I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14461

Only a Matter of Time? The Role of Time in School on Four-Day School Week Achievement Impacts

Paul N. Thompson
Jason Ward

JUNE 2021

I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 14461

Only a Matter of Time? The Role of Time in School on Four-Day School Week Achievement Impacts

Paul N. Thompson
Oregon State University and IZA

Jason Ward
RAND Corporation

JUNE 2021

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany
Phone: +49-228-3894-0
Email: publications@iza.org
www.iza.org

ABSTRACT

Only a Matter of Time? The Role of Time in School on Four-Day School Week Achievement Impacts*

Previous evidence has shown disparate achievement impacts of the four-day school week across state contexts. This paper examines the impacts of the four-day school week on achievement and achievement gaps across 12 states to contextualize these four-day school week impacts nationally. We estimate these effects using a difference-in-differences design with data from the Stanford Educational Data Archive and a proprietary longitudinal national four-day school week database of four-day school week use from 2009-2018. We find negative impacts for both math and ELA achievement when examining four-day school weeks nationally, but these aggregate effects appear to be masking important heterogeneity due to differences in overall time in school across states and districts. When stratifying the treatment sample into districts with low, middle, and high time in school, we find statistically significant negative impacts on math achievement for four-day school districts with low time in school, but no statistically significant impacts for four-day districts with middle or high time in school. Our findings suggest that maintaining time in school should be a key consideration for school districts contemplating fourday school week adoption.

JEL Classification: I21

Keywords: four-day school week, time in school, achievement

Corresponding author:
Paul N. Thompson
School of Public Policy
340 Bexell Hall
Oregon State University
Corvallis, OR 97331
USA
E-mail: paul.thompson@oregonstate.edu

* The research reported in the article was made possible (in part) by a grant from the Spencer Foundation (#202100126). The views expressed are those of the authors and do not necessarily reflect the views of the Spencer Foundation. This work has also been supported (in part) by Grant # 2001-20348 from the Russell Sage Foundation and the William T. Grant Foundation. Any opinions expressed are those of the principal investigator(s) alone and should not be construed as representing the opinions of either Foundation.

1. Introduction

Brick-and-mortar school environments foster human capital accumulation, social emotional development, and healthy behaviors through daily access to productive academic instruction, school meal programs, physical activity opportunities, and structured face-to-face social interactions with peers, teachers, and administrators. The positive impacts of educational input quality (e.g., high-quality teachers, small class sizes, high-quality school infrastructure; see, for example, Rockoff, 2004; Angrist and Lavy, 1999; Krueger, 1999; Jackson, 2020) and the quantity of exposure to these educational inputs (e.g., instructional time; see, for example, Lavy, 2015; Cattaneo, Oggenfuss, and Wolter, 2017) on student outcomes from early elementary through adolescence and adulthood are well-documented. In light of these benefits, the recent shift away from primarily in-person schooling to hybrid or remote schooling in many school districts across the United States in the wake of the COVID-19 pandemic has led to great concern over the ramifications of this lost exposure to the in-person school environment. This stark shift in the modality of educational provision led to, at a minimum, short-run impacts on academic achievement (Kuhfeld, et al., 2020; Chetty et al., 2020), socioeconomic gaps in student engagement with online learning resources (Barnum and Bryan, 2020; Bacher-Hicks, Goodman, and Mulhern, 2021), child maltreatment reporting (Baron, Goldstein, and Wallace, 2020), obesity (Takaku and Yokoyama, 2020) graduation decisions (Aucejo, et al., 2020) and self-reported well-being (Duckworth, et al. 2021).

While these COVID-related changes in educational provision are certainly unique in their breadth, other types of alternative school schedules also change the composition or quantity of in-person school exposure on a smaller scale. One such schedule, the four-day school week, eliminates one required school day per week — a Friday or Monday – with longer school hours[1] on the remaining four school days to adhere to state-mandated minimum yearly instructional hours requirements. An increasing trend in use of the four-day school week has been observed over the past two decades, with over 1,600 schools in 650 school districts, primarily in rural settings, operating on this schedule prior to the COVID-19 pandemic. Historically this shortened school schedule was primarily motivated by financial considerations,[2] but since the start of the pandemic other school districts have switched to a four-day model for public health reasons (e.g., "cleaning days"; see Lake and Heyward, 2021). Regardless of the rationale for these four-day school week schedules, understanding the ramifications for student achievement are of great importance amongst stakeholders. A sparse, but growing, literature has found mixed evidence on the effects of four-day school week adoption on student achievement across several state-specific studies – finding positive effects in Colorado and negative impacts in Oklahoma and Oregon.[3] Given the documented relationship between instructional time and academic achievement (Lavy, 2015; Cattaneo, Oggenfuss, and Wolter, 2017), these disparate four-day school week achievement impacts may be driven primarily by choices over how much time students spend in school under these four-day school week models (Thompson, 2021). However, other mechanisms such as weekend learning loss and early school start times could also play a role in achievement impacts.

---

[1] The average school day lasts seven hours and 45 minutes under a four-day school week model and six hours and 54 minutes under a five-day school week model (Thompson, et al. 2020).

[2] A recent survey by Thompson, et al. (2020) found nearly two-thirds of districts that adopt these four-day school schedules cite financial reasons; other key reasons they cite include teacher retention, student attendance, and rural-related issues (e.g., commuting time, farming/ranching).

[3] Anderson and Walker (2015) found positive impacts of the four-day school week on 4th grade reading and 5th grade math proficiency in Colorado during the pre-Great Recession period. More recent evidence on the effects of four-day school week adoption during and after the Great Recession found negative impacts in Oklahoma and Oregon. Morton (2020) found insignificant, negative impacts of the four-day school week using SEDA achievement data from Oklahoma. The point estimates, however, suggested a 0.052 standard deviation reduction in math and a 0.032 standard deviation reduction in reading. Using student-level achievement data in Oregon, Thompson (2021) found reductions of between 0.037 and 0.059 standard deviations in math and between 0.033 and 0.042 standard deviations in reading following the switch to the four-day school week.

This study also estimates the causal impact of four-day school week adoption on student achievement, but makes several key contributions to this emerging literature. First, this study provides some of the first multi-state estimates of four-day school week achievement impacts. We use a difference-in-differences research design to examine the achievement impacts of the four-day school week in 12 states using data on district-level student achievement from the Stanford Educational Data Archive (SEDA) and a proprietary, longitudinal, national database of four-day school week use from 2009-2018. This aggregate analysis finds negative effects for both math and ELA achievement – evidence consistent with the recent state-specific findings.

Second, we provide evidence that these aggregate effects may be masking important heterogeneity due to differences in overall time in school across states and districts. Both overall time in school requirements and instructional time vary significantly across the U.S. (Phelps, et al., 2012; Brixey, 2020). Additionally, districts adopting the four-day week vary in the extent to which they extend the four in-person days (Thompson, et al., 2020). When stratifying the treatment sample into terciles of low, middle, and high time in school groups, we find statistically significant negative impacts on math and ELA achievement for the low time in school four-day school week adopters, but no statistically significant impacts for the middle or high time in school four-day school week adopters. This new evidence suggests that the four-day school week is not detrimental for achievement per se, but that four-day school weeks implemented in districts with low overall time in school at baseline are likely to have meaningful negative consequences for student academic progress. Thus, maintaining adequate overall time in school should be a key consideration for school districts thinking about adopting this type of alternative school schedule.

2. Data and Methodology

We construct a school district-level panel data set consisting of school districts from a subset of states that allow four-day school weeks for the 2008-2009 through 2017-2018 school years. To do so, we combine data from the Stanford Educational Data Archive (SEDA; Reardon, et al., 2019) and a proprietary, longitudinal national four-day school week database. We obtained restricted use SEDA data that include estimates[4] of school district-level math and ELA achievement across grade 3 through 8 for a majority of the school districts in the continental United States.[5] The four-day school week database (see Thompson, et al., 2020) includes a panel data set of four-day school week use and a 2018-2019 cross-section of time in school for these four-day school week districts.

The analytic sample consists of all school districts with nonmissing school week, school district, and math and ELA achievement data from the 12 states with the highest prevalence of four-day school week school district coverage within the SEDA dataset.[6] Table 1 presents descriptive statistics for the achievement data and school district demographic and community characteristics for the school districts included in the analytic sample for the 2008-2009 through 2017-2018 school years. As suggested by Reardon, Kalogrides, and Shores (2019), descriptive statistics for mean math and ELA achievement and subsequent regression models are precision-weighted using the inverse of the estimated standard error squared.[7]

---

[4] In a process described by Reardon, Shear, Castellano, and Ho (2017), the achievement data are estimated from state accountability proficiency data. As states generally use different standardized assessments, the achievement data are converted to a common scale using state-level estimates from the National Assessment of Educational Progress (see Reardon, Kalogrides, and Ho, 2017). This two-step process creates test scores that are comparable across states and across time and provides a unique database from which to assess national four-day school week achievement effects.

[5] Given the small sizes of overall student populations by grade among many four-day districts, reporting restrictions in the public-use SEDA data means that the data set only provides meaningful estimates for 59.7 percent of the four-day school week districts we identify in our four-day school week adoption database. We therefore, use restricted-use data for these measures – similar data to that of Morton (2020) in Oklahoma – that are aggregated by school district across grades 3-8 (or 3-7 in certain instances) in order to increase the size of our estimation sample and the number of treated units. Use of these restricted data increase our coverage to 81.6 percent of the districts in our adoption database.

[6] These 12 states are Arizona, Georgia, Idaho, Kansas, Minnesota, Missouri, Montana, New Mexico, Nevada, Oklahoma, Oregon, and South Dakota. Colorado, the state with the highest four-day school week prevalence, and Wyoming, which ranks 11th in terms of number of four-day districts in our database, are both excluded from our analysis due to multiple years of missing data within the SEDA dataset.

[7] This type of precision weighting proportionally weights the observations up or down according to whether the estimated standard errors are less than or greater than the estimated mean. In our analytic sample 57,165 out of the total 91,378 school district-year-grade observations have standard errors greater than or equal to the estimated mean value.

Table 1 notes large differences in achievement across four- and five-day school week districts in the states in our sample. Four-day school week districts have math achievement that is 0.18 standard deviations below the mean and ELA achievement that is 0.15 standard deviations below the mean. In comparison, five-day schools have math and ELA achievement that is at the mean and 0.01 standard deviations below the mean, respectively. Given differences in enrollment, rurality, and the demographic make-up of these school districts, however, it is hard to causally link these achievement differences to the adoption of a four-day school week. To account for these potential confounders, we use a difference-in-differences regression analysis of the effects of four-day school weeks on achievement using the following two-way fixed effects estimator:[8]

$$A_{dgt} = \alpha + \beta FDSW_{dt} + \gamma \boldsymbol{X}_{dgt} + \mu_d + \theta_{gt} + \epsilon_{dgt} \quad (1)$$

where A_{dgt} is the school district d's math or English/language arts (ELA) achievement for grade g during school year t. The $FDSW_{dt}$ variable is a dummy variable equal to one if the school district was operating on a four-day school week schedule during school year t,[9] and the coefficient, β, is the estimate of interest. Vector $\boldsymbol{X}_{dgt}$ controls for time-varying district-level confounders (e.g., unemployment rate, average adult education level, racial/ethnic composition), while district fixed effects, μ_d, control for permanent differences across districts and grade-by-year fixed effects, θ_{gt},

[8] Given recent literature on concerns over the use of two-way fixed effects estimators when there is staggered treatment adoption (Goodman-Bacon, 2018; de Chaisemartin and D'Haultfoeuille, 2020; Callaway and Sant'Anna, 2020; Sun and Abraham, 2020), we test the sensitivity of the results of the two-way fixed effects estimator to alternative estimators that account for potential issues surrounding inference with staggered treatment timing. We use two of these alternative estimators – Sun and Abraham (2020) and Borusyak, Jaravel, and Spiess (2021) – that can generalize to the static difference-in-differences estimate. Using the eventstudyinteract and did_imputation packages in Stata to implement these strategies, we find that the two-way fixed effects estimates are generally robust to these alternative estimators (see Online Appendix Table A.1). This suggests that differential treatment timing is not very problematic for the estimation in our setting. Part of this may be due to the fact that the group of never treated districts (the "always five-day" group in Figure 1) is quite large relative to previously treated control districts, thus reducing the share of the problematic two-by-two difference in differences comparisons using early treated districts as controls.

[9] As our data suggest that 92 percent of the four-day school week school districts use the four-day school week districtwide, we restrict our treatment group to include only those districts. An alternative way of operationalizing this variable uses the fraction of total district enrollment on a four-day school week schedule. The results (presented in Appendix Table A.2 in the Online Appendix) are robust to operationalizing the treatment variable in this way.

control for common shocks across all districts. ϵ_{dt} is an idiosyncratic error term clustered at the school district level.

Table 1: Descriptive Statistics

	All School Districts	Four-day only	Five-day only
Math Achievement	-0.01	-0.18	0.00
	(0.34)	(0.33)	(0.34)
ELA Achievement	-0.02	-0.15	-0.01
	(0.29)	(0.29)	(0.29)
Rural	0.68	0.89	0.66
	(0.47)	(0.31)	(0.47)
Total Enrollment	1206	232	1306
	(3531)	(326)	(3691)
Fraction Black	0.05	0.02	0.06
	(0.13)	(0.06)	(0.14)
Fraction White	0.74	0.72	0.74
	(0.26)	(0.26)	(0.26)
Fraction Hispanic	0.11	0.15	0.11
	(0.17)	(0.21)	(0.17)
Fraction Economically Disadvantaged	0.57	0.64	0.56
	(0.21)	(0.21)	(0.21)
Observations	25658	2340	23275

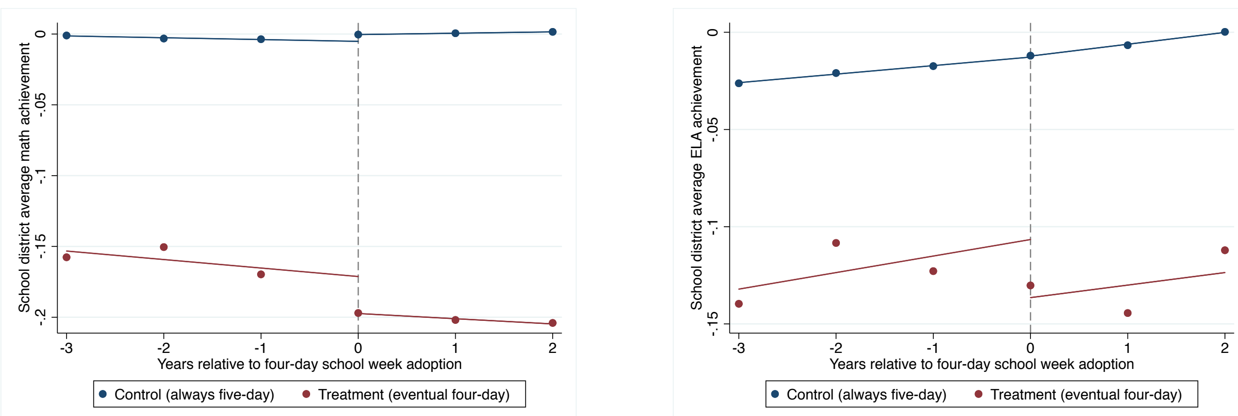
Note: "All School Districts" sample includes all school districts in the twelve four-day school week states; "Four-day only" sample includes only school districts that operate on a four-day school week for at least one year between 2009 and 2017; "Five-day only" sample includes only school districts that never operate on a four-day school week between 2009 and 2017. Standard deviations are given in parentheses.

Causal interpretation of the difference-in-differences estimates relies principally on the assumption that student achievement in districts that adopted four-day school weeks would have followed the trend observed among districts that did not adopt four-day school weeks (Lechner, 2011). To assess whether this identifying assumption holds in this setting, we compare the trends in mean achievement of treated schools to those of untreated schools before and after four-day school week adoption. As the timing of four-day school week adoption varies by school district, to compare achievement trends across treatment and control districts we stack adoption cohort-specific data sets together.[10] Figures 1a and 1b show these mean achievement trends for math and

[10] This follows a similar process as Thompson (2021). We first create individual four-day school week adoption year cohort data sets that include the full set of always five-day school week school districts and the group of four-day school week school districts that adopted the school schedule in the given year. A "years relative to four-day school week adoption" variable is created that captures event time for each cohort and then we stack these cohort data sets together across the "years relative to four-day school week adoption" variable.

ELA, respectively. Due to the lack of a change in the school schedule, the trend through event-time zero in mean achievement in control schools that always operate using a five-day school week is smooth. Prior to four-day school week policy adoption, there is a generally parallel mean math achievement pretrend for eventual four-day school week districts, and a noisier, potentially downward pretrend in mean ELA achievement. In event study results presented in Appendix Figure A.1 of the Online Appendix a similar pattern holds of a slightly noisier pretrend for ELA than math, but none of the pretrend estimates for either are statistically distinguishable from zero at conventional levels under any of the specifications we explore.[11] On balance, we suggest that this evidence is supportive of the key identifying assumption.[12]

Figure 1: Mean differences in achievement across treatment and control school districts



(a) Math Achievement (b) ELA Achievement

3. Results

The results of equation (1) using the full analytic sample are included in Panel A of Table 2. The results are highly imprecise for math and only marginally statistically significant for ELA when

[11] We use three estimators – the two-way fixed effects estimator, the Sun and Abraham (2020) estimator, and the Cengiz, et al. (2019) stacked cohort estimator – to estimate the event study version of equation (1), where the year before adoption is used as the reference point (point estimate in that year set to zero). Like the static difference-in-differences results, the event study results are very consistent across these three estimators, suggesting that differential treatment timing is not very problematic for the estimation in our setting.

[12] We note that in the descriptive results in these figures, the four-day week trends will exhibit more variation simply due to the much lower number of data points contributing to them.

only the four-day school week indicator and fixed effects are included. When including controls in the model, however, we find that math achievement falls by 0.032 standard deviations and ELA achievement falls by 0.029 standard deviations following the adoption of the four-day school week. These effects are a bit larger in magnitude when restricting the sample to include only rural school districts, finding that math achievement falls by 0.044 standard deviations and ELA achievement falls by between 0.031 standard deviations following the adoption of the four-day school week.

Table 2: Achievement Difference-in-Differences Results

	(1) Math	(2) Math	(3) Math	(4) ELA	(5) ELA	(6) ELA
			Panel A: Full Analytic Sample			
Four-day	-0.015	-0.032**	-0.044***	-0.016*	-0.029***	-0.031**
	(0.012)	(0.013)	(0.014)	(0.009)	(0.010)	(0.013)
Observations	25,596	21,983	14,119	25,247	21,683	13,963
R-squared	0.865	0.885	0.845	0.884	0.904	0.856
			Panel B: Low Time in School Treatment Group			
Four-day	-0.042**	-0.055***	-0.087***	-0.040**	-0.058***	-0.060**
	(0.018)	(0.020)	(0.025)	(0.016)	(0.018)	(0.025)
Observations	23,973	20,717	13,010	23,721	20,482	12,915
R-squared	0.869	0.889	0.848	0.888	0.906	0.858
			Panel C: Medium Time in School Treatment Group			
Four-day	0.016	0.024	0.004	0.003	0.006	0.003
	(0.019)	(0.021)	(0.021)	(0.014)	(0.018)	(0.021)
Observations	23,857	20,587	12,920	23,571	20,332	12,805
R-squared	0.869	0.889	0.849	0.888	0.907	0.859
			Panel D: High Time in School Treatment Group			
Four-day	-0.003	-0.041	-0.025	0.009	-0.015	-0.007
	(0.030)	(0.034)	(0.033)	(0.023)	(0.023)	(0.026)
Observations	24,072	20,780	13,103	23,767	20,510	12,977
R-squared	0.869	0.888	0.849	0.888	0.907	0.860
Controls		X	X		X	X
Rural Only			X			X

Notes: Each column of the table presents results from a separate regression containing the specified achievement dependent variable, controls (if specified), and school district and grade-by-school year fixed effects. Panel A presents results from the full analytic sample. The "Low Time in School Treatment Group" analyses (Panel B) include the stratified treatment sample containing four-day school week school districts with time in school in the bottom third of the distribution. The "Medium Time in School Treatment Group" analyses (Panel C) include the stratified treatment sample containing four-day school week school districts with time in school in the middle third of the distribution. The "High Time in School Treatment Group" analyses (Panel D) include the stratified treatment sample containing four-day school week school districts with time in school in the top third of the distribution. Robust standard errors, clustered at the school district-level are given in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

This multi-state evidence, which aligns with recent state-specific four-day school week research, suggests that four-day school weeks are detrimental to student achievement. A direct implication one might draw from this evidence is that the four-day school week is harmful to student learning and should be avoided. However, if school districts are determined to switch to the four-day school week for non-academic reasons, how might these school schedules be structured to mitigate these detrimental achievement effects? Given the lack of a one-size-fits-all approach to four-day school week implementation (see Thompson, et al. 2020), it may be the case that positive achievement impacts of well-functioning four-day school week models may be washed out by negative achievement impacts of poorly structured four-day school weeks. One specific element of these school schedules that has been cited as a key factor in the achievement impacts both broadly (Rivkin and Schiman 2015) and in the case of the four-day school week in Oregon (Thompson 2021) is time in school.[13]

To examine heterogeneity among this time in school parameter we group four-day school weeks by their amount of time in school. Using our data on time in school from the 2018-2019 cross-section, we create terciles of time in school within the treatment group stratifying the treatment sample into three groups – "low time in school" comprised of school districts in the bottom third of the four-day school week time in school distribution, "middle time in school" comprised of schools in the middle third of this distribution, and "high time in school" comprised of schools in the top third of this distribution.[14] We use these subgroups as separate treatment groups in separate regressions of equation (1).[15] Although we do not have time in

[13] This mechanism may also be a factor in the mixed findings from past work studying four-day school week effects among individual states.

[14] The terciles are generated using Stata xtile command. This command allocates tied values to the same tercile and therefore these groupings do not each fully account for 33.33 percent of the observations. Instead, 35.67 percent of the treatment districts are found in the bottom tercile, 31.25% are in middle tercile, and 33.08% are in top tercile.

[15] These results are generally robust to grouping the treatment group by quartile or quintile. Results with these groupings (Appendix Tables A.3 and A.4) are presented in the Online Appendix. One interesting difference that falls

school information for the five-day school week schools in the sample, we rely on the findings of Thompson (2021) that shows that time in school is relatively stable for five-day school weeks around the period of adoption for four-day school weeks. Thus, we assume that time in school for five-day weeks is uncorrelated with the timing of four-day school week adoption, and use all five-day week districts as controls for each four-day week groupings.

We find significant heterogeneity in the effects across these three treatment groupings when looking at the results of these analyses presented in Panels B through D of Table 2. The results indicate larger negative impacts on achievement for school districts that switch to the four-day school week with low instructional time, but find negligible impacts for those that maintain medium or high instructional time after the switch to the four-day school week.[16] We find that math achievement among the low time in school treatment schools fell by between 0.037 and 0.045 standard deviations after four-day school week adoption and ELA achievement among this group fell by 0.039 and 0.054 standard deviations. The negative impacts for this group are even more pronounced when restricting the sample to only rural schools. Using that sample, we find that math

---

out of the quintile treatment grouping is larger negative effects for the second lowest quintile compared to the lowest quintile of the distribution. Additional analysis suggests that the instructional time inadequacies may be able to be offset by greater use of fifth-day opportunities or later school start times among this lowest quintile group. School districts in the lowest quintile are more likely (25.5 percent to 19.2 percent) to provide off-day instructional activities for students and also start school about 10 minutes later than those in the second lowest quintile. This finding is not attributable to all states, however, as the states with a greater number of four-day school weeks (Oregon, Oklahoma, Arizona, and Montana) show large negative impacts for school districts in this lowest quintile of the distribution. Thus, future work should conduct greater analysis of other parameters of these four-day school week schedules that may compensate for low levels of time in school (e.g., off-day opportunities) and promote academic achievement.

[16] Similar conclusions are drawn when examining an event study version of equation (1) using the two-way fixed effects estimator, the Sun and Abraham (2020) estimator, and the stacked cohort estimator proposed by Cengiz, et al. (2019), As noted in Appendix Figure A.1 of the Online Appendix, there is a substantial reduction in math and ELA achievement – of between 0.059 and 0.091 standard deviations – in the year of four-day school week implementation for school districts with low time in school. In the following year these effects do diminish a bit for this group, but still find marginally significant impacts. It appears that there may be some transitory negative impacts in math during the year of four-day school week adoption for the medium time in school districts, but this effect reverses in sign in the following year. The high time in school treatment districts see increases of between 0.048 and 0.065 standard deviations in math achievement following four-day school week adoption. We find no statistically significant impacts on ELA achievement for either the medium or high time in school treatment groups.

and ELA achievement among the low time in school treatment schools fell by 0.07 and 0.056 standard deviations, respectively, after four-day school week adoption. No such significant effects were found among the treatment schools with middle or high time in school in either the full sample and rural only sample. This heterogeneity suggests that the blanket conclusion that four-day school weeks are detrimental to achievement drawn from previous literature and our results in Panel A of Table 2 misses some key nuances of these school schedules. Our results suggest that four-day school weeks that operate with adequate levels of time in school have no clear negative impacts on achievement and, instead, that it is operating four-day school weeks in a low-time-in-school environment that should be cautioned against.

4. Conclusion

This paper examined the impacts of four-day school weeks on achievement and achievement gaps in 12 states nationally. Using a difference-in-differences design with SEDA achievement data we find negative four-day school week achievement impacts for both math and ELA achievement when examining four-day school weeks nationally. These aggregate effects appear to mask heterogeneity in these achievement effects across the distribution of time in school among these four-day school week districts. When stratifying the treatment sample into low and middle/high time in school groups, we find statistically significant negative impacts on math and ELA achievement (between 0.041 and 0.07 standard deviations) for the low time in school four-day school week adopters, but no statistically significant impacts for the middle/high time in school four-day school week adopters.

One limitation of this analysis is that we are unable to differentiate between whether these effects are driven by the overall level of time in school used in conjunction with these school schedules, regardless of pre-adoption levels, or driven by school districts reducing time in school to facilitate the switch to the four-day school week. Future work that collects time in school information before and after the switch to the four-day school week will better assess these issues. Nonetheless, our results suggest that providing adequate instructional time in conjunction with the four-day school week should be a key focus of policymakers interested in adopting these four-day school week schedules. Adequate time in school may negate potential negative achievement impacts inherent to the four-day school week schedule (e.g., weekend learning loss; earlier school start times). Low time in school in conjunction with the four-day school week appears to be extremely problematic for academic achievement and school districts and states should be cognizant of these negative consequences. School districts may be able to carve out additional instructional time each day (e.g., limiting recess or lunch time) without changing overall time in school, but future work is needed to determine whether these creative scheduling techniques have any tangible achievement impacts. State policymakers may also play a critical role in improving the level of time in school by increasing state minimum annual hours requirements.

Our results also have implications related to lost exposure to in-person academic instruction due to COVID-19 public health concerns. Survey evidence from Spring 2020 (EdWeek Research Center; Garet, et al., 2020) suggests that instructional time was reduced by as much as half of what it was prior to the COVID-19 school closures. While many schools have been able to provide more instruction during the 2020-2021 school year, much has come through remote/asynchronous instruction, with in-person instructional time still greatly reduced in many districts. Thus, our results here likely represent a lower bound on the academic achievement impacts of lost time in school due to COVID-19.

References

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.

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' Rule to Estimate the Effect of Class Size on Scholastic Achievement. *The Quarterly Journal of Economics*, 114(2), 533-575.

Aucejo, E. M., French, J., Araya, M. P. U., & Zafar, B. (2020). The impact of COVID-19 on student experiences and expectations: Evidence from a survey. *Journal of Public Economics*, *191*, 104271.

Bacher-Hicks, A., Goodman, J., & Mulhern, C. (2021). Inequality in household adaptation to schooling shocks: Covid-induced online learning engagement in real time. *Journal of Public Economics*, *193*, 104345.

Barnum, M., & Bryan, C. (2020). America's great remote-learning experiment: What surveys of teachers and parents tell us about how it went. *Chalkbeat*

Baron, E. J., Goldstein, E. G., & Wallace, C. T. (2020). Suffering in silence: How COVID-19 school closures inhibit the reporting of child maltreatment. *Journal of public economics*, *190*, 104258.

Borusyak, K., Jaravel, X., & Spiess, J. (2021). *Revisiting Event Study Designs: Robust and Efficient Estimation*. Working Paper.

Brixey, E. (2020). 50-State comparison: Instructional time policies. Denver, CO: Education Commission of the States. Retrieved from https://www.ecs.org/50-state-comparison-instructional-time-policies/

Callaway, B., & Sant'Anna, P. H. (2020). Difference-in-differences with multiple time periods. *Journal of Econometrics*.

Cattaneo, M. A., Oggenfuss, C., & Wolter, S. C. (2017). The more, the better? The impact of instructional time on student performance. *Education Economics*, 25(5), 433-445.

Cengiz, D., Dube, A., Lindner, A., & Zipperer, B. (2019). The effect of minimum wages on low-wage jobs. *The Quarterly Journal of Economics*, *134*(3), 1405-1454.

Chetty, R., Friedman, J. N., Hendren, N., & Stepner, M. (2020). Real-time economics: A new platform to track the impacts of COVID-19 on people, businesses, and communities using private sector data. *NBER Working Paper*, *27431*.

De Chaisemartin, C., & d'Haultfoeuille, X. (2020). Two-way fixed effects estimators with heterogeneous treatment effects. *American Economic Review*, *110*(9), 2964-96.

Duckworth, A. L., Kautz, T., Defnet, A., Satlof-Bedrick, E., Talamas, S. N., Luttges, B. L., & Steinberg, L. (2021, February 8). Students Attending School Remotely Suffer Socially, Emotionally, and Academically. https://doi.org/10.31234/osf.io/rpz7h

Garet, M., Rickles, J., Bowdon, J., & Heppen, J. (2020). *National Survey of Public Education's Response to COVID-19*. American Institutes for Research

Goodman-Bacon, A. (2018). *Difference-in-differences with variation in treatment timing* (No. w25018). National Bureau of Economic Research.

Jackson, C. K. (2020). Does school spending matter? The new literature on an old question. In *Confronting inequality: How policies and practices shape children's opportunities.* (pp. 165-186). American Psychological Association.

Krueger, A. B. (1999). Experimental estimates of education production functions. *The Quarterly Journal of Economics*, 114(2), 497-532.

Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). Projecting the potential impact of COVID-19 school closures on academic achievement. *Educational Researcher*, *49*(8), 549-565.

Lake, R. & Heyward, G. (2021) *Analysis: In Thousands of Districts, 4-Day School Weeks Are Robbing Students of Learning Time for What Amounts to Hygiene Theater*, The 74, retrieved from https://www.the74million.org/article/analysis-in-thousands-of-districts-4-day-school-weeks-are-robbing-students-of-learning-time-for-what-amounts-to-hygiene-theater/

Lavy, V. (2015). Do differences in schools' instruction time explain international achievement gaps? Evidence from developed and developing countries. *The Economic Journal*, 125(588), F397-F424.

Lechner, M. (2011). The Estimation of Causal Effects by Difference-in-Difference Methods. *Foundations and Trends® in Econometrics*, *4*(3), 165-224.

Morton, E. (2020). Effects of four-day school weeks on school finance and achievement: Evidence from Oklahoma. *Educational Researcher*, 0013189X20948023.

Phelps, G., Corey, D., DeMonte, J., Harrison, D., & Loewenberg Ball, D. (2012). How much English language arts and mathematics instruction do students receive? Investigating variation in instructional time. *Educational Policy*, *26*(5), 631-662.

Reardon, S. F., Kalogrides, D., & Ho, A. D. (2017). Linking US School District Test Score Distributions to a Common Scale. CEPA Working Paper No. 16-09. *Stanford Center for Education Policy Analysis*.

Reardon, S. F., Kalogrides, D., Ho, A. D., Shear, B., Shores, K., & Fahle, E. (2019). Stanford Education Data Archive http://purl.stanford.edu/db586ns4974

Reardon, S. F., Kalogrides, D., & Shores, K. (2019). The geography of racial/ethnic test score gaps. *American Journal of Sociology*, *124*(4), 1164-1221.

Reardon, S. F., Shear, B. R., Castellano, K. E., & Ho, A. D. (2017). Using heteroskedastic ordered probit models to recover moments of continuous test score distributions from coarsened data. *Journal of Educational and Behavioral Statistics*, *42*(1), 3-45.

Rivkin, S. G., Schiman, J. C. (2015). Instruction Time, Classroom Quality, and Academic Achievement. *The Economic Journal*, 125(588), F425-F448.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review*, 94(2), 247-252.

Sun, L., & Abraham, S. (2020). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*.

Takaku, R., & Yokoyama, I. (2020). What school closure left in its wake: Contrasting evidence between parents and children from the first COVID-19 outbreak. *Available at SSRN 3693484*.

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.

Thompson, P. N., Gunter, K., Schuna Jr, J. M., & Tomayko, E. J. (2020). Are all four-day school weeks created equal? A national assessment of four-day school week policy adoption and implementation. *Education Finance and Policy*, 1-50.

Online Appendix for *Only a Matter of Time? The Role of Time in School on Four-Day School Week Achievement Impacts*

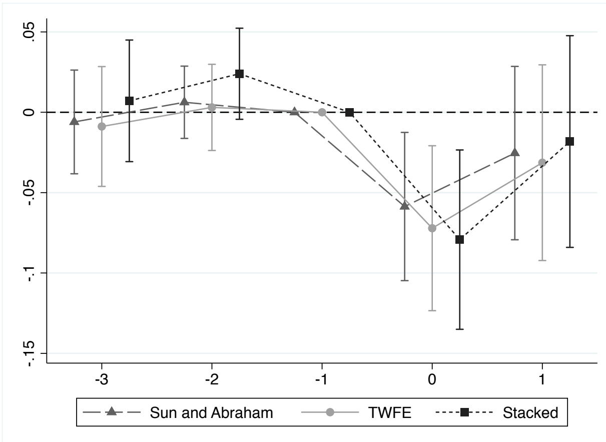
Appendix Table A.1: Sensitivity of Static D-in-D Results to Staggered Treatment Estimators

	(1)	(2)	(3)	(4)	(5)	(6)
	Math	Math	Math	ELA	ELA	ELA
			Panel A: Full Analytic Sample			
Four-day	-0.015	-0.018	-0.029***	-0.016	-0.022*	-0.019**
	(0.012)	(0.014)	(0.011)	(0.009)	(0.012)	(0.008)
Observations	25,596	24,475	25,596	25,247	24,190	25,247
R-squared	0.865		0.826	0.870		0.832
			Panel B: Low Time in School Treatment Group			
Four-day	-0.042**	-0.040**	-0.032***	-0.040**	-0.028**	-0.021**
	(0.018)	(0.017)	(0.025)	(0.016)	(0.016)	(0.010)
Observations	23,973	23,641	23973	23,721	23,385	23,721
R-squared	0.869		0.810	0.888		0.816
			Panel C: Medium Time in School Treatment Group			
Four-day	0.016	0.030	0.022	0.003	0.002	-0.008
	(0.019)	(0.023)	(0.014)	(0.014)	(0.019)	(0.012)
Observations	23,857	23,570	23,857	23,571	23,305	23,571
R-squared	0.869		0.809	0.888		0.837
			Panel D: High Time in School Treatment Group			
Four-day	-0.003	-0.011	-0.048	0.009	-0.018	-0.017
	(0.030)	(0.023)	(0.014)	(0.023)	(0.021)	-0.014
Observations	24,072	23,657	24,072	23,767	23,401	23,767
R-squared	0.869		0.830	0.888		0.838
Method	TWFE	BJS	SA	TWFE	BJS	SA

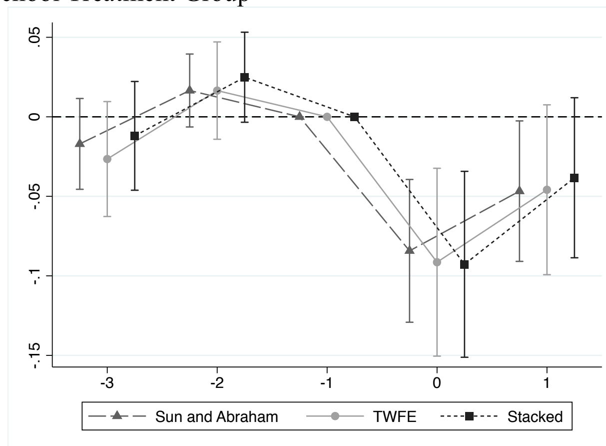
Notes: Each column of the table presents results from a separate regression containing the specified achievement dependent variable and school district and grade-by-school year fixed effects. The "TWFE" method indicates the results from the traditional two-way fixed effects estimator outlined in equation (1). The "BJS" method indicates the results from the static difference-in-differences version of the Borusyak, Jaravel, and Spiess (2021) estimator. The "SA" method indicates the results from the static difference-in-differences version of Sun and Abraham (2020). The Panel A presents results from the full analytic sample. The "Low Time in School Treatment Group" analyses (Panel B) include the stratified treatment sample containing four-day school week school districts with time in school in the bottom third of the distribution. The "Medium Time in School Treatment Group" analyses (Panel C) include the stratified treatment sample containing four-day school week school districts with time in school in the middle third of the distribution. The "High Time in School Treatment Group" analyses (Panel D) include the stratified treatment sample containing four-day school week school districts with time in school in the top third of the distribution. Robust standard errors, clustered at the school district-level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Figure A.1: Event Study Results

Panel A: Low Time in School Treatment Group

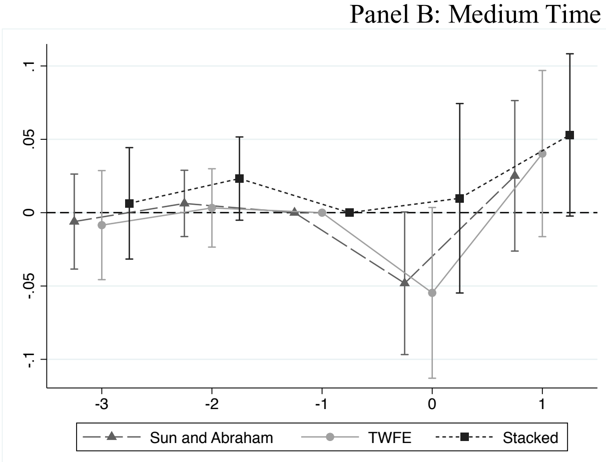


Math Achievement

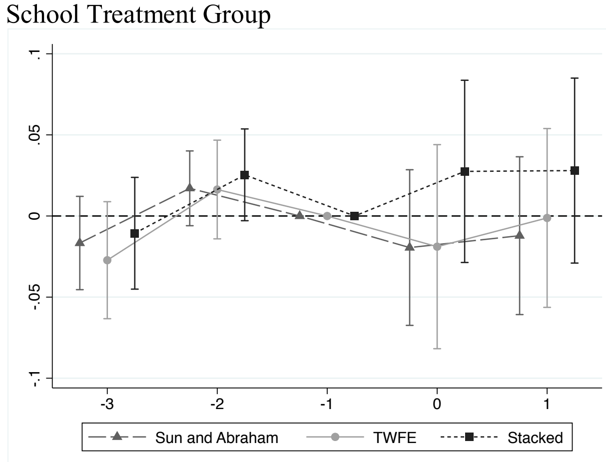


ELA Achievement

Panel B: Medium Time in School Treatment Group

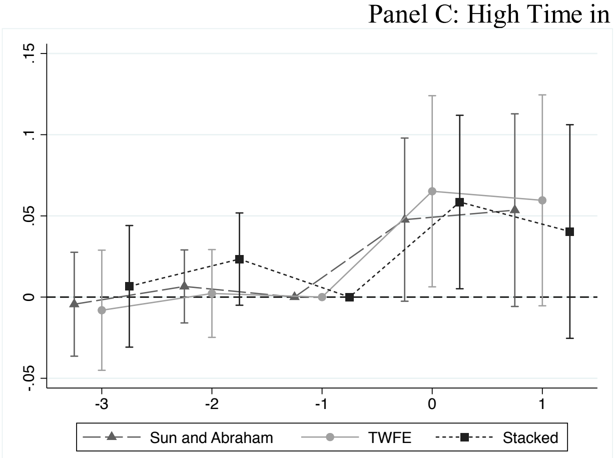


Math Achievement

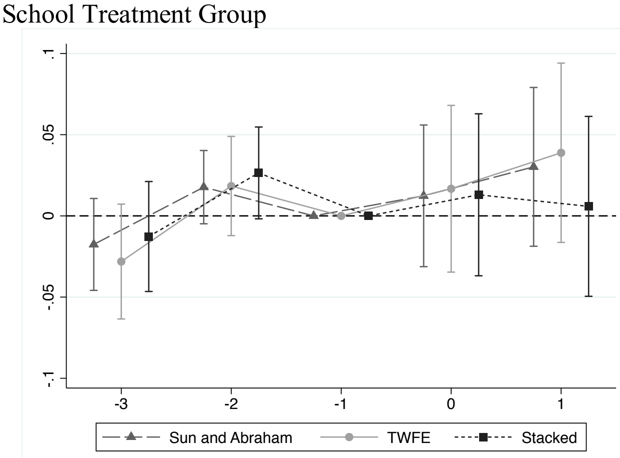


ELA Achievement

Panel C: High Time in School Treatment Group



Math Achievement



ELA Achievement

Notes: Coefficients are relative to year before four-day school week adoption. The "TWFE" line indicates the results from the traditional two-way fixed effects estimator of the event study specification of equation (1). The "Stacked" line indicates the results from using the stacked cohort method proposed by Cengiz, et al. (2019). The "Sun and Abraham" line indicates the results from the static difference-in-differences version of Sun and Abraham (2020). The "Low Time in School Treatment Group" analyses (Panel A) include the stratified treatment sample containing four-day school week school districts with time in school in the bottom third of the distribution. The "Medium Time in School Treatment Group" analyses (Panel B) include the stratified treatment sample containing four-day school week school districts with time in school in the middle third of the distribution. The "High Time in School Treatment Group" analyses (Panel C) include the stratified treatment sample containing four-day school week school districts with time in school in the top third of the distribution. Robust standard errors, clustered at the school district-level are given in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Appendix Table A.2: Results Using Fractional Treatment Variable

	(1) Math	(2) Math	(3) Math	(4) ELA	(5) ELA	(6) ELA
			Panel A: Full Analytic Sample			
Four-day	-0.014	-0.032**	-0.045***	-0.015	-0.029***	-0.031**
	(0.012)	(0.013)	(0.014)	(0.009)	(0.010)	(0.013)
Observations	25,639	22,018	14,153	25,281	21,709	13,988
R-squared	0.865	0.885	0.846	0.884	0.904	0.856
			Panel B: Low Time in School Treatment Group			
Four-day	-0.035**	-0.045**	-0.069***	-0.038***	-0.055***	-0.057**
	(0.016)	(0.018)	(0.022)	(0.015)	(0.017)	(0.022)
Observations	24,055	20,773	13,066	23,799	20,538	12,971
R-squared	0.869	0.889	0.848	0.888	0.906	0.858
			Panel C: Medium Time in School Treatment Group			
Four-day	0.016	0.022	-0.001	0.007	0.011	0.010
	(0.021)	(0.023)	(0.023)	(0.015)	(0.019)	(0.022)
Observations	23,798	20,550	12,883	23,509	20,288	12,761
R-squared	0.869	0.889	0.848	0.888	0.907	0.859
			Panel C: High Time in School Treatment Group			
Four-day	-0.000	-0.046	-0.026	0.010	-0.016	-0.009
	(0.029)	(0.033)	(0.032)	(0.023)	(0.023)	(0.025)
Observations	24,081	20,779	13,110	23,775	20,516	12,983
R-squared	0.869	0.888	0.849	0.888	0.907	0.861
Controls		X	X		X	X
Rural Only			X			X

Notes: Each column of the table presents results from a separate regression containing the specified achievement dependent variable, controls (if specified), and school district and grade-by-school year fixed effects. The “Four-day” variable is operationalized as the the fraction of total district enrollment on a four-day school week schedule. Panel A presents results from the full analytic sample. The “Low Time in School Treatment Group” analyses (Panel B) include the stratified treatment sample containing four-day school week school districts with time in school in the bottom third of the distribution. The “Medium Time in School Treatment Group” analyses (Panel C) include the stratified treatment sample containing four-day school week school districts with time in school in the middle third of the distribution. The “High Time in School Treatment Group” analyses (Panel C) include the stratified treatment sample containing four-day school week school districts with time in school in the top third of the distribution. Robust standard errors, clustered at the school district-level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table A.3: Results for Treatment Quartiles of Time in School

	(1) Math	(2) Math	(3) Math	(4) ELA	(5) ELA	(6) ELA
	Panel A: Low Time in School Treatment Group					
Four-day	-0.025	-0.027	-0.044*	-0.026	-0.030	-0.013
	(0.019)	(0.020)	(0.026)	(0.017)	(0.020)	(0.026)
Observations	23,811	20,598	12,919	23,561	20,364	12,825
R-squared	0.870	0.889	0.849	0.889	0.907	0.859
	Panel B: Medium Low Time in School Treatment Group					
Four-day	-0.029	-0.041*	-0.060**	-0.030*	-0.049**	-0.060**
	(0.022)	(0.025)	(0.026)	(0.018)	(0.022)	(0.025)
Observations	23,732	20,482	12,805	23,457	20,236	12,699
R-squared	0.869	0.889	0.849	0.888	0.907	0.860
	Panel C: Medium High Time in School Treatment Group					
Four-day	0.027	0.025	-0.005	0.009	0.007	0.009
	(0.028)	(0.030)	(0.031)	(0.023)	(0.022)	(0.028)
Observations	23,698	20,495	12,845	23,423	20,242	12,736
R-squared	0.869	0.889	0.848	0.888	0.907	0.859
	Panel D: High Time in School Treatment Group					
Four-day	-0.004	-0.050	-0.019	0.013	-0.015	-0.001
	(0.034)	(0.039)	(0.037)	(0.026)	(0.027)	(0.029)
Observations	23,881	20,621	12,970	23,591	20,364	12,853
R-squared	0.869	0.889	0.850	0.889	0.908	0.861
Controls		X	X		X	X
Rural Only			X			X

Notes: Each column of the table presents results from a separate regression containing the specified achievement dependent variable, controls (if specified), and school district and grade-by-school year fixed effects. The “Low Time in School Treatment Group” analyses (Panel A) include the stratified treatment sample containing four-day school week school districts with time in school in the bottom quartile of the distribution. The “Medium Low Time in School Treatment Group” analyses (Panel B) include the stratified treatment sample containing four-day school week school districts with time in school in the second lowest quartile of the distribution. The “Medium High Time in School Treatment Group” analyses (Panel C) include the stratified treatment sample containing four-day school week school districts with time in school in the second highest quartile of the distribution. The “High Time in School Treatment Group” analyses (Panel D) include the stratified treatment sample containing four-day school week school districts with time in school in the top quartile of the distribution. Robust standard errors, clustered at the school district-level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Appendix Table A.4: Results for Treatment Quintiles of Time in School

	(1) Math	(2) Math	(3) Math	(4) ELA	(5) ELA	(6) ELA
	Panel A: Low Time in School Treatment Group					
Four-day	-0.010	-0.011	-0.038	-0.013	-0.023	-0.012
	(0.021)	(0.023)	(0.032)	(0.020)	(0.025)	(0.033)
Observations	23,693	20,505	12,833	23,446	20,273	12,742
R-squared	0.870	0.889	0.849	0.889	0.907	0.860
	Panel B: Medium Low Time in School Treatment Group					
Four-day	-0.070***	-0.078***	-0.095***	-0.067***	-0.082***	-0.091***
	(0.024)	(0.025)	(0.029)	(0.020)	(0.022)	(0.027)
Observations	23,583	20,379	12,738	23,327	20,146	12,644
R-squared	0.870	0.890	0.850	0.889	0.908	0.860
	Panel B: Medium Time in School Treatment Group					
Four-day	0.014	0.005	-0.009	0.010	0.012	0.000
	(0.021)	(0.024)	(0.025)	(0.017)	(0.022)	(0.025)
Observations	23,664	20,439	12,781	23,386	20,188	12,670
R-squared	0.869	0.889	0.849	0.889	0.907	0.860
	Panel B: Medium High Time in School Treatment Group					
Four-day	0.009	-0.006	-0.010	0.009	-0.001	0.013
	(0.037)	(0.041)	(0.042)	(0.025)	(0.023)	(0.031)
Observations	23,834	20,594	12,923	23,550	20,338	12,811
R-squared	0.869	0.888	0.848	0.888	0.907	0.859
	Panel C: High Time in School Treatment Group					
Four-day	0.001	-0.048	-0.027	0.010	-0.030	-0.017
	(0.041)	(0.047)	(0.045)	(0.035)	(0.038)	(0.038)
Observations	23,568	20,391	12,770	23,296	20,143	12,662
R-squared	0.871	0.890	0.851	0.890	0.908	0.862
Controls		X	X		X	X
Rural Only			X			X

Notes: Each column of the table presents results from a separate regression containing the specified achievement dependent variable, controls (if specified), and school district and grade-by-school year fixed effects. The "Low Time in School Treatment Group" analyses (Panel A) include the stratified treatment sample containing four-day school week school districts with time in school in the bottom quintile of the distribution. The "Medium Low Time in School Treatment Group" analyses (Panel B) include the stratified treatment sample containing four-day school week school districts with time in school in the second lowest quintile of the distribution. The "Medium Time in School Treatment Group" analyses (Panel C) include the stratified treatment sample containing four-day school week school districts with time in school in the median quintile of the distribution. The "Medium High Time in School Treatment Group" analyses (Panel D) include the stratified treatment sample containing four-day school week school districts with time in school in the second highest quintile of the distribution. The "High Time in School Treatment Group" analyses (Panel C) include the stratified treatment sample containing four-day school week school districts with time in school in the top quintile of the distribution. Robust standard errors, clustered at the school district-level are given in parentheses. *** p<0.01, ** p<0.05, * p<0.1