes schools at any point. All specifications include school and entry-cohort fixed effects. Robust standard errors, clustered at the district level, are reported in parentheses.

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

Table A.5. Effect of 4DSW on High School Students' Academic Outcomes: Robustness Checks

	Taking Tests	**English Score**	**Math Score**	**Grade Repetition**	**Dropout**	**Graduation**	**On-Time Graduation**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Propensity Score Matching							
4DSW	0.003	0.015	0.015*	-0.026*	-0.027**	0.030***	0.032***
	(0.011)	(0.012)	(0.006)	(0.013)	(0.010)	(0.011)	(0.009)
N	616,684	377,922	377,922	616,684	616,684	616,684	616,684
Panel B: Random Forest Classifier							
4DSW	0.006	0.026*	0.023**	-0.016	-0.020*	0.026**	0.028***
	(0.010)	(0.014)	(0.012)	(0.013)	(0.011)	(0.010)	(0.009)
N	834,106	500,994	500,944	834,106	834,106	834,106	834,106
Panel C: Inverse Probability Matching							
4DSW	0.003	0.014	0.013*	-0.018	-0.018*	0.022**	0.024***
	(0.010)	(0.012)	(0.009)	(0.012)	(0.010)	(0.010)	(0.009)
N	1,224,731	720,473	720,473	1,224,731	1,224,731	1,224,731	1,224,731
Panel D: Excluding Control Variables							
4DSW	0.008	0.027*	0.025*	-0.001	-0.0212*	0.032***	0.029***
	(0.008)	(0.012)	(0.011)	(0.008)	(0.011)	(0.010)	(0.010)
N	1,224,731	720,473	720,473	1,224,731	1,224,731	1,224,731	1,224,731
Panel E: District Specific Linear Pre-trend							
4DSW	0.007	0.020*	0.021*	-0.008	-0.023*	0.025**	0.027**
	(0.013)	(0.011)	(0.010)	(0.006)	(0.012)	(0.011)	(0.012)
N	1,224,731	720,473	720,473	1,224,731	1,224,731	1,224,731	1,224,731
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Each panel reports difference-in-differences estimates of the effect of 4DSW adoption on high school students' academic outcomes under alternative specifications. Panel A matches treated and control schools using propensity score matching. Panel B applies a random forest classifier to construct matched comparison groups. Panel C uses inverse probability matching to reweight comparison units. Panel D re-estimates the baseline model without any controls. Panel E incorporates district-specific linear pre-trends following Goodman-Bacon (2021). Robust standard errors clustered at the school district level are in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.6. Effect of 4DSW on High School Students' Academic Outcomes: Additional Estimates

	Taking Tests	**English Score**	**Math Score**	**Grade Repetition**	**Dropout**	**Graduation**	**On-Time Graduation**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Excluding Partially Treated Cohorts							
4DSW	0.009	0.030**	0.028**	-0.015*	-0.020*	0.021**	0.022**
	(0.011)	(0.013)	(0.012)	(0.013)	(0.011)	(0.011)	(0.009)
N	1,182,640	1,182,640	1,182,640	1,182,640	1,182,640	1,182,640	1,182,640
Panel B: Pre-COVID Sample							
4DSW	0.017	0.023*	0.026*	-0.002	-0.014	0.030***	0.028***
	(0.010)	(0.014)	(0.015)	(0.008)	(0.013)	(0.011)	(0.011)
N	810,857	538,175	538,175	810,857	810,857	810,857	810,857
Panel C: Excluding BOCES and Charter Schools							
4DSW	0.009	0.023*	0.019*	-0.002	-0.021*	0.033**	0.031***
	(0.010)	(0.013)	(0.011)	(0.009)	(0.011)	(0.011)	(0.011)
N	1,194,796	704,181	704,181	1,194,796	1,194,796	1,194,796	1,194,796
Panel D: Rural/Small Districts Only							
4DSW	-0.001	0.024*	0.016	-0.013	-0.019**	0.034***	0.035***
	(0.010)	(0.014)	(0.013)	(0.010)	(0.008)	(0.009)	(0.009)
N	268,900	155,110	155,110	268,900	268,900	268,900	268,900
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes: Panel A excludes cohorts partially exposed at baseline (i.e., students already in grades 10–12 at policy start). Panel B restricts to cohorts graduating prior to the COVID-19 pandemic. Panel C drops BOCES (Boards of Cooperative Educational Services) and charter schools. Panel D restricts to rural/small districts. All specifications include school and entry-cohort fixed effects and a set of covariates. Robust standard errors, clustered at the school district level, are shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.7. Effect of 4DSW on Test Scores: ACT- and SAT-Based Cohorts

	ACT-Based Cohorts	**SAT-Based Cohorts**
	(1)	(2)
English Score	0.021*	0.028*
	(0.012)	(0.017)
Math Score	0.027*	0.026*
	(0.015)	(0.014)
School Fixed Effects	Yes	Yes
Entry Cohort Fixed Effects	Yes	Yes
N	419,636	300,455
Controls	Yes	Yes

Notes: This table reports estimates of the effect of 4DSW on standardized test outcomes, separately for ACT-based and SAT-based student cohorts. Standardized scores are normalized within each assessment. Robust standard errors clustered at the school district level are shown in parentheses.
* $p < 0.10$, $^{*}*$ $p < 0.05$, $^{*}**$ $p < 0.01$

Table A.8. Effect of 4DSW on Teacher Outcomes (All)

	Retention	**Transfer**	**Turnover**
	(1)	(2)	(3)
4DSW	0.019*	0.001	-0.007
	(0.010)	(0.004)	(0.007)
School Fixed Effects	Yes	Yes	Yes
Year Fixed Effects	Yes	Yes	Yes
Controls	Yes	Yes	Yes
N	667,533	667,533	667,533
Control Mean	0.751	0.053	0.199

Notes: The table reports difference-in-differences estimates of the impact of 4DSW on teacher outcomes (1st grade to 12th grade). The dependent variables are: an indicator for teachers retained at the same school (column 1), an indicator for teachers transferring to another school (column 2), and an indicator for teachers exiting the education sector (column 3). Robust standard errors are clustered at the school district level and shown in parentheses.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A.9. Effect of 4DSW on District Annual Instructional Hours

	Annual Instructional Hours
	(1)
4DSW	23.64
	(27.08)
District Fixed Effects	Yes
Year Fixed Effects	Yes
N	3,592
Controls	Yes

Notes: This table reports estimates of the effect of 4DSW on annual instructional hours. The dependent variable is total district-level annual instructional hours. Robust standard errors are clustered at the school district level and shown in parentheses.
* $p < 0.10$, $^{*}*$ $p < 0.05$, $^{*}**$ $p < 0.01$

Table A.10. Effect of 4DSW on Student Outcomes by Day Off

	Taking Tests	**English Score**	**Math Score**	**Grade Repetition**	**Dropout**	**Graduation**	**On-Time Graduation**
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Schools with Friday Off							
4DSW	0.005	0.021	0.020*	0.017*	-0.12	0.012	0.006
	(0.009)	(0.013)	(0.011)	(0.010)	(0.012)	(0.015)	(0.015)
Panel B: Schools with Monday Off							
4DSW	0.006	0.021	0.020*	-0.002	-0.012	0.025**	0.024**
	(0.009)	(0.013)	(0.011)	(0.009)	(0.013)	(0.010)	(0.010)
School Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Entry Cohort Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1,224,731	720,473	720,473	1,224,731	1,224,731	1,224,731	1,224,731
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Notes:This table reports the estimates of the effect of adopting the four-day school week (4DSW) on student outcomes, distinguishing between schools that take Friday off (Panel A) and those that take Monday off (Panel B). Robust standard errors are clustered at the school-district level and shown in parentheses. The coefficient estimates for graduation and on-time graduation are statistically different across the two schedule types ($p < 0.05$).

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

Table A.11. Effect of 4DSW on Teacher Salaries and Benefits

	Teacher Salaries	**Teacher Benefits**
	(1)	(2)
4DSW	-0.003	-0.002
	(0.040)	(0.054)
District Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
N	1,636	1,636
Controls	Yes	Yes

Notes: This table reports estimates of the effect of 4DSW adoption on the log of teacher salaries and benefits. Robust standard errors clustered at the district level are shown in parentheses.
* $p < 0.10$, $^{*}*$ $p < 0.05$, $^{*}**$ $p < 0.01$

Table A.12. Effect of 4DSW on District Expenditures: Instructional and Supplies Categories

	Instructional Expenditures	**Supplies Expenditures**
	(1)	(2)
4DSW	0.008	0.183**
	(0.018)	(0.095)
District Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
N	1,636	1,636
Controls	Yes	Yes

Notes: This table reports estimates of the effect of 4DSW adoption on the log of district expenditure outcomes. Instructional expenditures include classroom-related costs directly tied to teaching. Supplies expenditures capture spending on textbooks, consumables, technology, and other classroom materials. Robust standard errors clustered at the district level are shown in parentheses.
* $p < 0.10$, $^{*}*$ $p < 0.05$, $^{*}**$ $p < 0.01$

Table A.13. Effect of 4DSW on District Expenditures: Purchased and Community Services

	Purchased Services	**Community Services**
	(1)	(2)
4DSW	-0.117	-0.131
	(0.088)	(0.104)
District Fixed Effects	Yes	Yes
Year Fixed Effects	Yes	Yes
N	1,636	1,636
Controls	Yes	Yes

Notes: This table reports estimates of the effect of 4DSW adoption on the log of district expenditure outcomes. Purchased services include payments to external providers of educational or extracurricular programs, while community services encompass spending on student and family engagement outside of regular instruction. Robust standard errors clustered at the district level are shown in parentheses.
$^{*}\ p < 0.10$, $^{*}*\ p < 0.05$, $^{*}**\ p < 0.01$

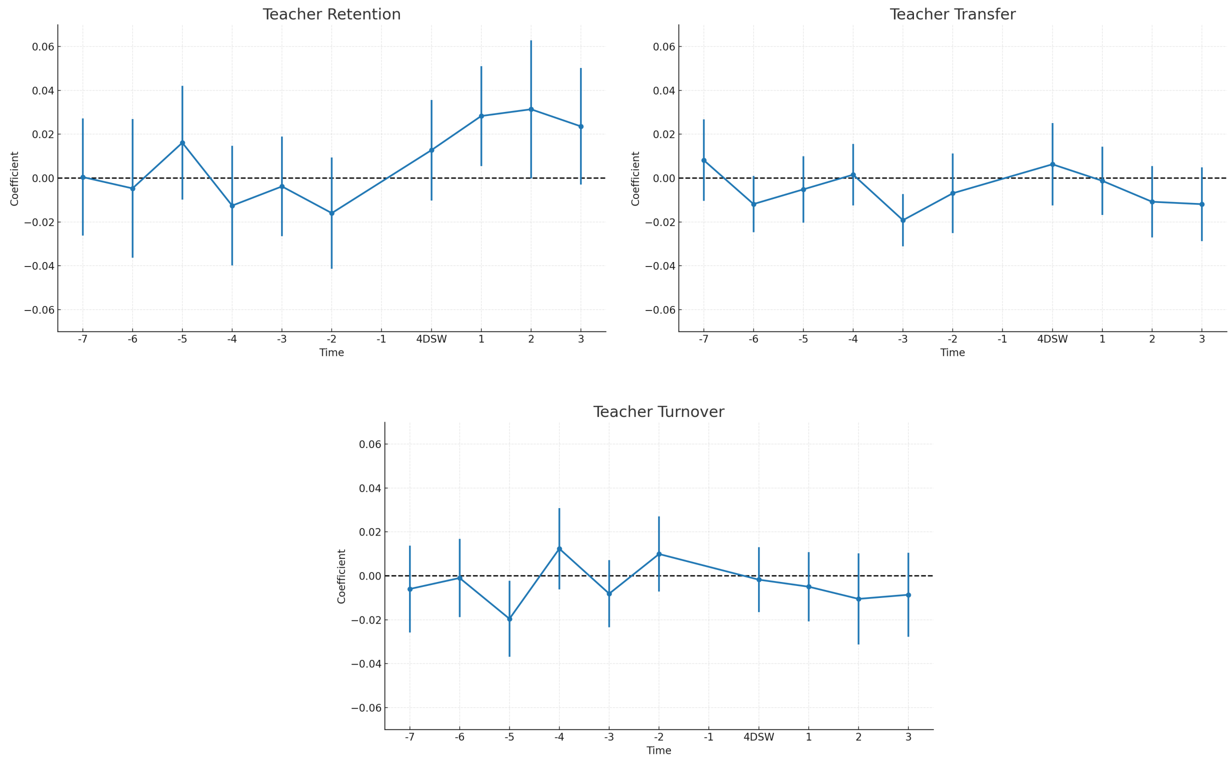

Figure A.1. Event study of the effect of 4DSW on teacher outcomes.

Notes: The figure graphs the regression coefficient estimates and their 90% confidence intervals (vertical lines). Year prior to 4DSW implementation is the omitted category.

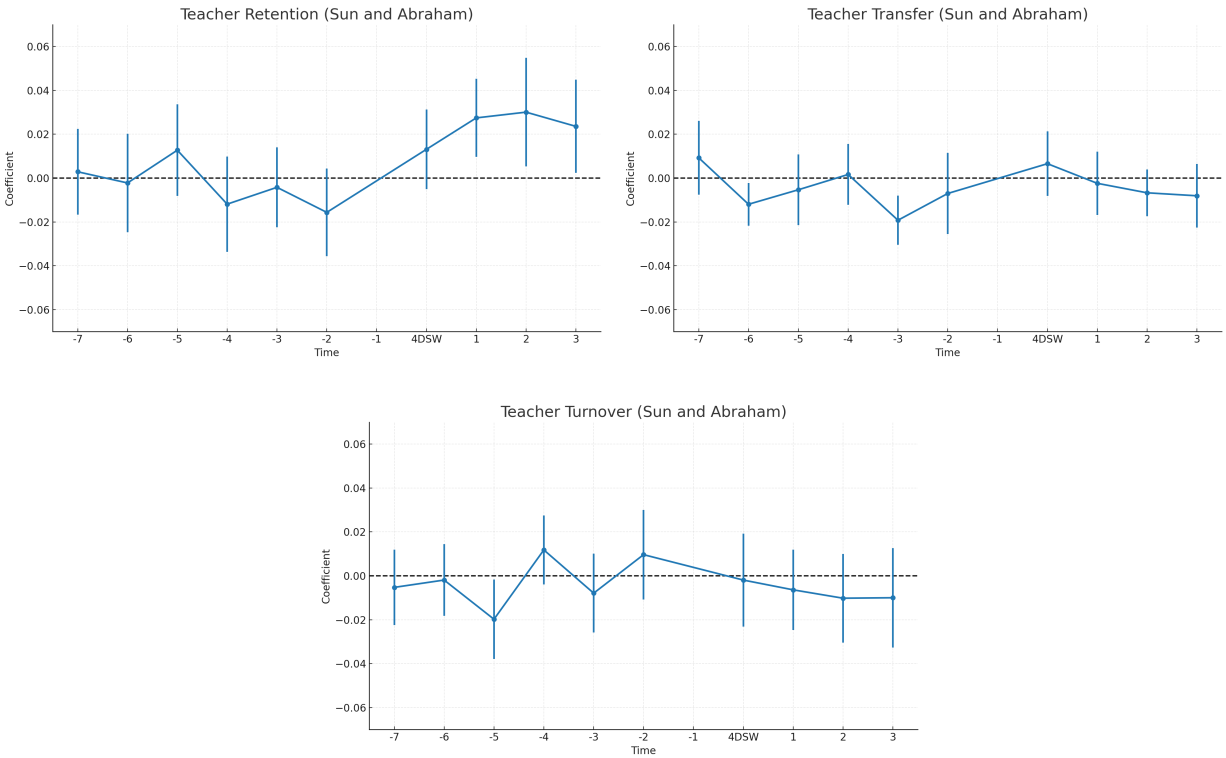

Figure A.2. Event study of the effect of 4DSW on teacher outcomes: Sun and Abraham (2021).

Notes: The figure graphs the regression coefficients obtained by using interaction weighted estimator from Sun & Abraham (2021) and their 90% confidence intervals (vertical lines). Year prior to 4DSW implementation is the omitted category.

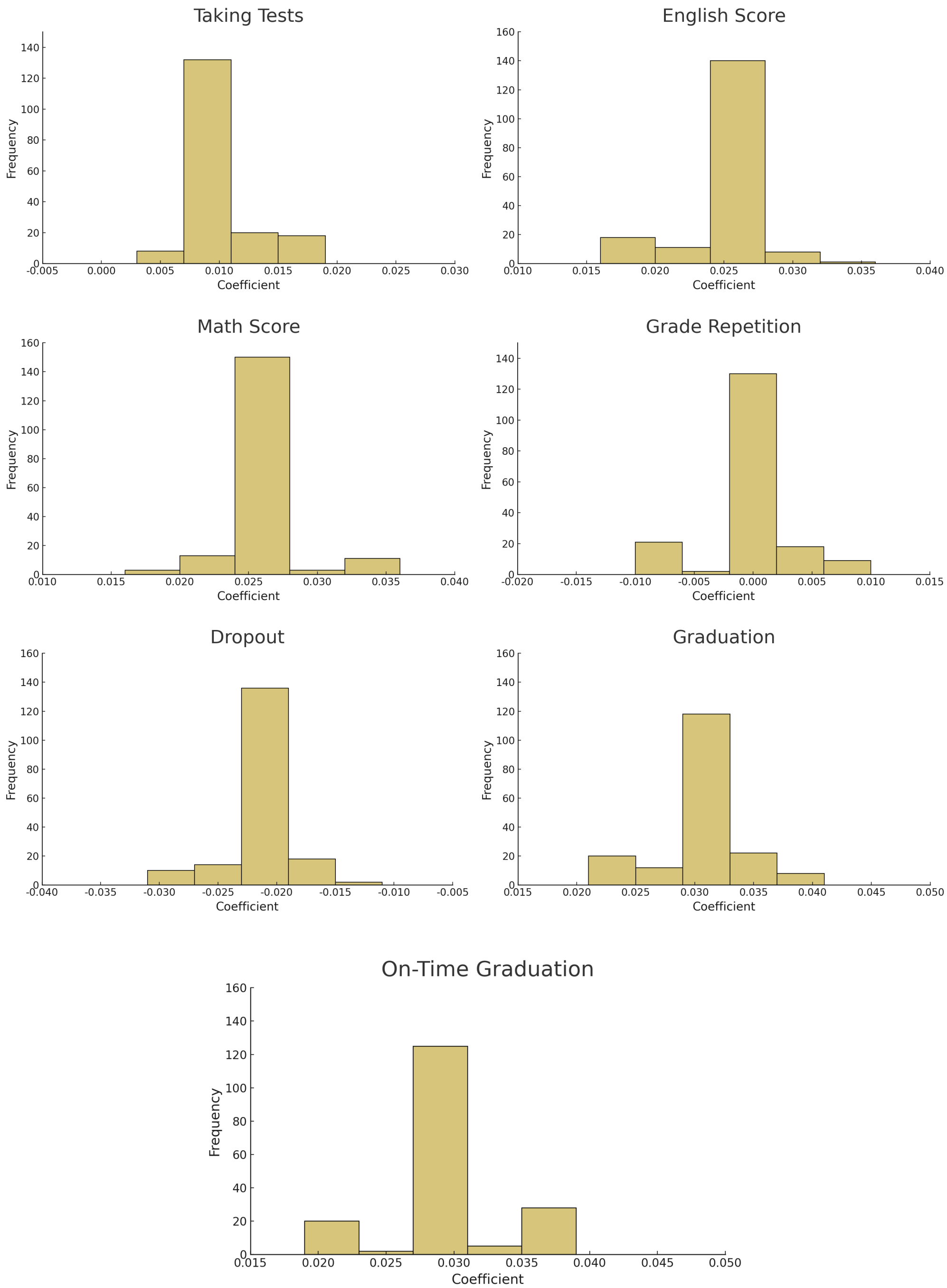

Figure A.3. Leave-one-district out estimations.

Notes: This figure shows the distribution of the leave-one-district-out coefficient estimates of 4DSW on student outcomes. The baseline specification for student outcomes is estimated repeatedly, each time removing a different school district.

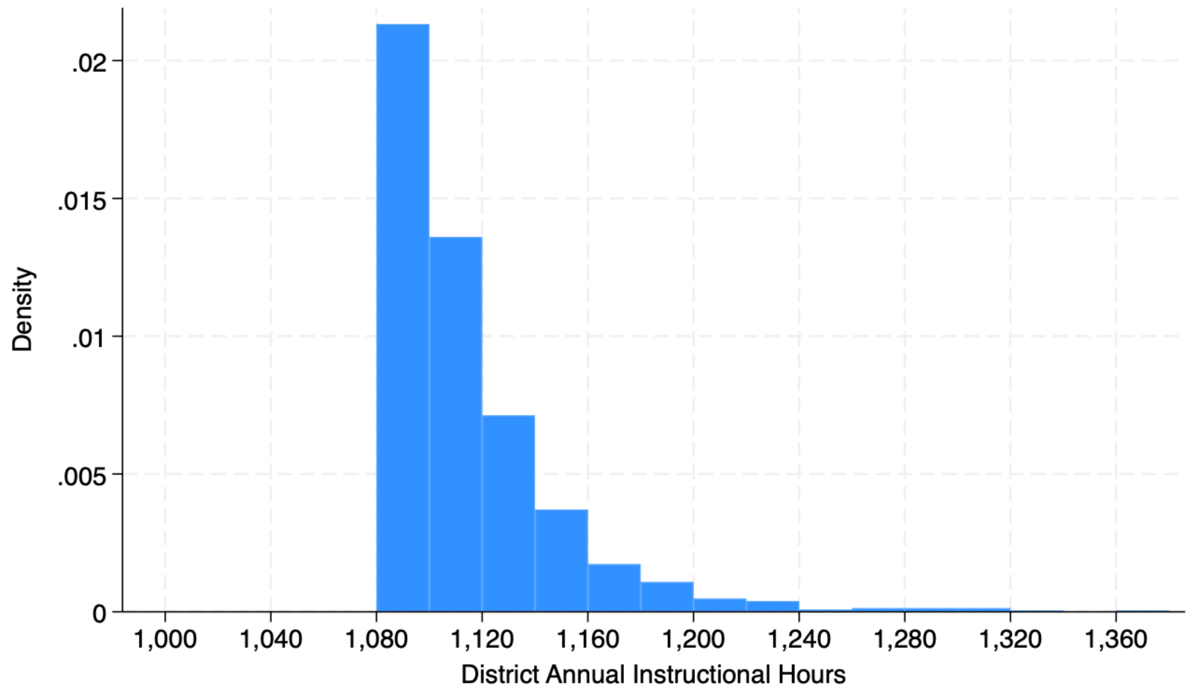

Figure A.4. Histogram of annual instructional hours.

Notes: This figure shows the distribution of school districts' annual instructional hours. In Colorado, public schools are required to provide a minimum of 1,080 instructional hours per year.