

WHEN THE BELL TOLLS: THE EFFECTS OF SCHOOL STARTING TIMES ON ACADEMIC ACHIEVEMENT

Peter Hinrichs

Georgetown Public Policy
Institute
Georgetown University
Washington, DC 20007
plh24@georgetown.edu

Abstract

A number of high schools across the United States have moved to later bell times on the belief that their previous bell times were too early for the “biological clocks” of adolescents. In this article I study whether doing so improves academic performance. I first focus on the Twin Cities metropolitan area, where Minneapolis and several suburban districts have made large policy changes but St. Paul and other suburban districts have maintained early schedules. I use individual-level ACT data on all individuals from public high schools in this region who took the ACT between 1993 and 2002 to estimate the effects of school starting times on ACT scores. I then employ school-level data on schedules and test scores on statewide standardized tests from Kansas and Virginia to estimate the effects of bell times on achievement for a broader sample. The results do not suggest an effect of school starting times on achievement.

1. INTRODUCTION

This article studies a relatively inexpensive policy that may potentially raise academic achievement: starting the school day later. High school students in the United States generally begin the school day quite early. One recent random nationwide survey of schools found that the average high school began the school day at 7:54 a.m. in 2001–2 (Wolfson and Carskadon 2005).¹ At the same time, sleep scientists believe that adolescents undergo a phase shift of the sleep-wake cycle and that early school schedules are not consistent with the circadian rhythms of adolescents (Dement and Vaughan 1999). Given the sleep habits of adolescents, it is certainly plausible that delaying the school day could increase the amount that students learn in school and consequently raise test scores. On the other hand, perhaps students are in fact able to adapt to early schedules, or perhaps the later schedules chosen by some schools may still be too early. Either of these possibilities could explain why shifting the school day may not raise test scores.

The possible benefits from shifting to a later school day have inspired legislation such as U.S. Representative Zoe Lofgren’s (D-CA) ZZZ’s to A’s Act, which would encourage secondary schools nationwide to begin the school day after 9:00 a.m., as well as a proposed law in Minnesota to prevent public schools in that state from beginning the school day before 8:00 a.m. Although neither of those pieces of legislation has ultimately been enacted, some individual school districts have decided to alter school schedules on their own. Most notably for the purpose of this study, Minneapolis Public Schools changed the high school schedule in 1997–98 from 7:15 a.m.–1:45 p.m. to 8:40 a.m.–3:10 p.m.² Another notable example is Arlington Public Schools in Virginia, which moved high school start times from 7:30 a.m. to 8:15 a.m. in 2001. There is a movement toward change in other districts as well. For instance, there is currently a parent-led effort to push back high school schedules in Virginia’s Fairfax County Public Schools.³

This article focuses on the Twin Cities of Minneapolis and St. Paul and their surrounding suburbs. Minneapolis’s policy change provides a strong basis for identifying the effects of school starting times, as Minneapolis’s “twin” city St. Paul has maintained an early 7:30 a.m.–2:00 p.m. schedule for many years. St. Paul provides a useful comparison for Minneapolis because the two cities are geographically contiguous and quite similar. Table 1 shows some of

-
1. Carskadon et al. (1998) found that tenth-grade boys at a high school in Rhode Island that began at 7:20 a.m. woke up on average at 6:05 a.m. and that girls woke up on average at 5:58 a.m.
 2. The schedule has since moved to 8:30 a.m.–3:00 p.m.
 3. This district is the nation’s thirteenth largest school district when measured by the number of students (see table 93 of the *Digest of Education Statistics, 2008*, http://nces.ed.gov/pubs2009/2009020_2a.pdf).

Table 1. Characteristics of the Population in Minneapolis and St. Paul

Variable	Minneapolis	St. Paul
Median household income (1999)	\$37,974	\$38,774
% of population aged 25 and over with BA (2000)	37.4%	32.0%
% white in population (2000)	62.5%	64.0%
% black in population (2000)	17.8%	11.4%
% Asian in population (2000)	6.1%	12.3%
% Hispanic in population (2000)	7.6%	7.9%
Unemployment rate (1990)	4.3%	4.6%
Unemployment rate (2000)	3.0%	3.2%
Unemployment rate (2005)	4.2%	4.4%
Median age of population in years (2000)	31.2	31.0
Land area in square miles	54.9	52.8
Population (2000)	382,618	287,151
Population growth (1990 to 2000)	3.9%	5.5%

Notes: Income data are from Table P53 of Census 2000 Summary File 3. The percentage of the population with at least a BA degree is the author's calculation based on Table P37 of Census 2000 Summary File 3. The racial composition statistics are the author's calculation based on Table P8 of Census 2000 Summary File 1. Unemployment data are from Table C-3 of County and City Data Book: 2007. Age data are from Table P13 of Census 2000 Summary File 1. Land area data are from Table C-1 of County and City Data Book: 2007. Population data are from Table P1 of Census 2000 Summary File 1 and Table P001 of Census 1990 Summary Tape File 1.

the similarities. Journalistic accounts suggest that political momentum and personal opinions of the leaders involved explain why Minneapolis made a large policy change but St. Paul did not.⁴ An additional fortunate feature of the Twin Cities metropolitan area is that several large suburban districts have made large changes to later bell schedules in various years, while others have maintained early schedules.⁵ Thus the identifying variation used in this article does not come about solely from Minneapolis's policy change.

In this article I estimate whether pushing back school start times results in higher student achievement. I use individual-level ACT test score data on all individuals from public high schools in the Twin Cities metropolitan area who took the ACT between 1993 and 2002 to estimate the effects of school starting

4. A newspaper article about Minneapolis's school starting time change states, "Minneapolis officials said they found a way to make a drastic change without spending any more money" (Peterson 1997b). Meanwhile, another newspaper article explains, "Though St. Paul school officials have considered later times, they've said the expense for additional buses and inconvenience in changing elementary school schedules are too great. But the idea is still on the table" (Smith 1998).
5. The suburban districts that are coded as ever having moved the school day back by thirty minutes or more from one year to the next during the time period under consideration are Bloomington, Eden Prairie, Edina, Minnetonka, and Roseville. Chaska is also coded as gradually moving back its schedule.

times on ACT scores. I then employ school-level data on starting times and test scores on statewide standardized tests from Kansas and Virginia to estimate the effects of bell times on achievement for a broader sample of students. The data differ between the states, and the states have their separate advantages. For example, Minnesota has a particularly large amount of within-school variation in school schedules over time, while Kansas was chosen because a large amount of data on school schedules and test scores on required statewide exams is readily available. But despite the differences in the data among the three states, the results are consistent in suggesting that there is not an effect of school starting times on achievement. The results are robust to a variety of potential issues, including selection into taking the ACT and measurement error in the Minnesota school schedule data. The Minnesota results are especially precise: a 95 percent confidence for the effect on ACT scores of moving the school day back one hour is $(-.2306, .1825)$.⁶

Researchers at the University of Minnesota's Center for Applied Research and Educational Improvement (CAREI) conducted a study of starting times in Minneapolis concurrently with Minneapolis's policy change (Wahlstrom, Wrobel, and Kubow 1998; Wahlstrom et al. 2001; Wahlstrom 2002). The study finds a modest increase in attendance rates in Minneapolis over time, and teachers report that students seem less sleepy after the policy change. But insofar as the study focuses on academic achievement, it considers only unadjusted differences in grades at a point in time between districts or unadjusted changes in grades over time within a single district.⁷ Economists have also recently begun to study whether students perform better in classes that meet later in the day.⁸ Cortes, Bricker, and Rohlfs (2009) show that students in Chicago public schools receive lower grades and are more likely to be absent in first-period classes compared to other classes. Dills and Hernandez-Julian

6. Expressed in terms of standard deviations, this confidence interval is $(-.0512, .0406)$.

7. Wahlstrom, Wrobel, and Kubow (1998) find that a sample of students in Minneapolis has higher self-reported grade point averages than a sample of students at an unnamed comparison district at a point in time. Wahlstrom (2002) surveys a random sample of students at a broader set of secondary schools in the Twin Cities and also finds an association between later start times and higher grades. Wahlstrom et al. (2001) find a small but statistically insignificant increase in grades in Minneapolis from the three years before the policy change to the three years after the policy change. In addition, Wolfson and Carskadon (2003) review research by sleep scientists on the relationship between sleep habits and grades. The general pattern in these studies is that the students who have better sleep habits do in fact receive higher grades, but these articles generally measure only simple correlations.

8. More broadly, research on school starting times is related to recent research by economists on the structure of the school day (Eren and Millimet 2007) and the amount of time spent in school (Krashinsky 2006; Eren and Millimet 2007; Marcotte 2007; Pischke 2007; Hansen 2008; Marcotte and Hemelt 2008; Sims 2008). Even more broadly, economists studying time use have expanded their analysis beyond the simple hours trade-off between labor and leisure to include sleep (Biddle and Hamermesh 1990; Brochu, Deri Armstrong, and Morin 2009), other activities (Aguar and Hurst 2007; Krueger 2007; Guryan, Hearst, and Kearney 2008), and timing of activities throughout the day (Hamermesh 1998, 1999a, 1999b, 2002; Hamermesh, Myers, and Pocock 2008).

Table 2. Summary Statistics for Minnesota

Variable	Mean (SD)
ACT composite score	21.94 (4.50)
School start time (hours after 7:00 a.m.)	0.646 (0.319)
Length of school day	6.66 (0.20)
School is new	0.003
Male	0.443
Missing gender	0.003
Asian	0.067
Black	0.030
Hispanic	0.013
White	0.794
Missing race	0.069
Family income < \$30,000	0.121
Family income > \$60,000	0.430
Missing income	0.147
English not home language	0.044
Missing home language	0.069
Grade 10	0.006
Grade 11	0.617
<i>N</i>	196,617

Note: The table shows means with standard deviations below in parentheses.

(2008) and Carrell, Maghakian, and West (2010) have found that college students receive higher grades in classes that meet later in the day. Unlike the other existing research by economists showing that grades are higher in later periods, I focus on test score effects of changing the timing of the entire school day. And unlike the CAREI study of Minneapolis, I make use of a control group in which starting times did not change; I also include individual- and school-level covariates.

2. MINNESOTA ACT TEST SCORE ANALYSIS

Data and Empirical Methods

The Minnesota test score analysis uses three main types of data: ACT test score and student survey data, school starting and ending time data, and school- and district-level control variables. Table 2 shows summary statistics for these data.

The ACT data are student-level data on all students in grades 10–12 in public schools from the forty-eight districts (seventy-three schools) of the seven-county Twin Cities metropolitan area who took the ACT between 1993–94 and 2001–2.⁹ The data set contains one observation for each test sitting. I am not able to distinguish first-time test takers from repeat test takers. The forty-eight districts include the large urban districts of Minneapolis and St. Paul, a number of mid-sized suburban districts, and some smaller districts that are on the periphery of the Twin Cities metropolitan area. The data include individual-level test score and demographic information for those who took the test. The main outcome is the composite ACT score. I use ACT data because Minneapolis’s policy change occurred before statewide accountability tests gained prominence in Minnesota. An advantage of these data is that the ACT is a high-level test that has real consequences for individual students. Importantly, such test scores are at least somewhat malleable, as previous research has in fact found effects of various policies or conditions on ACT or SAT scores.¹⁰ Even more to the point, people have claimed that moving the school day later does raise scores on college entrance exams. For example, George Will wrote in his *Washington Post* column on 4 March 2010, “When Edina, Minn., changed its high school start from 7:25 a.m. to 8:30 a.m., math/verbal SAT scores rose substantially” (Will 2010). In addition, the ACT is more of an achievement test and less of an aptitude test than the SAT. One limitation is that not every student takes the ACT. However, Minnesota is an “ACT state,” and the ACT participation rate is thus high: between 1994 and 2002, the annual participation rate among high school graduates in Minnesota varied from a low of 59 percent in 1995 and 1996 to a high of 66 percent in 2000 and 2001.

The starting and ending time data for Minnesota were acquired by contacting schools and school districts directly, and school- and district-level covariates come from the National Center for Education Statistics’s Common Core of Data (CCD). The CCD includes information on spending, demographic composition, and percentage of students at the school enrolled in free lunch under the National School Lunch Program.

I use the Minnesota data to regress ACT scores on school starting times, controlling for various covariates. The main ACT test score equations I estimate are of the form

-
9. All regressions drop two observations from tenth graders in White Bear Lake, a district that has separate schools for grades 9–10 and grades 11–12. The two campuses in this district have slightly different schedules.
 10. For instance, Card and Payne (2002) find that school finance equalization reduces SAT score gaps, Card and Rothstein (2007) find that reducing racial segregation narrows the black-white SAT score gap, Krueger and Whitmore (2001) find that lower class sizes in early grades in the Tennessee Student Teacher Achievement Ratio (STAR) experiment results in higher ACT or SAT scores later on, and Pallais (2009) finds that a merit scholarship program in Tennessee raised ACT scores.

$$y_{ist} = starttime_{st}\gamma + x_{ist}'\beta_1 + z_{st}'\beta_2 + \delta_t + \alpha_s + \eta_{st} + \varepsilon_{ist}. \quad (1)$$

Here y_{ist} is the test score from student i from school s in year t , $starttime_{st}$ is the number of hours after 7:00 a.m. that the school day begins, x_{ist} are individual-level covariates, z_{st} are district- or school-level covariates such as the length of the school day, δ_t are year fixed effects, α_s are school fixed effects, η_{st} denotes a full set of school-specific linear time trends, ε_{ist} is the error term, and the other components of the equation are parameters to be estimated. I estimate these models by ordinary least squares (OLS), and the standard errors allow for clustering at the school level. I also estimate models for various demographic groups because it may be the case that social or biological factors that facilitate going to sleep and waking up at certain times differ by demographic group.¹¹

It is conceptually possible that the start time, end time, and length of school day have their own independent effects on outcomes. However, it is not possible to simultaneously identify the effects of these three variables due to the perfect collinearity among them.¹² I estimate parameters on the start time and the length of school day variables; thus the coefficient on the start time variable estimates the effect of making the school day earlier without changing its length. This implementation is in line with the goal of this article, which is to estimate the effect of the timing of a fixed-length school day rather than the effect of shortening the school day by having it begin later in the morning.

Results

Figure 1 plots average test scores in Minneapolis and St. Paul over time. This figure does not give much indication that Minneapolis's policy change had an effect on test scores. This is supported by table 3, which includes the full sample and reports results from regressions of composite ACT scores on school starting times when controlling for various covariates.¹³ The estimates are quite precise. For example, in the full specification of column 8, a 95 percent confidence interval for the effect of moving the school day back one hour is $(-.2306, .1825)$. When expressed in terms of standard deviations of ACT scores, the confidence interval is $(-.0512, .0406)$. The insignificant coefficients on the treatment variable are in contrast to the estimated coefficients on the covariates, which are generally statistically significant and of the expected sign.

11. Wahlstrom (2002, p. 196) writes about Minneapolis's policy change, "Overall, the group most positively affected by the later start time comprised the African American students, whose rates in every outcome indicator showed statistically significant positive gain."
12. Data from suburban districts are necessary for identifying the effect of the length of day, since there is no variation within or between Minneapolis and St. Paul.
13. The "school is new" variable for Minnesota is coded from knowledge of what schools opened in which years in addition to relying on the CCD.

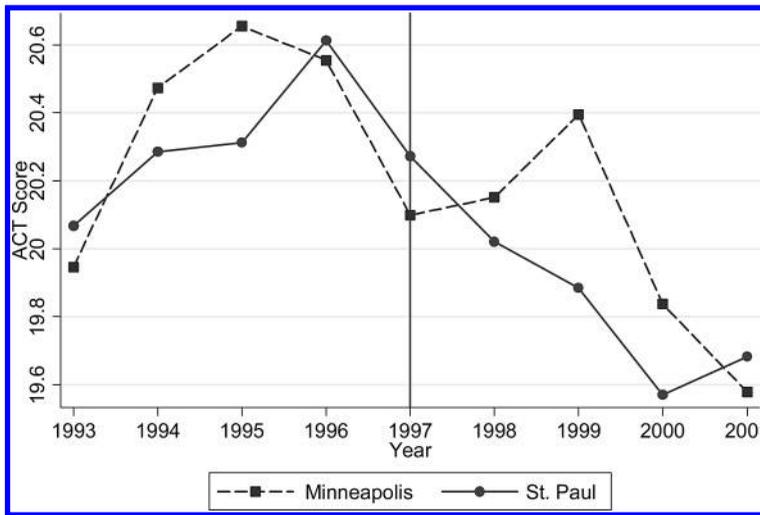


Figure 1. Average ACT Scores in Minneapolis and St. Paul

For example, students from high-income families score higher than students from the excluded category of middle-income families, who in turn score higher than students from low-income families. White students score higher than those in the excluded racial category.¹⁴ Asians do not score significantly differently from the excluded category, while blacks and Hispanics score lower. Students in tenth grade score higher than those in eleventh grade, who in turn score higher than those in twelfth grade. This result is presumably due to more capable students taking the test earlier.

The ACT results hold up under a variety of robustness checks and alternative specifications. The full results are available upon request. In short, I find no effect of school schedules on test scores when doing the following:

- Stratifying the sample by subgroups defined by race, income, or gender;
- Handling selection bias by controlling for the school-level participation rate or the inverse Mills ratio of the school-level participation rate;¹⁵
- Running quantile regressions to estimate effects on features of the test score distribution other than the mean;
- Handling measurement error in school schedules by instrumenting with a second measurement;¹⁶

14. The excluded category consists of those who identify themselves as American Indian, multiracial, or “other.”

15. I also find no effect of school starting times on ACT-taking rates.

16. The second measurement comes from the Internet Archive’s “Wayback Machine” (www.archive.org/web/web.php), Blake and Ballman (1994), Peterson (1997a, 1997b), and Draper (1998, 2001). See, for instance, Ashenfelter and Krueger (1994) on instrumenting with a second measurement.

Table 3. Effect of School Starting Times on Composite ACT Scores

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
School start time (hours after 7:00 a.m.)	-0.3089 [0.4821]	-0.0796 [0.1657]	0.0792 [0.1276]	-0.0213 [0.1011]	-0.2193 [0.4751]	-0.0681 [0.1663]	0.0912 [0.1281]	-0.0241 [0.1036]
Length of school day					1.2115 [0.4812]*	0.3862 [0.3016]	0.4024 [0.2549]	-0.1130 [0.4201]
School is new			-0.2836 [0.3562]	-0.1297 [0.3366]			-0.2903 [0.3540]	-0.1287 [0.3371]
Male			0.2585 [0.0271]**	0.2546 [0.0274]**			0.2583 [0.0271]**	0.2546 [0.0274]**
Missing gender			-0.7286 [0.2232]**	-0.7756 [0.2203]**			-0.7302 [0.2229]**	-0.7753 [0.2203]**
Asian			0.0319 [0.1609]	0.0355 [0.1593]			0.0325 [0.1609]	0.0354 [0.1593]
Black			-2.4793 [0.1770]**	-2.4658 [0.1821]**			-2.4794 [0.1771]**	-2.4657 [0.1821]**
Hispanic			-0.4326 [0.1499]**	-0.4419 [0.1510]**			-0.4330 [0.1498]**	-0.4419 [0.1510]**
White			0.7346 [0.0973]**	0.7272 [0.0975]**			0.7348 [0.0972]**	0.7272 [0.0975]**
Missing race			1.0369 [0.1115]**	1.0390 [0.1129]**			1.0384 [0.1116]**	1.0390 [0.1129]**
Family income < \$30,000			-0.9207 [0.0556]**	-0.9238 [0.0556]**			-0.9206 [0.0556]**	-0.9238 [0.0556]**
Family income > \$60,000			0.6981 [0.0512]**	0.7019 [0.0511]**			0.6982 [0.0512]**	0.7019 [0.0511]**

Missing income	0.1892 [0.0600]**	0.1797 [0.0586]**	0.1897 [0.0604]**	0.1797 [0.0586]**
English not home language	-1.4302 [0.1222]**	-1.4076 [0.1220]**	-1.4309 [0.1224]**	-1.4075 [0.1220]**
Missing home language	-0.8694 [0.0675]**	-0.8692 [0.0668]**	-0.8690 [0.0675]**	-0.8692 [0.0668]**
Grade 10	2.7504 [0.1869]**	2.7513 [0.1886]**	2.7522 [0.1872]**	2.7509 [0.1885]**
Grade 11	1.1933 [0.0369]**	1.1890 [0.0368]**	1.1932 [0.0369]**	1.1890 [0.0368]**
Year fixed effects	No	Yes	No	Yes
School fixed effects	No	Yes	No	Yes
School time trends	No	No	No	No
N	196,617	196,617	196,617	196,617

Notes: The table shows OLS estimates of equation 1. Standard errors that allow for clustering at the school level are in brackets. The omitted racial category is those who identify themselves as American Indian, multiracial, or "other."
*significant at 5%; **significant at 1%.

Table 4. Summary Statistics of Attendance Data

<i>Variable</i>	<i>Mean (SD)</i>
Attendance rate	92.32 (5.40)
School start time (hours after 7:00 a.m.)	0.652 (0.343)
Length of school day	6.65 (0.19)
School is new	0.002 (0.039)
Percent nonwhite in school	18.70 (20.92)
Percent on free lunch	13.31 (16.00)
<i>N</i>	644

Note: The table shows school-level means (with standard deviations in parentheses) weighted by the number of students at the school.

- Controlling for a dummy for a large schedule change in the current school year, controlling for last year's schedule, stratifying by the time of year the test was taken, using a quadratic in the start time, or dropping the large urban districts of Minneapolis and St. Paul; and
- Interacting the starting time with demographic characteristics of the school district, school spending, or latitude and longitude.

3. ATTENDANCE EFFECTS IN MINNESOTA

There is a possibility that later school starting times will raise the probability that high school students will attend school. As noted earlier, Wahlstrom et al. (2001) found that the attendance rate in Minneapolis was rising over the time period in which the schedule change took place. To try to determine whether this finding is spurious or whether it is attributable to the school starting time change, I employ school-level data from Minnesota on average daily attendance rates. The data are publicly available and were downloaded from the Web site of the Minnesota Department of Education. Table 4 shows summary statistics. As a first pass at determining whether the increased attendance in Minneapolis was spurious, figure 2 shows that attendance was also increasing in St. Paul over the time period. I also estimate equations of the form

$$\gamma_{st} = starttime_{st}\gamma + z_{st}'\beta + \delta_t + \alpha_s + \eta_s t + \varepsilon_{st}. \quad (2)$$

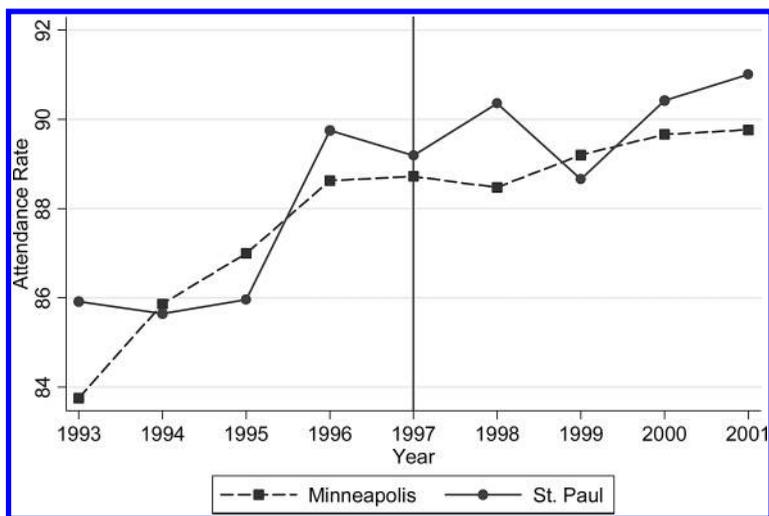


Figure 2. Attendance in Minneapolis and St. Paul

Here y_{st} is the attendance rate from school s in year t , $starttime_{st}$ is the number of hours after 7:00 a.m. that school begins, z_{st} are district- or school-level covariates such as the length of the school day, δ_t are year fixed effects, α_s are school fixed effects, η_{st} denotes a full set of school-specific linear time trends, ε_{st} is the error term, and the other components of the equation are parameters to be estimated. Table 5 shows regression results using the same sample of schools as in the ACT test score analysis. The regression results do not give evidence that delaying school start times will improve overall attendance.¹⁷

4. SCHOOL-LEVEL TEST SCORE EFFECTS IN OTHER STATES

Kansas State Assessment Analysis

Table 6 shows summary statistics for the Kansas data. I obtained school-level test score data on the Kansas Mathematics Assessment, the Kansas Reading Assessment, the Kansas Science Assessment, and the Kansas Social Studies Assessment from the Kansas Department of Education for every public high school in the state from 2000–2001 to 2005–6.¹⁸ This includes data for the

17. Cortes, Bricker, and Rohlfs (2009) find that attendance is lower in Chicago public schools in first period than in later periods. One possible reason for the difference in results between that piece and this one is that Cortes, Bricker, and Rohlfs have attendance data by period, whereas I have only overall attendance data. It could be the case that attendance is lower in earlier periods regardless of what time school starts.

18. The Kansas data cover the entire state but due to confidentiality rules, data for a group at a school will be missing if there are fewer than ten students in that group in the school. There were also a small number of schools in the school schedule file that did not merge to the test score files, but these schools are presumably small. I drop one school, Hope Street Charter Academy, that is coded

Table 5. Effect of School Starting Times on Attendance in Minnesota

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
School start time (hours after 7:00 a.m.)	-0.9847 [0.9438]	1.2184 [0.7710]	0.9846 [0.8213]	0.3418 [0.6862]	-0.7324 [0.9481]	1.2406 [0.7831]	1.0161 [0.8319]	0.4642 [0.7868]
Length of school day					2.7773 [1.7309]	0.7193 [3.2707]	0.8884 [3.4219]	3.4430 [8.5226]
School is new			1.1934 [1.5064]	0.9046 [1.5324]			1.1861 [1.5000]	0.9052 [1.5368]
Percent nonwhite in school			0.2153 [0.1540]	0.2148 [0.2445]			0.2164 [0.1553]	0.2148 [0.2445]
Percent on free lunch			-0.1019 [0.1911]	-0.0934 [0.2413]			-0.1035 [0.1921]	-0.0958 [0.2403]
School fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School time trends	No	No	No	Yes	No	No	No	Yes
N	644	644	644	644	644	644	644	644

Notes: The table shows least squares estimates of equation 2 with school-level attendance rates on the left-hand side weighted by the number of students at the school. Standard errors that allow for clustering at the school level are in brackets.
* significant at 5%; ** significant at 1%.

Table 6. Summary Statistics for Kansas

Variable	Mean (SD)
Reading test score	79.90 (4.08)
School start time (hours after 7:00 a.m.)	0.981 (0.247)
Length of school day	7.06 (0.17)
School is new	0.005 (0.072)
White female share in grade 11	0.399 (0.101)
Nonwhite male share in grade 11	0.089 (0.093)
Nonwhite female share in grade 11	0.087 (0.094)
Fraction on free lunch	0.189 (0.141)
<i>N</i>	1,666

Note: The table shows school-level means (with standard deviations in parentheses) weighted by the number of test takers.

eleventh-grade reading and tenth-grade mathematics exams for each year in the sample beginning in 2001–2, as well as data for the eleventh-grade social studies and tenth-grade science exams for every other year in the sample beginning in 2000–2001. The starting and ending time data from Kansas were provided by the Kansas State Department of Education. School- and district-level controls come from the CCD.

I use the Kansas data to regress school average scores on school starting times, controlling for school-level control variables from the CCD. I estimate an equation like equation 2 but with test scores rather than attendance on the left-hand side. I weight these regressions by the number of students taking the test, I again allow for clustering at the school level, and I estimate models for various subgroups in addition to models for all students.

An advantage of the Kansas data over the Minnesota data is that there is likely less measurement error with the start and end time data because the Kansas State Department of Education keeps careful records of school schedules. Moreover, selection bias is likely less problematic with the Kansas tests than with the ACT. A disadvantage is that there are not as many districts in Kansas making large policy changes.

in 2002 as having a schedule of 8:30 a.m.–8:00 p.m. I drop another school, Jennings High, whose district is coded as spending \$45,000 per pupil in 2005.

The results obtained for Minnesota are largely supported by Kansas. Table 7 reports regression results for the Kansas reading exam. The results are precise enough for some of the covariates to be significant, but the results do not suggest an effect of school start times on test scores. In results not reported here, I find no effect on scores on the mathematics, science, and social studies assessments. I also find no effect when stratifying by gender or when limiting the sample to those eligible for free or reduced price lunches. In addition, many of the same robustness checks and alternative specifications I did for Minnesota are possible to do for Kansas. I find no effect when I interact the starting time with the demographic composition of the district, spending per pupil, or latitude and longitude. I also find that the lack of an effect is not due to schools in the first year of a large change and that there is no effect when controlling for the previous year's starting time.

On the whole, the results from Minnesota are confirmed by Kansas. But interestingly, when I use a quadratic in the starting time, both the linear and quadratic term are marginally significant. The results from the quadratic model suggest that delaying school start times until 8:07 a.m. would raise test scores but that delaying them further would lower test scores. Although this is plausible, there are at least two reasons to be skeptical of this result. First, significance is lost when controlling for a quadratic in the length of the school day. Second, although a spline regression with knots every forty-five minutes has a significant negative coefficient on the segment from 8:30 a.m. to 9:15 a.m., none of the other segments in that regression or in a spline regression with knots every thirty minutes is significant.

Virginia End-of-Course Exam Analysis

I have also estimated the effects of school starting times on test scores on statewide end-of-course exams for a sample of schools in Virginia. The sample consists of the seventy-five schools in the nineteen districts in Virginia's Region 4, which corresponds roughly to the Virginia suburbs of Washington, DC. This region was chosen because there have been several districts there, most notably Arlington Public Schools in 2001–2, that have moved to later high school schedules. At the same time, data are available on widely taken standardized tests. The school start and end time data were acquired by contacting schools and districts directly, and test score data for the years 2000–2001 to 2006–7 were downloaded from the Virginia Department of Education Web site. I am also able to control for school-level covariates from the CCD. The results, which are not reported here but are available upon request, are somewhat imprecise, but they do not give evidence for an effect of the timing of the school day on test scores.

Table 7. Effect of School Starting Times on Kansas Reading Scores

Variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
School start time (hours after 7:00 a.m.)	1.4405 [1.6072]	-0.4779 [0.8721]	-0.7480 [0.8971]	1.9903 [1.5515]	1.4463 [1.6076]	0.2116 [1.3095]	-0.1702 [1.2817]	0.9514 [1.8497]
Length of school day					0.5210 [1.3567]	1.2541 [1.3100]	1.0448 [1.2636]	-1.9397 [2.2033]
School is new			-1.1690 [0.2581]**	2.6145 [2.1658]			-1.1703 [0.2582]**	2.6231 [2.1682]
White female share in grade 11			3.2916 [1.2467]**	2.8943 [1.6135]			3.2945 [1.2489]**	2.9491 [1.5917]
Nonwhite male share in grade 11			-5.6151 [3.4737]	-3.6022 [4.6596]			-5.5653 [3.4608]	-3.6477 [4.6904]
Nonwhite female share in grade 11			-2.0856 [4.1965]	-0.9070 [5.5869]			-1.9462 [4.2042]	-1.0853 [5.6099]
Fraction on free lunch			-7.8157 [2.9289]**	-0.9441 [4.7351]			-7.8008 [2.9183]**	-0.8743 [4.7361]
School fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Year fixed effects	No	Yes	Yes	Yes	No	Yes	Yes	Yes
School time trends	No	No	No	Yes	No	No	No	Yes
N	1,666	1,666	1,666	1,666	1,666	1,666	1,666	1,666

Notes: The table shows estimates of equation 2. Regressions are weighted by the number of students taking the exam. Standard errors that allow for clustering at the school level are in brackets.
*significant at 5%; **significant at 1%.

5. DISCUSSION

Even though the results do not suggest an effect of school starting times on outcomes, several caveats are worth bearing in mind. First, the time of day a test is taken may matter. The ACT is generally given early on Saturday mornings at a uniform time, whereas the Kansas State Assessments are given during school hours.¹⁹ A problem with using ACT data is that the early testing time may not be in sync with the usual sleep-wake cycles of students who attend late-starting high schools. Thus even if later start times increase the amount that students learn, this may not necessarily translate into higher scores on exams taken in the morning. On the other hand, with the Kansas State Assessments I may be confounding the effect of what time of day is best for learning with what time of day is best for taking tests. Presumably this would work to the advantage of later-starting high schools, which should also tend to give tests later in the day due to the later schedule. If this is true, the biases would work in opposite directions for the Minnesota and Kansas results, so it is reassuring that the empirical results for the two states are consistent with one another. Second, the effects of moving to later school schedules may be lower in the Midwest than in other regions of the United States, since Midwesterners may naturally operate on earlier schedules. Early schedules could be either a natural carryover of habits from the time when the Midwest had a more agrarian economy, or they could be due to the fact that prime-time television shows are aired at an earlier nominal time in the Central Time Zone than in the other U.S. time zones. This latter possibility is supported by the results of Hamermesh, Myers, and Pocock (2008), who find that television schedules have a real effect on the timing of sleep and work.

But taking the results at face value, there are several reasons why there might not actually be an effect of school starting times on achievement. First, when schools start early and students lose sleep as a result, students may learn less per unit of time but they may learn more overall outside school due to being awake longer. Second, students may be able to adapt to early starting times by re-optimizing their sleep patterns. For example, they may be able to catch up on sleep over the weekend. The approach of Biddle and Hamermesh (1990), who view sleep as being endogenous and derive a demand curve for it, may be particularly relevant here.²⁰ Third, students may be able to adapt to early schedules in other ways. Environmental stimulation can keep people awake when they are sleep deprived, as can chemical stimulation in the form of

19. ACT testing rooms begin to admit students at 8:00 a.m., and testing begins at 9:00 a.m.

20. Also see Brochu, Deri Armstrong, and Morin (2009), who use Canadian time use data to show that sleep time rises when the unemployment rate is higher.

soda and coffee (Dement and Vaughan 1999).²¹ Fourth, although the biological clocks of adolescents may be better suited for learning later in the day, their teachers may perform better earlier in the day. This may have a counteracting effect.²² Fifth, there may be an effect of changing school starting times through synchronizing or desynchronizing the schedules of students with those of their parents, either in the morning or in the afternoon. For example, later starting times could result in less time spent with parents in the morning without changing the amount of time spent with parents in the afternoon or evening. On the other hand, it could result in more time spent with parents. Sixth, “zero hour” classes or other before-school activities may nullify the effects of later starting times. Seventh, with later school schedules, students may miss instructional time due to being dismissed early to participate in athletic competitions or other extracurricular activities. Eighth, there may be no effect of starting times in my data because even the relatively later starting times may still be too early for the biological clocks of adolescents. Ninth, to the extent that the tests I use in this analysis measure innate ability or to the extent that schools do not teach what is tested, there is only limited scope for most school policies to affect test scores. A final possibility is based on sleep scientists’ model of sleep. According to the model, there is a homeostatic process that makes people more sleepy the longer they go without sleep, and there is the biological clock that tends to make most people sleepy around the middle of the afternoon and the middle of the night (Borbely 1982). The interplay between these two forces governs how sleepy people are at various times. With later school starting times, students may be more likely to be in school during their mid-afternoon dip in alertness. This negative effect could counteract the presumed positive effect of not needing to wake up as early in the morning.

It should be noted that even if there is not a measurable effect on test scores, there might be other benefits to switching to a later school day. For example, feeling better rested and less irritable may have effects on health and happiness even if it does not result in improved academic performance. Postponing the schedule may lead to less tardiness and truancy. In addition, later school starting times may lead to less crime and fewer teenage pregnancies because students are “incapacitated” into later in the afternoon. There may also be better athletic performance or fewer automobile accidents.²³ But there may be

21. Although there is a movement away from soda vending machines inside schools, there are potential benefits from having them there.

22. Wahlstrom, Wrobel, and Kubow (1998) found that teachers in Minneapolis had a mixed reaction to the school start time change.

23. Alternatively, moving the school day back may result in more automobile accidents if doing so leads to more adolescents being on the road during the morning and evening rush hours.

costs as well. If a school unilaterally moves its bell times, it may become more difficult to coordinate athletic events with other schools. There is also concern about after-school sports practices, especially in the dark winter months. And it may create problems for other extracurricular activities or make it more difficult for students to work after school.²⁴

6. CONCLUSION

This article has employed a quasi-experiment in Minnesota and data from Kansas and Virginia to estimate the effects of school starting times on test scores. It has also considered the effects on school attendance in Minnesota. The results suggest that there are no effects on the outcomes considered. This holds across several states and types of tests as well as under a variety of robustness checks and alternative specifications. Although I find no effect, this result is interesting and somewhat surprising in light of the fact that others have argued that there is an effect.

In the era of school accountability, schools have real incentives to raise student achievement. They may turn to later bell times as one policy to improve achievement. But the results here suggest that doing so would be ineffectual. However, this does not rule out the possibility that moving the school day later would affect other outcomes. But the clear policy implication from this article is that doing so would not raise achievement as proxied by test scores.

I thank Joshua Angrist, Duncan Chaplin, David Figlio, Nora Gordon, William Gormley, Daniel Hamermesh, Andrew Healy, Heinrich Hock, Harry Holzer, and two anonymous referees for helpful suggestions and discussions. I also thank lunch seminar participants at Georgetown University, seminar participants at Mathematica Policy Research, and conference participants at the 2011 American Economic Association, and the 2010 American Education Finance Association, Association for Policy Analysis and Management, Eastern Economic Association, Midwest Economics Association, and Southern Economic Association conferences. I thank Mohamad Alloush, Ryan Greenfield, and especially Thomas McCall for indispensable research assistance. I thank Eric Peterson, David Shawver, and Robert Ziomek at ACT, Inc., and Tiffany Ellerman and Joycelyn Parish from the Kansas State Department of Education for data assistance. I also thank the forty-eight Minnesota school districts for agreeing to participate in the study, as well as everyone at Minneapolis Public Schools I spoke to about their policy change. I thank Georgetown University for funding.

REFERENCES

Aguiar, Mark, and Erik Hurst. 2007. Measuring trends in leisure: The allocation of time over five decades. *Quarterly Journal of Economics* 122(3): 969–1006.

24. However, it is unclear that reducing work hours should be considered a disadvantage, especially if it is concomitant with an increase in time spent sleeping.

- Ashenfelter, Orley, and Alan Krueger. 1994. Estimates of the economic return to schooling from a new sample of twins. *American Economic Review* 84(5): 1157–73.
- Biddle, Jeff E., and Daniel S. Hamermesh. 1990. Sleep and the allocation of time. *Journal of Political Economy* 98(5): 922–43.
- Blake, Laurie, and Laura S. Ballman. 1994. Getting to school later could be the new rule for city adolescents. *Star Tribune*, 12 May.
- Borbely, A. A. 1982. A two process model of sleep regulation. *Human Neurobiology* 1(3): 195–204.
- Brochu, Pierre, Catherine Deri Armstrong, and Louis-Philippe Morin. 2009. The “trendiness” of sleep: An empirical investigation into the cyclical nature of sleep time. University of Ottawa Working Paper No. 0909E.
- Card, David, and A. Abigail Payne. 2002. School finance reform, the distribution of school spending, and the distribution of student test scores. *Journal of Public Economics* 83(1): 49–82.
- Card, David, and Jesse Rothstein. 2007. Racial segregation and the black-white test score gap. *Journal of Public Economics* 91(11–12): 2158–84.
- Carrell, Scott E., Teny Maghakian, and James E. West. 2010. A’s from zzzz’s? The causal effect of school start time on the academic achievement of adolescents. Unpublished paper, University of California, Davis.
- Carskadon, Mary A., Amy R. Wolfson, Christine Acebo, Oma Tzischinsky, and Ronald Seifer. 1998. Adolescent sleep patterns, circadian timing, and sleepiness at a transition to early school days. *Sleep* 21(8): 871–81.
- Cortes, Kalena E., Jesse Bricker, and Chris Rohlfs. 2009. The role of specific subjects in education production functions: Evidence from morning classes in Chicago public high schools. Unpublished paper, Syracuse University.
- Dement, William C., and Christopher Vaughan. 1999. *The promise of sleep: A pioneer in sleep medicine explores the vital connection between health, happiness, and a good night’s sleep*. New York: Dell.
- Dills, Angela K., and Rey Hernandez-Julian. 2008. Course scheduling and academic performance. *Economics of Education Review* 27(6): 646–54.
- Draper, Norman. 1998. Later school starting times produce worrisome problems for many parents. *Star Tribune*, 15 March.
- Draper, Norman. 2001. Start times a-changin’ at many metro schools. *Star Tribune*, 17 July.
- Eren, Ozkan, and Daniel L. Millimet. 2007. Time to learn? The organizational structure of schools and student achievement. *Empirical Economics* 32(2–3): 301–32.
- Guryan, Jonathan, Erik Hearst, and Melissa Kearney. 2008. Parental education and parental time with children. *Journal of Economic Perspectives* 22(3): 23–46.
- Hamermesh, Daniel S. 1998. When we work. *American Economic Review (AEA Papers and Proceedings)* 88(2): 321–25.

- Hamermesh, Daniel S. 1999a. Changing inequality in markets for workplace amenities. *Quarterly Journal of Economics* 114(4): 1085–123.
- Hamermesh, Daniel S. 1999b. The timing of work over time. *Economic Journal* 109(452): 37–66.
- Hamermesh, Daniel S. 2002. Timing, togetherness, and time windfalls. *Journal of Population Economics* 15(4): 601–23.
- Hamermesh, Daniel S., Caitlin Knowles Myers, and Mark L. Pockock. 2008. Cues for timing and coordination: Latitude, Letterman, and longitude. *Journal of Labor Economics* 26(2): 223–46.
- Hansen, Benjamin. 2008. School year length and student performance: Quasi-experimental evidence. Unpublished paper, University of California, Santa Barbara.
- Krashinsky, Harry. 2006. How would one extra year of high school affect performance in university? Evidence from a unique policy change. Unpublished paper, University of Toronto.
- Krueger, Alan B. 2007. Are we having more fun yet? Categorizing and evaluating changes in time allocation. *Brookings Papers on Economic Activity* 2007(2): 193–215.
- Krueger, Alan B., and Diane M. Whitmore. 2001. The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from Project STAR. *Economic Journal* 111(468): 1–28.
- Marcotte, Dave E. 2007. Schooling and test scores: A mother-natural experiment. *Economics of Education Review* 26(5): 629–40.
- Marcotte, Dave E., and Steven W. Hemelt. 2008. Unscheduled school closings and student performance. *Education Finance and Policy* 3(3): 316–38.
- Pallais, Amanda. 2009. Taking a chance on college: Is the Tennessee education lottery scholarship program a winner? *Journal of Human Resources* 44(1): 199–222.
- Peterson, David. 1997a. Starting times. *Star Tribune*, 7 May.
- Peterson, David. 1997b. All eyes on Minneapolis' start times. *Star Tribune*, 8 May.
- Pischke, Jorn-Steffen. 2007. The impact of length of the school year on student performance and earnings: Evidence from the German short school years. *Economic Journal* 117(523): 1216–42.
- Sims, David P. 2008. Strategic responses to school accountability measures: It's all in the timing. *Economics of Education Review* 27(1): 58–68.
- Smith, Maureen M. 1998. St. Paul students, parents offer ideas to cut absenteeism. *Star Tribune*, 10 March.
- Wahlstrom, Kyla L. 2002. Accommodating the sleep patterns of adolescents within current educational structures: An uncharted path. In *Adolescent sleep patterns*, edited by Mary A. Carskadon, pp. 172–97. Cambridge: Cambridge University Press.

Wahlstrom, Kyla, Mark Davison, Jiyoung Choi, and Jesse N. Ross. 2001. Minneapolis Public Schools start time study: Executive summary. Unpublished paper, Center for Applied Research and Educational Improvement (CAREI), University of Minnesota.

Wahlstrom, Kyla, Gordon Wrobel, and Patricia Kubow. 1998. Minneapolis Public Schools start time study: Executive summary. Unpublished paper, Center for Applied Research and Educational Improvement (CAREI), University of Minnesota.

Will, George F. 2010. How to ruin a child: Too much esteem, too little sleep. *Washington Post*, 4 March.

Wolfson, Amy R., and Mary A. Carskadon. 2003. Understanding adolescents' sleep patterns and school performance: A critical appraisal. *Sleep Medicine Reviews* 7(4): 491–506.

Wolfson, Amy R., and Mary A. Carskadon. 2005. A survey of factors influencing high school start times. *NASSP Bulletin* 89(642): 47–66.

This article has been cited by:

1. Eric R Eide, Mark H Showalter. 2013. Sleep and Student Achievement. *Eastern Economic Journal* **38**:4, 512-524. [[CrossRef](#)]
2. Finley Edwards. 2012. Early to rise? The effect of daily start times on academic performance. *Economics of Education Review* **31**:6, 970-983. [[CrossRef](#)]