When Less Is More: School Schedule Reform and Student Outcomes

Meng Song* Ruinan Zhao†

Working Draft - Please Do Not Circulate

Abstract

We examine how school schedule reforms shape students' outcomes by studying the four-day school week, in which students attend school four rather than five days per week. We use comprehensive administrative data from Colorado, where this schedule is most prevalent in the United States. The dataset spans nearly two decades of entry cohorts and more than one million unique student records, providing us with a unique opportunity to study the policy's long-run effects. Leveraging quasi-random variation in adoption across districts in a difference-in-differences framework, we estimate how a shortened school week affects students' educational trajectories. In contrast to much of the existing literature, we find modest improvements in standardized test scores, lower dropout rates, and higher on-time and overall graduation rates. The duration of exposure plays a key role: students with longer and more consistent exposure experience the largest gains in persistence and attainment. Improvements in teacher retention and classroom expenditures also contribute to these gains. At the same time, heterogeneous effects indicate that disadvantaged students benefit less, raising concerns about equity. Overall, the findings suggest that a shorter school week need not necessarily harm student outcomes and can improve both performance and completion when designed and implemented carefully.

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.

This paper uses confidential administrative student and teacher records from the Colorado Department of Education. The study was approved by the Institutional Review Board at the University of California, Riverside. The findings and conclusions are those of the authors and do not necessarily reflect the views of the Colorado Department of Education.

1. Introduction

Across the United States, the four day school week (4DSW) has emerged as one of the fastest growing educational reforms, reshaping classroom routines, family schedules, and teacher labor markets. Since 1999, the number of districts adopting this schedule has increased sixfold, reaching more than 850 districts across 24 states (National Conference of State Legislatures, 2023). Under this model, students attend school four days per week rather than five, with longer instructional days intended to preserve total instructional time. Proponents emphasize potential benefits such as cost savings, improved teacher morale, and greater flexibility for families, while critics argue that reducing the school week may harm learning and exacerbate inequality. The rapid diffusion of the 4DSW therefore raises a central policy question: how does this schedule affect student achievement, persistence, and completion, and under what conditions can students adjust without losses in learning?

Existing research on the 4DSW has focused primarily on elementary school students, with mixed findings. Several studies document declines in mathematics achievement following adoption, while effects on reading are less consistent (Thompson, 2021; Morton, 2023). Other work reports modest improvements in attendance but limited cost savings, largely driven by reductions in transportation and utilities expenditures (Ward, 2019; Morton, 2021). Fischer and Argyle (2018) further show that the 4DSW can increase juvenile property crime, particularly in urban and low income settings. Evidence on high school students remains scarce and is largely confined to two studies. Using data from Oregon, Thompson et al. (2022) find declines in test scores, lower on time graduation, and higher chronic absenteeism, while Morton (2023) reports modest reductions in disciplinary incidents in Oklahoma but no effects on ACT scores or attendance.

A common feature of much of this literature is its reliance on outcomes that can be measured shortly after policy adoption. Elementary test scores are observed frequently and provide timely indicators of immediate academic responses, making them well suited for evaluating short-run effects. However, these outcomes are less informative about cumulative adjustment and

longer-run attainment, which unfold over multiple years and depend on sustained exposure to the policy. In many existing settings, 4DSW adoption was relatively recent or short lived, further constraining the ability to observe how students, teachers, and schools adapt once transitional dynamics have dissipated. As a result, prior studies provide limited evidence on whether short-run responses persist or translate into meaningful changes in educational trajectories.

We therefore focus on high school students, for whom both the accumulation of policy exposure and the stakes of educational outcomes are especially pronounced. High school is the stage at which changes in school schedules are most likely to translate into durable and economically meaningful outcomes. Unlike elementary and middle school test scores, which are often transitory and subject to later remediation, high school outcomes reflect discrete and frequently irreversible margins of educational attainment, including grade progression, dropout, and diploma completion. These margins directly shape access to postsecondary education, vocational training, and early career opportunities and are central to long run human capital formation. A large literature shows that high school completion strongly predicts lifetime earnings, labor market stability, health, criminal involvement, and intergenerational mobility, while non completion imposes persistent economic and social costs (Heckman, 2006; Lochner and Moretti, 2004; Oreopoulos and Salvanes, 2011; Chetty et al., 2017). Furthermore, high school is also the stage at which behavioral and institutional responses to schedule design are most likely to emerge. Adolescents face meaningful opportunity costs of time, including part time work, caregiving responsibilities, and longer commutes, making changes in the structure of the school week more salient than in earlier grades. At the same time, high school instruction relies more heavily on subject specialization, course sequencing, and teacher continuity, rendering teacher labor supply, burnout, and retention particularly consequential. As a result, the effects of a shortened school week are more likely to operate through persistence, engagement, and completion than through short run test score responses alone. Evaluating high school outcomes therefore provides a clearer assessment of whether schedule reforms generate lasting gains in human capital or merely short run academic fluctuations that may fade over time.

This paper makes four contributions to the literature on school scheduling, education policy, and human capital formation. First, we utilize comprehensive administrative data from Colorado that link the full population of public school students and teachers statewide from 2004 to 2023, comprising more than one million unique student records. To our knowledge, this is the first study to use Colorado's statewide longitudinal student data to examine the effects of the 4DSW, in a setting with the highest concentration and longest history of adoption in the United States. The combination of population-level coverage, cohort-level linkage from school entry through completion, and substantial variation in adoption timing allows us to study high school outcomes in a way that is not feasible in most prior settings.

Second, we provide new evidence that helps reconcile conflicting findings in the existing literature by examining the 4DSW in a distinct institutional and exposure setting. Prior studies from Oregon and Oklahoma document null or negative effects on some student outcomes, highlighting potential concerns related to instructional time compression, adjustment costs, and implementation challenges. These studies are conducted in contexts where the structure of the four-day schedule, the extent of instructional-hour preservation, and the timing of student exposure differ meaningfully across states. In contrast, Colorado's long history of adoption, binding instructional-hour requirements, and cohort-level exposure patterns allow us to study students who experience the 4DSW in a more stable and sustained environment, often beginning before entry into high school and continuing through completion. In this setting, districts adopting the 4DSW maintain instructional time, and we find modest but statistically significant improvements in English and mathematics achievement, a decline in dropout rates of approximately two percentage points, and increases in both on-time and overall graduation. Together, these results suggest that differences in institutional design and exposure structure play a central role in shaping the effects of the 4DSW, helping to explain why impacts vary across states.

Third, we document substantial heterogeneity in treatment effects. Students with longer and more consistent exposure prior to high school experience the largest gains in persistence and attainment, consistent with a cumulative human capital accumulation process and grad-

ual adaptation to the schedule. Although most students benefit, some disadvantaged groups, including English learners and homeless students, experience smaller gains, indicating that equity considerations remain central to the evaluation of this reform.

Fourth, we examine mechanisms that help explain these effects. Analyses of implementation practices, teacher mobility, and district spending point to two primary channels: increased teacher retention and greater classroom investment, both of which contribute to more stable instructional environments. We further show that outcomes are similar across districts that take Monday versus Friday off, but that attainment gains are more persistent among Monday off adopters, suggesting that the organization of instructional time may interact with broader institutional responses to the reform. Together, these findings help reconcile conflicting evidence across states and clarify the conditions under which the four day school week can support student success.

Overall, this paper provides policy relevant evidence on one of the most rapidly expanding school reforms in the United States. By documenting both average effects and heterogeneity across student groups and implementation contexts, the study offers guidance for states and districts considering adoption of the 4DSW.

The remainder of the paper proceeds as follows. Section 2 describes the institutional background of the 4DSW nationally and in Colorado. Section 3 presents a conceptual framework. Section 4 outlines the data and empirical strategy. Section 5 reports the main results and robustness checks. Section 6 investigates 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 per week rather than the traditional five. To satisfy state instructional hour requirements, districts typically lengthen each instructional day by 60 to 90 minutes (CDE, 2024). The model was first introduced during the Great Depression as a cost saving measure and reemerged during the energy crisis of the 1970s, when districts sought to reduce fuel and transportation expenses (Donis

Keller and Silvernail, 2009; Tharp, 2014). Since then, adoption has been concentrated primarily in rural areas, where districts cite not only fiscal considerations but also advantages for teacher recruitment and work life balance (Tharp, 2014; Walker, 2020). Interest rose again following the 2008 financial crisis, when many states faced substantial budget shortfalls (Anderson and Walker, 2015; Walker, 2020).

More than 1,600 districts nationwide now operate on a four day schedule, with Colorado, New Mexico, Montana, and Oregon accounting for the largest shares of adoption (Thompson, 2021; CDE, 2024). Proponents argue that the 4DSW improves teacher morale, reduces absenteeism, and increases time available for lesson preparation and professional development (Anderson and Walker, 2015; Thompson, 2021). Critics emphasize potential costs, including declines in academic achievement, increased childcare burdens for working families, and heightened inequities for low income students who rely on school based services such as meals and structured supervision (Tharp, 2014; Anderson and Walker, 2015). Consistent with these concerns, the existing empirical evidence is mixed and highly context specific, with effects varying by grade level, subject, and state policy environment (Thompson, 2021; Morton, 2022).

Colorado represents the most extensive adopter of the 4DSW in both absolute terms and the share of districts participating. The state legislature formally authorized alternative school calendars in 1980. Although initial adoption was limited, the Colorado Department of Education (CDE) gradually formalized the approval process, requiring districts operating fewer than 160 instructional days to obtain state approval (CDE, 2024). Adoption accelerated in the early 2000s, and by the 2023–2024 school year, 132 of the state's 187 districts, or more than 70 percent, operated on a four day schedule, more than tripling adoption over two decades (CDE, 2024). Adoption is concentrated in rural areas: more than 80 percent of students in rural districts attend schools on a four day calendar, compared with approximately 11 percent in non rural districts. Figure 1 illustrates the geographic distribution of adoption across the state.

The operational structure of Colorado's 4DSW differs systematically from that of traditional five day districts. Standard calendars typically consist of 160 instructional days with approxi-

mately six hours per day. In contrast, four day districts average roughly 144 instructional days with approximately 7.5 hours per day, most commonly operating 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 academic year, distinguishing Colorado from other states in which adoption has coincided with reductions in total instructional time.

Flexibility on the non instructional day is a defining feature of the policy. Teachers frequently use the fifth day for lesson planning, grading, and professional development (Walker, 2020). Students may participate in enrichment activities, extracurricular programs, community service, or family responsibilities. Some districts also offer academic supports such as tutoring, remedial instruction, or supervised childcare to offset potential learning losses, although the availability and quality of these services vary substantially across districts (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 institutional background and implementation of the four day school week (4DSW), this section outlines a conceptual framework that clarifies the channels through which the policy may affect high school outcomes. We adopt a production function and potential outcomes perspective to organize the mechanisms of interest and to guide interpretation of the empirical results.

Let Y_{ist} denote an outcome for student i in school s and year t. The potential outcome under school schedule $D_{st} \in \{0, 1\}$, where 0 denotes a traditional five day schedule and 1 denotes the

4DSW, is given by

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

where T_{st} represents total instructional time, ϕ_{st} captures instructional productivity, Q_{st} denotes teacher quality and effectiveness, and C_{ist} reflects student inputs outside of school, including how the non instructional day is used. X_{ist} denotes student characteristics, Z_{st} captures school and district context such as rurality and budget constraints, and ε_{ist} is an idiosyncratic shock.

Instructional time can be expressed as $T_{st} = d_{st}h_{st}$, where d_{st} is the number of instructional days per week and h_{st} is the length of each school day. Adoption of the 4DSW reduces d_{st} from five to four, typically accompanied by longer instructional days. Some districts maintain approximately constant total instructional time, while others experience increases or decreases in weekly hours. Learning per instructional hour depends on both instructional productivity and day length, denoted $g(\phi_{st}, h_{st})$, with diminishing returns to longer days. Weekly in school learning can therefore be written as

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

Other inputs may adjust in response to the policy. Teacher quality may improve if a shorter week reduces burnout or improves recruitment and retention, but may decline if compressed schedules increase instructional strain:

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

Similarly, student out of school inputs may rise if the additional day is used for rest, enrichment,

or academic support, or fall if students face increased work or caregiving responsibilities:

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

Because adoption of the 4DSW is concentrated in rural districts, contextual factors Z_{st} may also differ systematically between adopting and non adopting schools. These differences may interact with the policy through labor markets, access to childcare, and local capacity to provide fifth day supports.

The average treatment effect of the 4DSW can be expressed as the sum of changes across these channels:

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

The sign of τ is theoretically ambiguous. Total instructional time may remain constant or change, instructional productivity may improve or deteriorate with longer days, teacher quality may respond positively or negatively to labor market incentives, and student inputs outside of school vary with household constraints. These offsetting forces imply no clear theoretical prediction for the overall effect of the 4DSW. As a result, empirical analysis is essential for determining whether, in practice, the four day school week enhances or hinders student achievement, persistence, and attainment.

4. Data and Method

4.1 Data

Our analysis draws on comprehensive administrative records from the Colorado Department of Education (CDE), covering the universe of public school students and teachers from the 2003–2004 through the 2022–2023 academic years. The student-level data include detailed demographic characteristics (race, gender), academic attributes (gifted status, English learner sta-

tus, migrant status), and educational outcomes, including standardized test participation and performance, grade repetition, dropout, and graduation. Unique student identifiers allow us to construct longitudinal records spanning elementary through high school, enabling precise measurement of educational trajectories and cumulative exposure to the 4DSW. The teacher-level data provide information on demographics, professional characteristics such as years of experience and subjects taught, and detailed employment histories.

Information on 4DSW adoption is obtained directly from the CDE and includes district identifiers and the year of first implementation. While a small number of districts adopted a four-day schedule prior to 2000, our observation window begins in the 2003–2004 academic year, with the majority of adoptions occurring after the 2008 financial crisis. Although districts operating under reduced-calendar waivers are not formally required to verify implementation, discussions with CDE officials indicate that nearly all waiver recipients transition to a four-day schedule, suggesting limited scope for treatment misclassification.

We supplement these administrative records with school- and district-level characteristics, including enrollment size, student composition, and rural designation. These variables are used to account for systematic differences across schools and districts that may be correlated with both 4DSW adoption and student outcomes.

Although the data include all Colorado public school students in grades 1 through 12, our baseline sample consists of first-time ninth-grade entrants from the 2004–2005 through the 2019–2020 cohorts. The 2004–2005 cohort is the first for which statewide standardized test scores are available, and restricting entry cohorts through 2019–2020 ensures that each cohort can be followed through the end of high school. This design allows us to observe complete high school outcomes, such as dropout, on-time graduation, and overall graduation.

To ensure a well-defined treatment, we exclude students enrolled in districts that adopted the 4DSW and subsequently returned to a five-day schedule, eliminating cases in which treatment status turns on and off over time. These districts account for fewer than one percent of

the full sample.[1]

Finally, we examine heterogeneity in exposure timing by distinguishing between students first exposed to the 4DSW during high school and those first exposed during middle school.[2]

4.2 Outcomes

We examine student outcomes across three domains: performance, progression, and attainment. Performance is measured using standardized English and mathematics test scores from statewide high school assessments. Prior to the 2016–2017 academic year, Colorado administered the ACT to all high school students; beginning in 2016–2017, the state transitioned to the SAT. To ensure comparability across test formats and cohorts, all scores are standardized within administration year across all districts to have mean zero and unit standard deviation.

Progression outcomes capture students' movement through high school and include grade repetition and dropout. Grade repetition is defined as whether a student repeats any grade during high school. Dropout is defined using CDE records indicating whether a student exits the public school system prior to earning a high school diploma. Attainment outcomes measure high school completion and include on-time graduation, defined as completing high school within four years of entering ninth grade, and overall graduation, which additionally includes students who complete in more than four years.

We also examine teacher outcomes, which serve primarily to illuminate potential mechanisms underlying student effects.[3] These outcomes include (1) retention, defined as remaining in the same school from one year to the next; (2) transfer, defined as leaving the current school to take a position in another Colorado public school; and (3) exit, defined as leaving the Col-

[1] Ninth grade marks the transition into high school, a stage at which families typically have limited information about school-specific rigor, teacher quality, or schedule design. As a result, initial enrollment decisions are less likely to reflect detailed knowledge of school effectiveness, reducing concerns about endogenous sorting based on the four-day schedule. Consistent student identifiers allow us to link records across elementary, middle, and high school, enabling measurement of cumulative exposure to the 4DSW. Because four-day schedules are disproportionately adopted in rural districts, robustness checks implement propensity score matching and a random forest classifier to restrict the comparison group to schools that are observationally similar to treated schools.

[2] Students first exposed during elementary school are excluded because evaluating their high school outcomes requires a substantially longer observation window, yielding a smaller and more selective sample.

[3] Teacher outcomes are used for mechanism analysis and are discussed in detail in Section 6.

orado public school system entirely. These categories are mutually exclusive. The teacher data further include detailed characteristics such as gender, race, grade level, subjects taught, degree attainment, and years of experience.

Finally, following prior work on Oregon's four-day school week (Thompson, 2021), we examine whether changes in total instructional time help explain differences in student outcomes across settings. In Oregon, declines in English and mathematics performance were attributed largely to reductions in instructional hours. By contrast, Colorado imposes binding minimum instructional-hour requirements on districts adopting the 4DSW, potentially accounting for the more favorable outcomes we observe. We therefore test directly whether adoption of the 4DSW alters total instructional hours in Colorado. In addition, we analyze school district expenditures to assess whether a shorter school week affects the allocation of resources across instructional and non-instructional uses. Together, these measures allow us to evaluate whether Colorado's experience reflects distinct implementation choices and resource patterns that help reconcile our findings with prior evidence.

4.3 Summary Statistics

Table I presents summary statistics for 4DSW districts before and after adoption, alongside traditional five-day districts. These comparisons provide descriptive context for the empirical analysis and highlight baseline differences between adopting and non-adopting districts.

Panel A reports student composition and academic outcomes. Overall, the demographic profiles of students in 4DSW and five-day districts are broadly comparable, though modest differences remain. The share of male students is nearly identical across groups at approximately 51 percent. Racial and ethnic composition differs slightly: 4DSW districts serve a smaller share of White and Asian students and a larger share of Hispanic students, with somewhat lower Black representation relative to five-day districts. Other indicators, including the shares of English learners, migrant students, gifted students, and homeless students, are closely aligned across the two groups.

Academic outcomes suggest that students in 4DSW districts begin from a modestly weaker baseline. Average English and mathematics test scores are approximately 1.5 points lower than those observed in five-day districts. Graduation outcomes also lag slightly prior to adoption, with lower on-time and overall graduation rates and marginally higher dropout rates. These descriptive gaps are consistent with the fact that 4DSW adoption is more common among smaller and more rural districts, which historically face greater resource constraints.

Panel B reports teacher characteristics and labor market outcomes. Teacher demographics are similar across 4DSW and five-day districts, with only minor differences in gender and racial composition. Differences are more pronounced in educational attainment: teachers in 4DSW districts are more likely to hold a bachelor's degree and less likely to hold a master's degree than their counterparts in five-day districts. Teachers in 4DSW districts also have slightly more experience on average, with approximately 9.7 years compared with 9.1 years in five-day districts.

Teacher labor market outcomes are likewise broadly comparable across the two groups. Retention rates are modestly higher in 4DSW districts, at approximately 76 percent, while exit rates are lower at around 18 percent. Transfer rates differ little between 4DSW and five-day districts. These patterns are suggestive of greater teacher stability in 4DSW districts, a channel we examine more formally in the mechanisms analysis.

Taken together, the descriptive statistics indicate that 4DSW districts serve student populations that are similar to those in five-day districts along many observable dimensions, but begin from modestly weaker academic baselines. These patterns are consistent with the concentration of 4DSW adoption in smaller and more rural districts. Importantly, the presence of baseline differences underscores the need to control for compositional characteristics and to rely on within-district variation over time in the empirical analysis, rather than cross-sectional comparisons alone.

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 out-

[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.

comes, 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 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:

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

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.

[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:

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

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

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]

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:

$$Y_{dt} = \alpha_1 \textbf{year_before_4DSW} + \sum_{k=1}^{3} \gamma_k (\textbf{year_after_4DSW} = k) + \beta_2 X_{dt} + \theta_d + \lambda_t + \varepsilon_{dt}, \quad (5)$$

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 coef-

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

ficient 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 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

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

[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.

does not induce selective migration or attrition, supporting the credibility of the identification strategy.

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 specifications include school and ninth-grade entry cohort fixed effects, along with a rich set of student, school, and district controls.

We begin with achievement outcomes. We find no evidence that the 4DSW affects standardized test participation, with estimates that are small and statistically indistinguishable from zero. Test performance, however, improves modestly under full exposure to the policy. English scores increase by 0.027 standard deviations, significant at the 5 percent level, and mathematics scores by 0.025 standard deviations, significant at the 10 percent level. Although modest in absolute terms, these effects are economically meaningful. Using the common benchmark that a 0.25 standard deviation gain corresponds to approximately one grade-level year of learning, these estimates imply roughly one to one-and-a-half months of additional progress (Bloom et al., 2008). The magnitudes are comparable to those reported for other widely studied school-based reforms, such as accountability policies under No Child Left Behind (Hanushek and Raymond, 2005; Dee and Jacob, 2011). By contrast, estimates for students first exposed to the 4DSW after ninth grade are smaller and imprecisely estimated, indicating that achievement gains are concentrated among cohorts with sustained exposure.

Turning to progression outcomes, we find no detectable effect on grade repetition. Dropout, however, declines sharply under full exposure to the 4DSW. The estimated reduction of 2.1 percentage points corresponds to an 18 percent decrease relative to the control mean dropout rate of 11.8 percent. This effect is both statistically and substantively important, particularly in Colorado, where students are legally permitted to leave school at age 17. The result suggests that the 4DSW plays a meaningful role in retaining marginal students during the period when early

exit becomes feasible. Partial exposure yields smaller and statistically insignificant effects, reinforcing the importance of sustained implementation.

The largest effects emerge in attainment outcomes. Under full exposure, overall graduation increases by 3.2 percentage points and on-time graduation by 2.9 percentage points, both statistically significant at the 1 percent level. These estimates correspond to increases of approximately 4.3 percent and 4.0 percent relative to baseline graduation rates of 74.7 and 73.2 percent, respectively. Students first exposed to the 4DSW after ninth grade also experience gains, with graduation and on-time graduation rising by 1.3 and 1.4 percentage points. Although smaller in magnitude, these effects indicate that the policy improves completion even for partially exposed cohorts, with the largest and most consistent benefits accruing to students exposed continuously from ninth grade onward.

To place these magnitudes in context, we benchmark our estimates against findings from related work. Jackson et al. (2020) show that attending a school one standard deviation higher in predicted test-score value added increases high school graduation by approximately 1.3 percent for ninth-grade students in Chicago public schools. By comparison, the 4.3 percent increase in graduation associated with full exposure to the 4DSW in Colorado is equivalent to attending a school with roughly 3.3 standard deviations higher test-score value added. Scaling these effects to the population level further underscores their importance. With approximately 68,000 students per annual cohort in Colorado (CDE, 2025), full exposure to the 4DSW would prevent an estimated 1,428 dropouts, generate 2,176 additional graduates, and yield 1,972 more on-time graduates. Even modest percentage changes at the individual level therefore translate into substantial aggregate gains in human capital, consistent with evidence on the high returns to high school completion (Lochner and Moretti, 2004; Oreopoulos and Salvanes, 2011) and the large long-term costs of dropping out (Heckman, 2006).

Finally, our findings differ sharply from prior evidence in Oregon. Thompson et al. (2021) document declines of approximately 0.09 standard deviations in test scores and reductions of four to five percentage points in graduation rates, with no corresponding change in dropout.

In Colorado, by contrast, both achievement and attainment improve. This contrast highlights that the effects of the 4DSW depend critically on how the policy is implemented rather than on the schedule itself. In Oregon, adoption was accompanied by substantial reductions in instructional hours, directly lowering time spent in school. In Colorado, districts largely maintain required instructional hours and often reallocate savings toward classroom resources and teacher support, effectively ruling out lost instructional time as the primary channel (see Section 6). These institutional differences underscore that school schedules interact with broader organizational and resource decisions. As emphasized by Jackson et al. (2020), schools influence students through multiple pathways: academic, behavioral, and organizational, and the relative importance of these channels varies across contexts. Understanding these mechanisms is therefore central to explaining why the same policy can produce divergent outcomes across states and to identifying the design features that make alternative schedules most effective.

5.2 Duration of Exposure

The results in Table II indicate that students fully exposed to the 4DSW throughout high school experience larger gains in persistence and attainment than those exposed only later, suggesting that the duration of exposure plays an important role in shaping outcomes. This pattern 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 separate effects for cohorts first exposed in middle school and for those whose exposure begins only after entry into high school.[11] Table III summarizes the results.

The evidence reveals a clear pattern of heterogeneous but consistently positive effects. For students first exposed to the 4DSW during high school, the estimates indicate modest improvements in both performance and attainment. Test participation remains unchanged, while English and mathematics scores increase by approximately 0.02 to 0.03 standard deviations. Dropout declines slightly but is not statistically distinguishable from zero. Graduation and on-time grad-

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

uation, however, rise by roughly 1.6 to 1.7 percentage points. These results suggest that even limited exposure to the 4DSW can improve educational outcomes during high school, though the magnitude of the effects remains relatively modest.

By contrast, students first exposed to the 4DSW in middle school experience larger and more precisely estimated gains. English and mathematics scores increase by about 0.02 standard deviations, while dropout falls by approximately 1.9 percentage points. Graduation and on-time graduation rates rise by more than 4 percentage points, with both effects statistically significant at conventional levels. Relative to students first exposed only in high school, earlier exposure is associated with substantially larger improvements in persistence and completion, consistent with a cumulative effect of sustained exposure to the schedule.

Panel C provides additional evidence of a dose–response relationship between exposure length and educational outcomes. Each additional year of exposure to the 4DSW is associated with a 0.6 to 0.7 percentage point increase in graduation and on-time graduation rates and a 0.4 percentage point reduction in dropout, all statistically significant. These estimates imply that the benefits of the 4DSW accrue gradually over time rather than emerging immediately upon adoption. Longer exposure allows small annual gains to compound into meaningful improvements in high school completion.

Taken together, these findings highlight that both the timing and duration of exposure are central to understanding the effects of the 4DSW. Earlier adoption, beginning before high school, appears most conducive to sustained improvements in attainment. Adolescence is a formative period during which students develop learning routines, self-regulation, and time-management skills that strongly influence later educational outcomes (Cunha and Heckman, 2007). Prolonged exposure to a consistent school schedule may facilitate these behavioral adjustments and allow them to reinforce one another over time. In addition, longer implementation horizons provide schools and districts with greater scope to adapt instructional practices, staffing, and organizational routines to the four-day structure. Districts that adopt earlier are more likely to complement the schedule change with investments in planning, professional development,

and curriculum coordination, which may enhance the effectiveness of the reform (Jackson, 2018). By contrast, students first encountering the policy during high school may have already established routines that are less responsive to changes in school structure. As a result, policymakers should expect smaller short-run effects for partially exposed cohorts and consider transitional supports, such as tutoring or coordination with families, to facilitate adjustment. Overall, the evidence suggests that sustained and consistently supported implementation, rather than short-term or partial exposure, is critical for realizing the longer-run educational benefits of the 4DSW.

5.3 Do All Students Benefit Equally?

Table IV examines heterogeneity in the effects of the 4DSW across student subgroups. The analysis assesses whether improvements in achievement, progression, and attainment are broadly shared or disproportionately accrue to particular populations.

By gender, the estimated effects are broadly similar across most outcomes. Both male and female students experience statistically significant improvements in progression and attainment, including reductions in dropout and increases in graduation and on-time graduation. Gains in test performance are somewhat larger and more precisely estimated for male students, particularly in mathematics, though the direction of the effects is positive for both groups. Overall, the results indicate that the benefits of the 4DSW are not meaningfully differentiated by gender, with only modest differences in test score responses.

Across racial and ethnic groups, graduation outcomes improve for all major populations, though the magnitude and precision of the estimates vary. White and Hispanic students exhibit consistent and statistically significant gains in both achievement and attainment. Estimated effects for Black students are also positive but less precisely estimated, reflecting smaller sample sizes and greater underlying variation in outcomes. Taken together, these results suggest that the positive effects of the 4DSW are not confined to a single racial or ethnic group but extend across the primary populations served by Colorado public schools.

More pronounced differences arise for subgroups defined by socioeconomic and educational disadvantage. English learners and homeless students exhibit smaller and less consistent gains, and in some cases negative point estimates, particularly for achievement outcomes. Because these groups constitute a relatively small share of the sample, the estimates are imprecise and should be interpreted cautiously. Even so, the pattern suggests that students who face greater barriers to learning or rely more heavily on school based instructional and support services may benefit less from a shortened school week. For these students, reduced in school structure or limited access to academic supports on the non instructional day may offset some of the gains observed for other groups. These results point to the potential value of complementary supports, such as targeted tutoring, language services, or supervised programming, in shaping how schedule reforms affect vulnerable students.

5.4 Robustness Checks

To assess the credibility of the baseline estimates, we conduct a comprehensive set of robustness and complementary analyses designed to address concerns related to covariate balance, model specification, treatment definition, sample composition, measurement, and inference.

We begin by addressing potential imbalance between treated and control units. Adoption of the 4DSW is concentrated in smaller and more rural districts, which implies that treated students represent a relatively small share of the statewide population of non-adopting students, even though adoption spans a large fraction of districts. To improve comparability, we implement propensity-score matching based on pre-treatment school characteristics that predict adoption. Each treated school is matched to its nearest control schools in terms of estimated propensity scores, without replacement, and the least comparable control schools are excluded.[12] The resulting estimates, reported in Table A.5 Panel A, closely mirror the baseline results, indicating that differences in observable characteristics are unlikely to drive the main

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

findings.

To further assess balance, we re-estimate adoption propensities using a random forest classifier, which flexibly captures nonlinearities and higher-order interactions among covariates without imposing parametric structure. Applying the same matching procedure based on these nonparametric predictions,[13] the estimates reported in Table A.5 Panel B remain nearly identical to the baseline. As an alternative approach, we also estimate inverse probability weighted models that retain the full sample while reweighting observations so that the control group mirrors the covariate distribution of treated schools. The IPW estimates in Table A.5 Panel C again align closely with the baseline, reinforcing the conclusion that the results are robust to alternative methods for achieving covariate balance.

Next, we examine sensitivity to model specification. Estimating the baseline specification without any control variables yields coefficients that are very similar to those from the fully controlled model, indicating that the results are not driven by covariate adjustment. Adding district-specific linear pre-trends to account for potential differences in underlying outcome trajectories across districts produces nearly identical estimates (Table A.5 Panel E). To assess whether the findings are driven by any single adopting district, we sequentially exclude each treated district and re-estimate the model. As shown in Figure A.3, the estimated effects remain stable across exclusions, suggesting that no individual district disproportionately influences the results.

We also examine sensitivity to the definition of treatment. The baseline specification distinguishes between full exposure and partial exposure for students who begin high school under a five-day schedule but experience adoption while still enrolled. 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 restriction provides a cleaner comparison that does not rely on assumptions about the functional form of partial exposure. The resulting estimates, reported in Table A.6, remain consistent with the baseline findings.

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

To evaluate the influence of sample composition, we conduct several additional checks. Restricting the analysis to students entering high school in 2015 or earlier, and thus graduating prior to the Covid pandemic, yields estimates that closely match the baseline. Excluding charter schools and Boards of Cooperative Educational Services, which operate under distinct governance and funding structures, produces similar results. Although 4DSW adoption in Colorado is predominantly rural, several urban districts also adopt the schedule. Restricting the sample to small rural districts produces estimates that are consistent with the baseline, indicating that the results are not driven by urban or suburban adopters.[14]

We further assess the robustness of achievement results to changes in test format. Colorado administered the ACT prior to the 2016–2017 academic year and transitioned to the SAT thereafter. Although the baseline analysis standardizes scores across assessments, we re-estimate the models excluding all SAT-based cohorts and, separately, excluding all ACT-based cohorts. In both cases, the estimates remain virtually unchanged,[15] indicating that test format transitions do not drive the achievement findings.

Finally, we conduct a placebo inference test by randomly reassigning adoption years across districts 1,000 times and re-estimating the baseline specification. Figure 4 shows that the true estimates lie in the extreme tails of the placebo distribution, with fewer than ten percent of simulated draws exceeding the observed effects in absolute value. This result provides additional reassurance that the estimated impacts are unlikely to arise from chance correlations.

Collectively, these robustness exercises demonstrate that the estimated effects of the 4DSW are stable across a wide range of samples, specifications, treatment definitions, and inference procedures.

6. Potential Mechanisms and Discussion

6.1 Potential Mechanism

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

[15] Results excluding ACT-based cohorts are similar and available upon request.

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

[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.

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 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 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

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

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 resources 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

[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.

[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.

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 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

[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.

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 large-scale causal evidence on the effects of the 4DSW on high school students, drawing on nearly two decades of linked administrative records from Colorado that cover

more than one million unique students across successive cohorts. The scale and longitudinal structure of the data allow us to evaluate both short-run academic performance and longer-run educational attainment, outcomes that are central to human capital formation and economic mobility. We find that adoption of the 4DSW leaves test participation unchanged, modestly improves achievement, and generates economically meaningful increases in persistence and high school completion. When scaled to the size of a typical statewide cohort, these effects imply substantial gains in educational attainment and aggregate human capital.

The results also clarify the conditions under which these benefits are most likely to arise. Students with longer and more consistent exposure to the 4DSW, particularly those exposed prior to entering high school, experience the largest and most reliable gains in persistence and attainment. In contrast, students first exposed during high school exhibit smaller improvements. These patterns indicate that the timing and duration of exposure are critical for understanding the effects of calendar reforms and suggest that short-term or partial adoption may yield limited benefits relative to sustained implementation.

The distribution of gains further highlights important equity considerations. While improvements in persistence and completion are observed for most student groups, the benefits are smaller and less consistent for English learners and homeless students, populations that rely more heavily on school-based structure and support services. These findings indicate that calendar reforms may improve average outcomes while leaving some students behind if not paired with complementary supports. From a policy perspective, this heterogeneity underscores the importance of designing alternative schedules alongside targeted interventions, such as academic supports or supervised programming on non-instructional days, to ensure that flexibility-enhancing reforms do not exacerbate existing disparities.

Evidence on mechanisms helps explain why Colorado's experience differs from prior findings in other states. Colorado's binding instructional-hour requirements ensure that total classroom time is preserved under the 4DSW, ruling out reductions in instructional time as a primary channel. Instead, teacher retention emerges as a central mechanism. Districts adopting the

4DSW experience higher retention, strengthening continuity in instruction and relationships between students and teachers, which appears to contribute to improved persistence and completion. We also document modest increases in instructional supply expenditures, suggesting some reinvestment in classroom environments, though the magnitude of these changes limits strong conclusions about their independent effects. In contrast, we find little evidence of systematic expansion in purchased or community services that would indicate widespread enrichment programming on the non-instructional day. The organization of the schedule itself also matters, as districts taking Monday off exhibit more sustained gains in attainment than those taking Friday off, highlighting that calendar design interacts with broader institutional responses.

Taken together, the findings offer several policy-relevant lessons for states and districts considering adoption or expansion of the 4DSW. First, alternative school calendars can improve high school persistence and completion when implemented under institutional conditions that preserve instructional time and support teacher stability. Second, the benefits of such reforms are more likely to materialize with sustained exposure rather than short-term experimentation. Third, distributional consequences warrant careful attention, as vulnerable students may require additional supports to benefit from increased scheduling flexibility. More broadly, the results underscore that the effects of calendar reforms are not inherent to the schedule itself but depend critically on how the policy is designed, implemented, and integrated into existing educational systems. Evaluations of alternative schedules should therefore consider not only average effects but also timing, duration, and equity implications to inform policies that promote both efficiency and inclusion.

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 grade 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	**Transfer**	**Exit**
	(1)	(2)	(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.
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

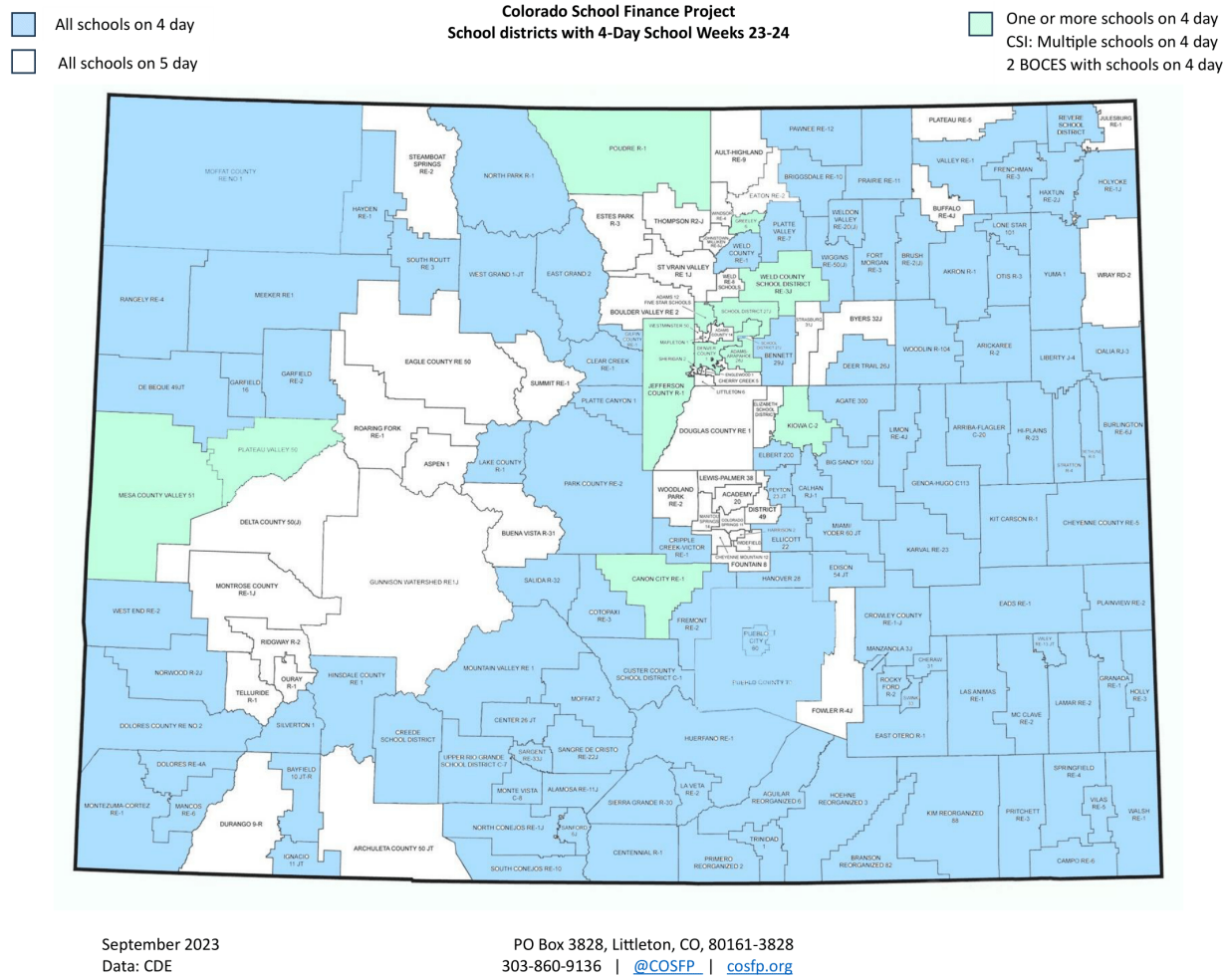


Figure 1. Map of 4DSW Districts in Colorado.

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.

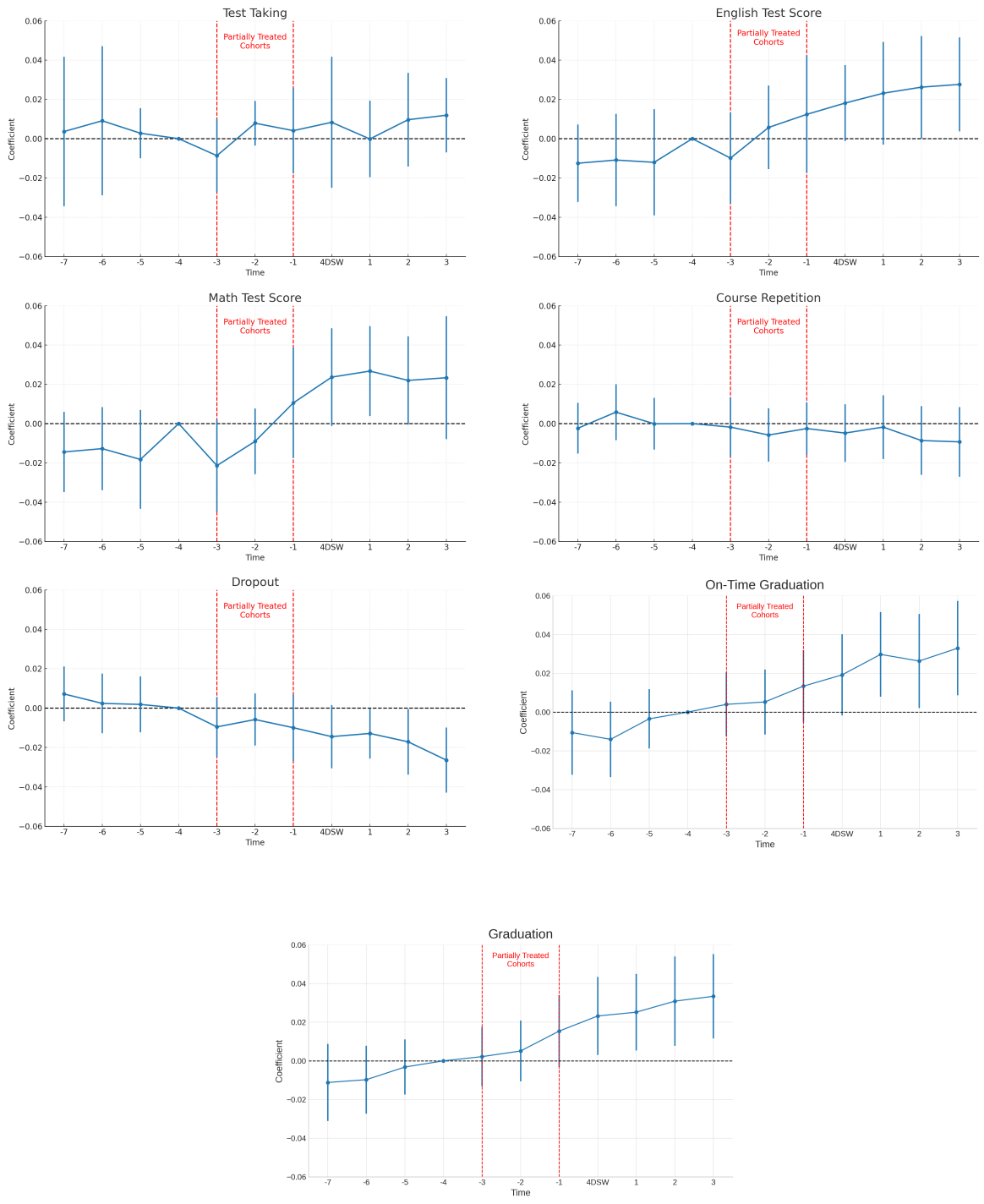


Figure 2. Event study of the effect of 4DSW on student outcomes.

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.

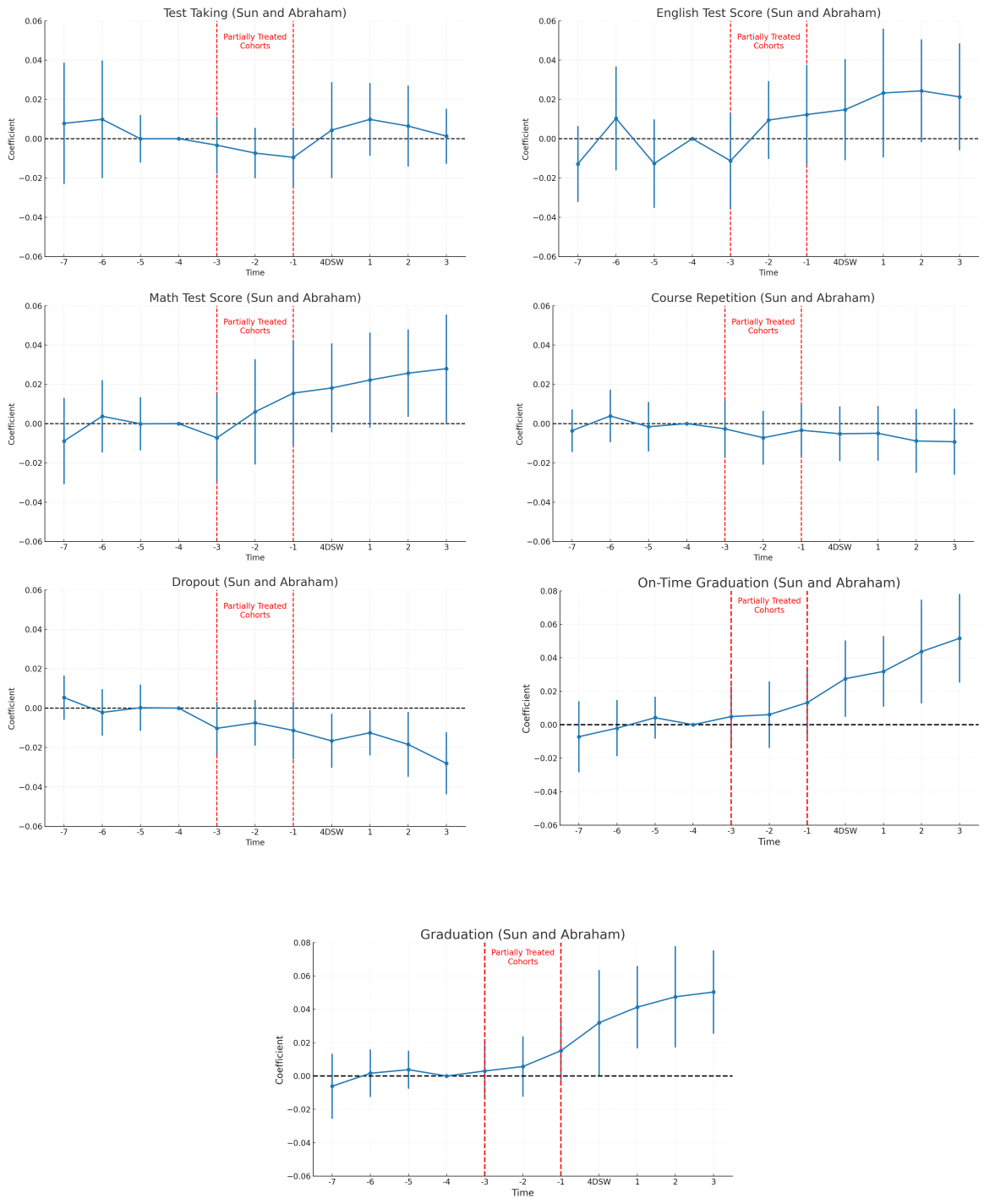


Figure 3. Event study of the effect of 4DSW on student outcomes: Sun and Abraham (2021).

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.

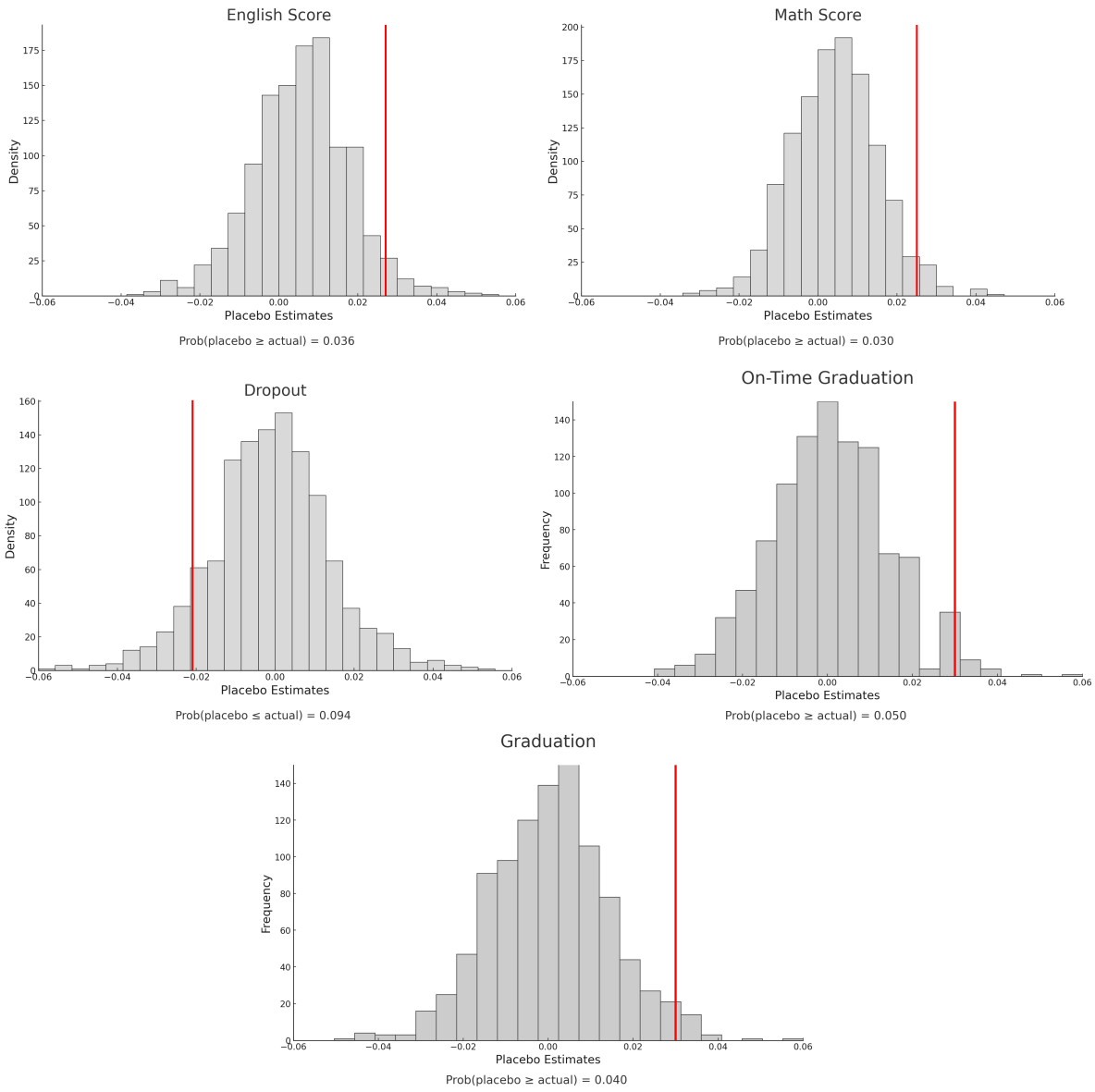


Figure 4. Placebo estimates.

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** (1)	**Teacher Characteristic**	**Teacher Sorting** (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*	0.031**	0.032**
	(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 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.
* $p < 0.10$, $^{*}*$ $p < 0.05$, $^{*}**$ $p < 0.01$

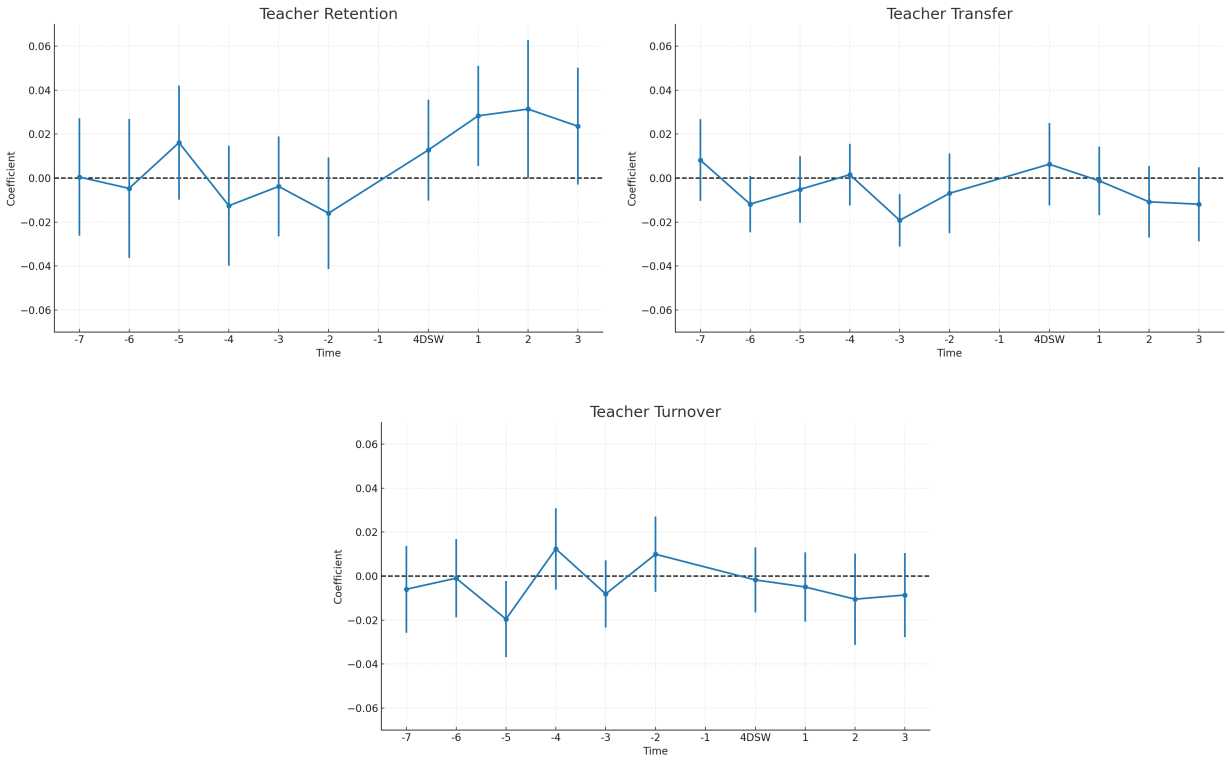


Figure A.1. Event study of the effect of 4DSW on teacher outcomes.

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

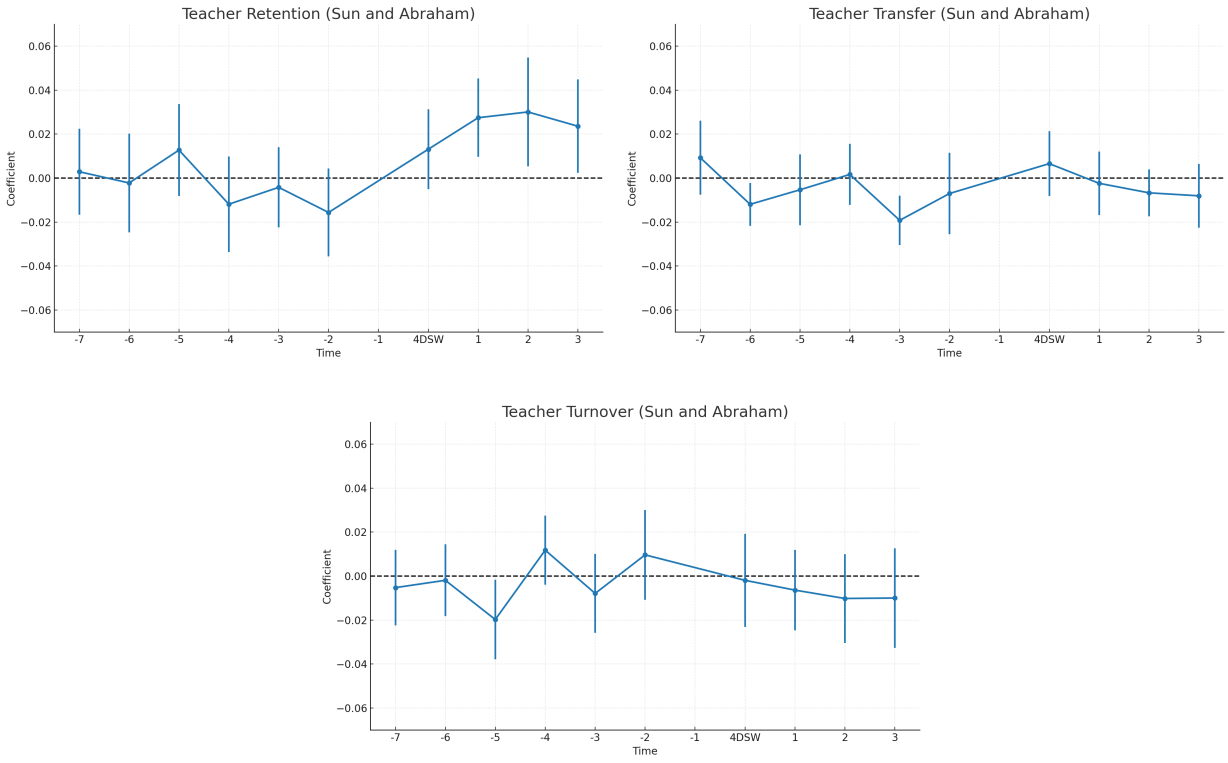


Figure A.2. Event study of the effect of 4DSW on teacher outcomes: Sun and Abraham (2021).

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.

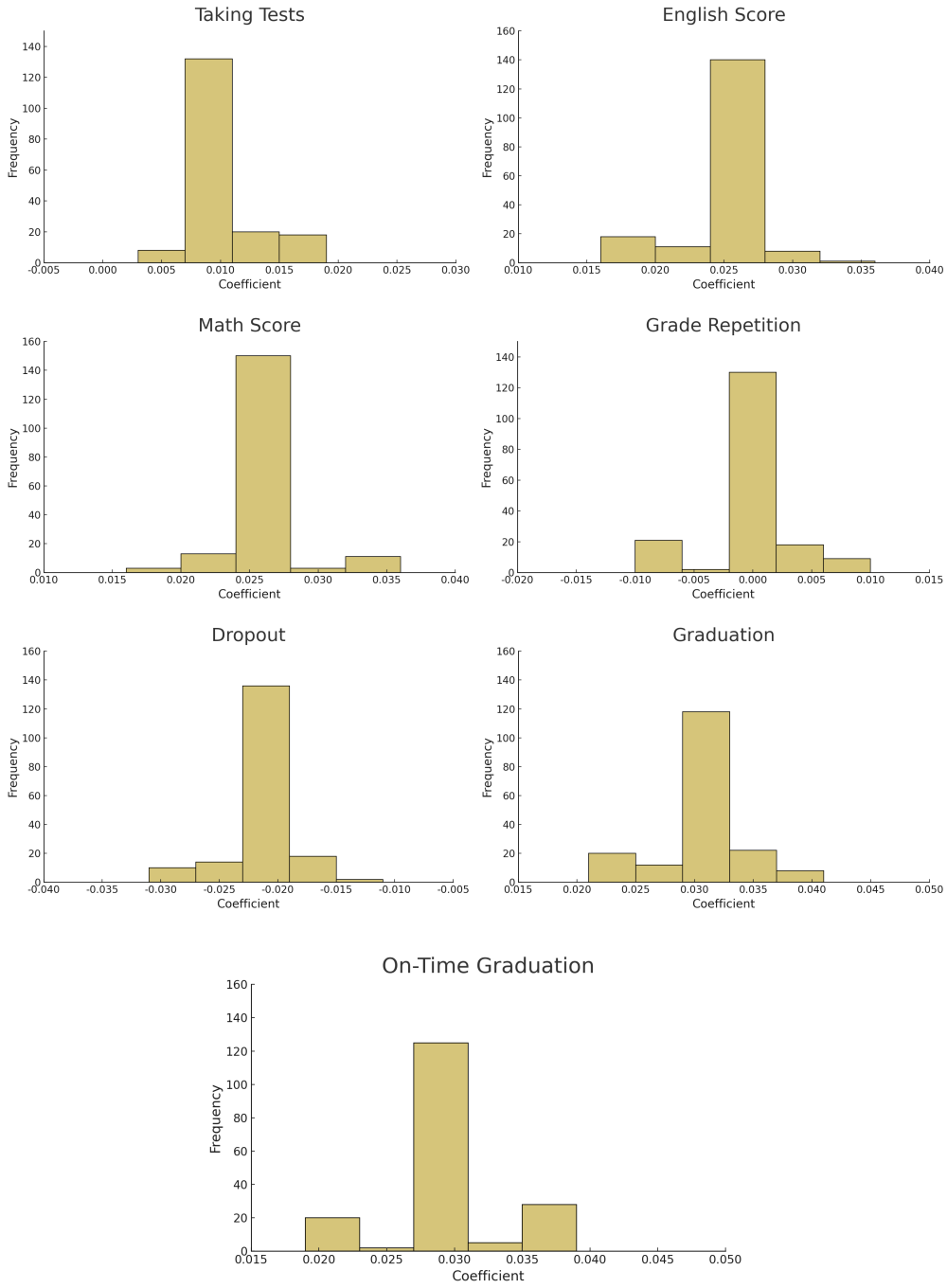


Figure A.3. Leave-one-district out estimations.

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.

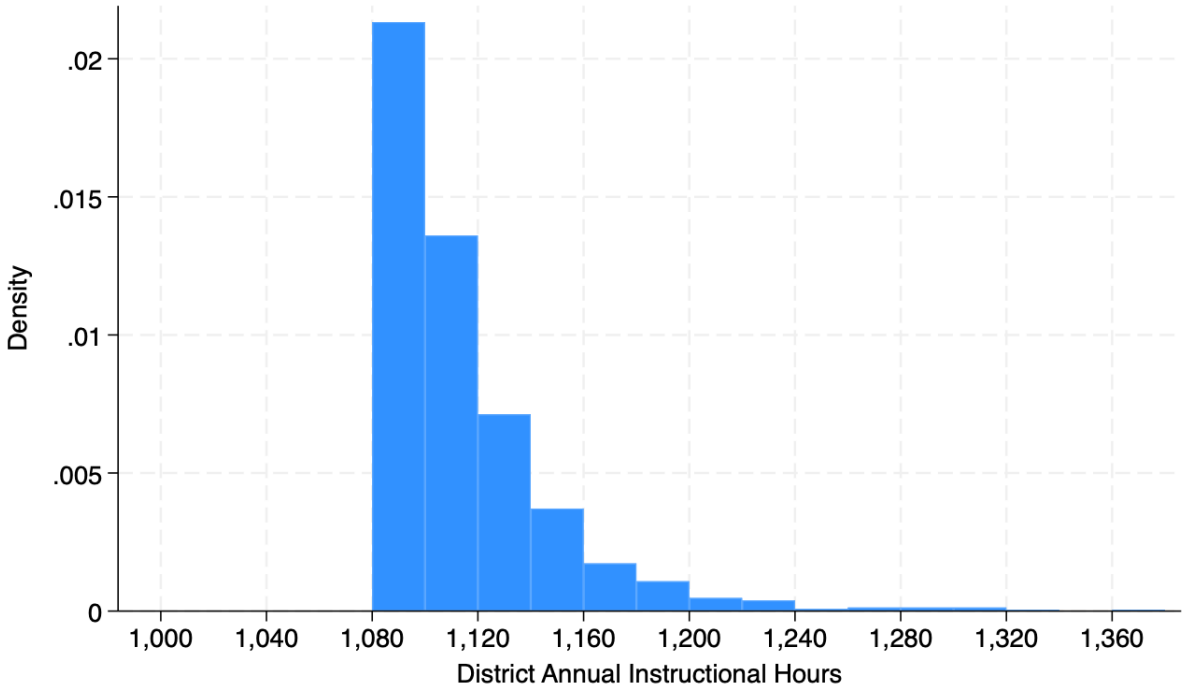


Figure A.4. Histogram of annual instructional hours.

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.