Economics of Education Review 31 (2012) 970–983

Contents lists available at SciVerse ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

Early to rise? The effect of daily start times on academic performance

Finley Edwards*

Department of Economics, Colby College, 8800 Mayflower Hill, Waterville, ME 04901, United States

ARTICLE INFO

Article history:
Received 25 March 2011
Received in revised form 18 June 2012
Accepted 6 July 2012

JEL classification:
I2

Keywords:
Start times
Test scores
Education production function
Middle school
Sleep

ABSTRACT

Local school districts often stagger daily start times for their schools in order to reduce busing costs. This paper uses data on all middle school students in Wake County, NC from 1999 to 2006 to identify the causal effect of daily start times on academic performance. Using variation in start times within schools over time, the effect is a two percentile point gain in math test scores – roughly fourteen percent of the black–white test score gap. I find similar results for reading scores and using variation in start times across schools. The effect is stronger for students in the lower end of the distribution of test scores. I find evidence supporting increased sleep as a mechanism through which start times affect test scores. Later start times compare favorably on cost grounds to other education interventions which result in similar test score gains.

1. Introduction

What time should the school day begin? There is considerable variation in daily start times, both across the nation and within individual communities, with some schools beginning at 7:30 a.m. or earlier, and others beginning at 9:00 a.m. or later.[1] The issue of timing is by no means trivial. Districts stagger the start times of different schools in order to reduce the number of buses used and thus reduce transportation costs. However, if beginning the school day early in the morning has a negative impact on academic performance, staggering start times may not be justified on a cost–benefit basis.

In recent years, school start times have received considerable attention in the popular press (see, for example, Kalish (2008) and Trudeau (2007)). Proponents of later start times argue that students in early starting schools do not receive enough sleep and that performance can be increased by beginning the school day at a later time. Several school districts have responded by delaying the start of their school day, and a 2005 congressional resolution (H. Con. Res. 200) was introduced by Rep. Zoe Lofgren, recommending that all secondary schools start at 9:00 or later nationwide. Despite this attention, the relationship between start times and academic performance is not well understood.

This paper uses data on all middle school students in Wake County, North Carolina from 1999 to 2006 to examine how start times affect the performance of middle school students (grades 6–8) on standardized tests. Wake County is uniquely suited for this purpose because it has considerable variation in start times across schools within a single large school district, as well as variation in start times within schools over time. The differences in start time across schools are a result of the bus schedule. As the student population has grown, the school district has changed the start times for many individual schools in order to maintain a balanced bus schedule, generating variation in start times within schools. Using both sources of variation in start times, I find that a 1 h delay in start time increases standardized test scores on both math and reading tests by three percentile points. Since start times may be correlated with other characteristics that determine test scores, I also

* Tel.: +1 207 859 5236.
E-mail addresses: fedwards@colby.edu, Finleyedwards@gmail.com

[1] Throughout this paper, all times are a.m., unless otherwise indicated.

0272-7757/$ – see front matter
http://dx.doi.org/10.1016/j.econedurev.2012.07.006

estimate the effect using only variation in start times within the same schools over time, and find a two percentile point improvement in math and a one and a half percentile point improvement in reading. The effect of start times on academic performance is robust to different specifications and sources of variation. The magnitude of the effect is similar to the difference in test scores for one additional year of parental education. The impact of later start times on test scores is persistent. Conditional on a high school fixed effect, a 1 h later start time in grade eight is associated with an increase in test scores in grade ten similar in magnitude to the increase in grade eight. Quantile regression results show that the effect of later start times is stronger for the lower end of the distribution of test scores, suggesting that later start times may be particularly effective in meeting minimum competency requirements.

Public attention on school start times largely focuses on high school start times. In contrast, this paper examines how start times affect middle school students, in order to examine possible mechanisms through which later start times might raise student performance. The rationale typically given for start times affecting academic performance is primarily biological. Earlier start times may result in fewer hours of sleep, since students may not fully compensate for earlier rising times with earlier bed times. In particular, adolescents have difficulty adjusting to early bed times due to the timing of the release of the hormone melatonin (Dahl & Lewin, 2002). A reduced amount of sleep has been demonstrated to reduce students' cognitive ability (Meijer, Habekothe, & Van Den Wittenboer, 2000), which in turn could reduce learning, resulting in lower test scores. Since most students enter adolescence during their middle school years, the adolescent hormone explanation predicts that the effect of later start times will be larger for older students. I find evidence supporting this explanation: among middle school students, the impact of start times is greater for older students. However, I also find evidence of other potential mechanisms as well. Students who begin school later have fewer absences, watch less television and spend more time on homework each week. These factors may also explain why later-starting students have higher test scores. Later start times also have the potential to be a cost effective method of increasing test scores. For example, Krueger (1999) finds similar increases in test scores from smaller class sizes, but at cost an order of magnitude larger than the cost of changing start times.

2. Why might start times be important?

One reason that schools begin the school day at different times is the use of tiered busing systems. Many school districts stagger the times their schools begin the school day so that the same bus and driver can serve multiple schools. For example, a school district might start a high school at 7:30, a middle school at 8:15 and an elementary school at 9:00. Staggering start times in this manner has the potential to drastically reduce transportation costs compared to an alternative schedule where all three schools start at 8:15.

Only one nationally representative dataset records school start times: The 2001 Before- and After-School Programs and Activities section of the National Household Education Survey (ASPA-NHES), conducted by the National Center for Educational Statistics. This dataset does not contain any information on academic achievement, so it cannot be used to measure the impact of start times on student performance. However, it is useful in establishing the variation of start times across the nation and for comparing the data used in this paper to national norms. Table 1 lists summary statistics for middle school start times for the national sample in 2001. Nationally, the median middle school student begins school at 8:00. Over one fourth of students begin school at 8:30 or later, while more than 20 percent begin at 7:45 or earlier.

In addition to the national sample, Table 1 also lists statistics for middle school start times in Wake County from 1999 to 2006. As would be expected when comparing a specific district to the national distribution, start times in Wake County are more concentrated than they are nationwide. In Wake County, 53.1 percent of middle school students start school at 7:30, and another 22.4 percent begin at 8:15. In the national sample, only 26.7 percent of students start at the national modal time of 8:00. More importantly, start times are consistently earlier in Wake County than nationwide: the median Wake County student begins school earlier than over 90 percent of students nationwide. Since the marginal impact of start times on academic performance may decrease (or increase) for later start times, care should be taken in imputing the conclusions reached in this paper to schools that have a later start time. Put another way, the gain from later start times found here derive largely from changes from 7:30 to 8:15. The same gains may not occur from changing start times from 8:30 to 9:15, for example.

Another source of information on nationwide start times is a survey of high school start times by Wolfson and Carskadon (2005). They randomly selected 4116 schools and asked them to report start times retrospectively since 1965. While the survey deals with high schools instead of middle schools and has response problems (fewer than ten percent of schools selected are included in their final sample), it does provide useful information about the types of schools that begin early or late. They find that large schools (>1000 students) begin on average 15 min earlier than small (<1000 students) schools, and that rural schools begin 15 min later than urban or suburban schools. Schools that are not part of a tiered busing system begin on average 15 min later than schools with a two-tiered system and 20 min later than schools using a three-tiered system. Their findings are consistent with the WCPSS being an urban/suburban school district with a three-tiered busing system and having somewhat earlier start times than the national sample.

There is credible evidence (Baroni et al., 2004; Danner & Phillips, 2008) that students who begin the school day earlier sleep less. To obtain the same amount of sleep as students who have a later start time, early-starting students would need to go to sleep earlier. Activities such as sports, work, family and social schedules may make it difficult for students to adjust the time they go to bed. In addition, the onset of puberty brings two factors that can make this adjustment particularly difficult for adolescents: an increase in the needed amount of sleep and a

Table 1
Nationwide and wake county student start time statistics.

	Percentiles				
	10%	25%	50%	75%	90%
Nationwide	7:35	7:55	8:00	8:30	8:45
Wake County, NC	7:30	7:30	7:30	8:15	8:15

	S. Dev. (min)	Mode	Percentage at mode	Sample size
Nationwide	27.7	8:00	26.7%	4568
Wake County, NC	20.4	7:30	53.1%	173,791

Students are in grade 6–8. Nationwide data comes from the ASPA-NHES (2001). Wake County data is pooled data from 1999 to 2006.

change in the natural timing of the sleep cycle. Hormonal changes, in particular the secretion of melatonin, shift the natural circadian rhythm of adolescents, making it increasingly difficult to fall asleep early in the evening. It is well established in the physiological literature that less sleep is associated with a decrease in cognitive performance, both in a laboratory settings and through self-reported sleep habits (Pilcher & Huffcutt, 1996).[2] Specific to academic achievement, numerous studies have reported a negative correlation between self-reported hours of sleep and grades among both middle and high school students.[3]

Since students who start school earlier typically sleep less, and less sleep is associated with decreased academic performance, one would expect that students in early-starting schools would perform worse on standardized tests. However, there is little empirical evidence directly linking school start times and academic performance. Allen (1992) and Wolfson, Spaulding, Dandrow, and Baroni (2007) examine a small number of schools and find a positive correlation between later start times and student grades. This approach is inherently limited; any increase in performance could reflect other unobserved factors rather than the impact of a later start time.

The study most widely cited in the popular press (Wahlstrom, 2000) examined two Minnesota school districts (Minneapolis Public Schools and Edina) which changed the start times in their secondary schools to start over an hour later.[4] Wahlstrom is concerned with the impact of start times on a wide range factors, including attendance, sleep behavior, school discipline and extra-curricular activities, in addition to academic performance. The impact of a delayed start time is measured by comparing the mean grade obtained by high school students during the three years prior to the start time change and mean grades from three years after the change. Wahlstrom finds a positive but statistically insignificant increase in mean grades.

Recent papers by Hinrichs (2011), Carrell, Maghakian, and West (2011), and Wong (2011) avoid many of the issues of extant studies. Hinrichs uses individual ACT data on students in the Minneapolis metro area for the same policy change as Wahlstrom. However, he includes additional school districts, allowing him to control for secular time trends. Hinrichs also uses school-level data on Kansas assessment tests in the 10th and 11th grades. Using a variety of specifications, he finds no effect of later start times in any of his specifications. Carrell et al. (2011) examine how the time of the first class of the day affects college freshman. They use data from the United States Air Force Academy, where freshman are randomly assigned to class periods.[5] They find that a 1 h delay in the first class of the day increases grades by 0.15 standard deviations. Wong (2011) uses cross sectional data and finds effects ranging from 0.02 to 0.08 standard deviations for a 1 h later start time, although his results generally lack statistical precision.

A related literature examines how time of day affects student performance. Dills and Hernandez-Julian (2008) find that college students receive higher grades in classes that meet later in the day. Cortes, Bricker, and Rohlfs (2012) examine students in Chicago public high schools and find that grades are reduced in first-period classes.

This paper complements existing studies in several ways. First, I am better able to control for unobservable factors by using multiple sources of variation in start times, both across schools and within schools over time. Second, I examine students in middle school (6th–8th grades), while existing studies tend to examine high school students, often in a single grade. If adolescent hormones determine how start times affect academic performance, middle school students may respond differently than high school students. Third, I use standardized test scores instead of letter grades or grade point average. Fourth, I examine changes in start times that occurred in different years for different schools, and are not intended to increase student achievement. This feature of the data makes Hawthorne effects[6] unlikely and will allow me to separate out year effects. Lastly, my unique data set allows me to examine several specific mechanisms through which start times may affect academic performance, in particular whether

[2] Laboratory studies tend to focus on large amounts of sleep loss. The amount of sleep lost from starting school earlier would be much less. Laboratory results should be viewed as establishing a credible relationship between sleep and cognitive performance.

[3] See Wolfson and Carskadon (2003) for a survey of these studies.

[4] Edina school district changed their start times from 7:25 to 8:30 and Minneapolis changed from 7:15 to 8:40.

[5] Some students have a first period class and others do not, but all students must attend mandatory breakfast prior to the first class period.

[6] A Hawthorne effect occurs when subjects modify their behavior because they are being studied, and not because of a change in an explanatory variable.

the “adolescent hormone” hypothesis explains the effect of start times on academic performance.

3. Data and institutional background

The data set used in this paper is combined from two sources. The first source is administrative data for every student in North Carolina between 1999 and 2006.[7] The data contain detailed demographic variables for each student as well as end of grade test scores in both reading and math.[8] The raw test scores vary considerably by year, so I use the statewide data to construct percentile scores for each student within their grade and the current year.[9] The second source of data is the start time for each school in the Wake County Public School System (WCPSS), by year.[10] Start times were matched to the school code for each student in the administrative data.

There are several factors that make these data well suited for examining the effect of start times on academic performance. Examining a single school district avoids problems resulting from correlation of start times with unobserved characteristics at the district level. Since start times are generally determined at the district level, it may be that well-run districts tend to have earlier (or later) start times, or they may be more (or less) likely to change the start times of their schools. The WCPSS is also large enough to measure effects precisely. It has considerable variation in start times both across schools and within schools over time. By focusing on middle school students, I am able to test the “adolescent hormones” hypothesis of why later start times affect test scores. The WCPSS is the eighteenth largest public school district in the United States with 120,504 students (kindergarten through twelfth grade) in the 2005–06 school year. It encompasses all public schools in Wake County, a mostly urban and suburban county that includes the cities of Raleigh and Wake Forest. Start times for schools in the district are proposed by the transportation department (which also determines bus schedules) and are approved by the school board. While the school board could in theory make changes to the bell schedule, they did not do so during the sample period of 1999–2006. The practice of combining school start times and bus scheduling is supported by the Operations Research literature concerning the “school bus problem,” where it is well established that costs can be reduced up to 30% if start times and bus routes are chosen simultaneously (Fugenschuh, 2009; Keller & Muller, 1979).

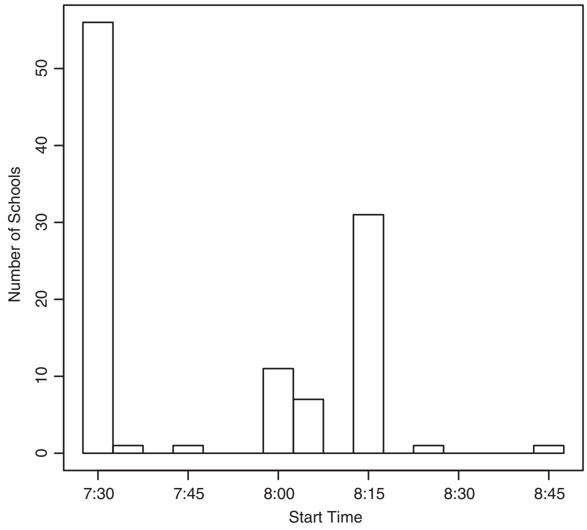

Fig. 1. Histogram of middle school start times.
Source: Author’s calculations from data provided by the WCPSS transportation department. Figure includes all non-magnet middle schools in Wake County, NC from 1999 to 2006. Each school appears once per school year.

Since 1995, WCPSS has operated under a three-tiered system. While there is some variation in the exact start times, most Tier I schools begin at 7:30, Tier II schools at 8:15 and Tier III at 9:15. Tiers I and II are composed primarily of middle and high schools, and Tier III is composed entirely of elementary schools. Fig. 1 is a histogram of the distribution of start times of middle schools for the pooled sample.[11] Just over half of middle schools begin at 7:30, with substantial numbers of schools beginning at 8:00 and 8:15 as well. All of the schools have the same length of the school day.

Wake County is divided into 1134 student assignment nodes. Each node is associated with an elementary, middle and high school that is the “base” school for all students living within the node. Nodes are matched to schools based on facility utilization, distance and diversity. Each year, as the population changes and new schools are built, a limited number of nodes are reassigned to new schools. Nodes remain with a given school for a minimum of three years between reassignments. A majority (72.5% in 2006) of students attend their base school. Alternatively, students may choose to apply to attend a magnet school.[12]

[7] The administrative data were provided by the North Carolina Education Research Center.

[8] The data also contain information on teachers, potentially allowing me to control for teacher and classroom characteristics. However, I am unable to link students to the teacher data set for roughly 28% of my sample. In addition, the teacher that is linked is the teacher who supervised the exam, which may not be the teacher the student had for the school year. I construct a “class size” index which is the number of students coded with the same teacher, class period, and year. For 15% of students matched to teachers, the “class size” was over 50, suggesting for many students the teacher linked to their record was not their true teacher, but rather a supervisor of a test for multiple classes. As a result, I do not include teacher or class characteristics in the results shown. Results are generally similar when teacher characteristics are included.

[9] Specifically, I pool the raw test scores for all students in North Carolina in a given year and grade. I then assign each student in Wake County a percentile rank based on where their raw test score falls in the statewide distribution. Percentile scores were constructed separately for math and reading. An alternate normalization would be to construct Z-scores by subtracting the statewide mean and dividing by the statewide standard deviation. Results using Z-scores are presented in Table 7 and are very similar to those for percentile ranks.

[10] The start time data were provided by the WCPSS transportation department.

[11] Most schools are represented seven times (once for each year) although sometimes with different start times.

[12] A small number (<5%) of students transfer to a base school other than the one they are assigned. Such transfers are only allowed for exceptional

Table 2
Changes in start times.

		New start time							
		7:30	7:35	7:45	8:00	8:05	8:15	8:25	8:45
Old start time	7:30	–	–	–	1	–	2	–	–
	7:35	–	–	–	–	–	1	–	–
	7:45	–	–	–	–	–	–	–	–
	8:00	1	–	–	–	–	4	1	–
	8:05	–	–	–	–	–	–	–	–
	8:15	2	–	–	–	–	–	–	–
	8:25	–	–	–	1	–	–	–	–
	8:45	–	–	–	–	–	1	–	–

Source: Author's tabulation.
Each cell contains the number of schools that changed their start time from the given old start time to the new start time. In addition to these changes eleven schools started at the same time in each sample year, of these eight started at 7:30 and one each at 7:45, 8:00, and 8:05.

Magnet schools use a specialized curriculum and typically have smaller enrollments. Magnet school admittance is determined by lottery.[13] Bus transportation is provided to students whether they attend the base school or a magnet school. Since buses serving magnet school must cover a larger geographic area, ride times tend to be longer for magnet school students. As a result, almost all magnet schools begin at the earliest start time. For example in 2004, nine out of ten magnet schools began at 7:45 or earlier compared with nine out of sixteen base schools. Students at magnet schools tend to have higher test scores, which may cause a spurious negative relationship between start times and test scores. Furthermore, since students can choose to apply to magnet schools, it is possible that they chose a magnet school partially based on start time. For these reasons, I exclude magnet schools from my sample. Five schools began magnet programs during the sample period. These schools are included in the sample prior to becoming magnet schools and excluded after. Results including magnet schools are presented in Table 7. Results including magnet schools are similar to the main results, but with slightly lower magnitudes.

Over the seven years examined in this paper, WCPSS grew from 20,530 students enrolled in twenty-two middle schools during the 1999–2000 school year to 27,686 students enrolled in twenty-eight middle schools in 2005–2006. In addition to population growth, the WCPSS increased the number of magnet schools from five to thirteen by opening three new magnet schools and converting five existing non-magnet schools into magnet schools. Three new non-magnet schools were also opened. As a result of population growth, the transportation department changed start times to accommodate new schools and increased enrollment at existing schools. An explanation of the need to change start times is given in the minutes of the Board of Education:

> The proposed bell schedule modifications open and close schools at times the Transportation Department can more effectively and efficiently provide bus service to students. Advantages of the bell schedule modifications include long-term savings in operating costs, better utilization of equipment, and the ability to provide better service. It is important to note that new school openings will dictate a need to change opening and closing times. (Wake County Board of Education, 2004)

Existing non-magnet schools had their start times change fourteen times in the sample, with some schools starting earlier and others later. In the portion of my analysis that uses school or student fixed effects, it is these changes that identify the effect of later start times. Table 2 tabulates the changes in start times. Each cell contains the number of schools that switched from the corresponding old start time to the corresponding new start time. Four schools changed their start time two or three times. Those schools appear more than once in the table. There were a total of fourteen changes in start times by nine schools. Seven of those changes were of 30 min or more. Eleven schools did not change their start time in the sample period, and so do not appear in the table. Of the eleven, eight started at 7:30 and one each at 7:45, 8:00, and 8:05.

Table 3 lists means of selected demographic variables by start time for 2000 and for 2006. Panel A includes student characteristics, while Panel B includes (unweighted) school characteristics.[14] For the purposes of tabulation, I separate start times into two groups: Tier I (7:30–7:45) and Tier II (8:00–8:45). Buses that serve Tier I schools would have time to serve two additional schools, but buses that serve Tier II schools would only have time for one more school. For almost all of the student demographics, there are differences between earlier and later starting schools in 2000. The differences are generally small in magnitude, but precisely measured. For example, 24% of students in Tier I schools are black, compared to 21% of Tier II students. Students in earlier starting schools are more likely to be female, belong to an ethnic minority, be eligible for free lunches and have less educated parents. In 2006, the demographic characteristics of the two tiers are generally closer together, but several differences still remain.

circumstances and for grandfathered students who live in a reassigned node. Bus transportation is not provided for transfer students. I do not have information on which students have transferred to a different base school than the one to which they are assigned.

[13] Preference within the lottery is given to some students based on sibling enrollment, distance and diversity.

[14] Tabulations for other years in the sample are similar to 2000 and 2006, but are not presented.

Table 3
Means by start time for 2000 & 2006.

	2000		2006	
	Tier I	Tier II	Tier I	Tier II
Panel A: Student characteristics				
Female	0.495	0.481	0.494	0.490
	(0.005)	(0.006)	(0.006)	(0.006)
Age	13.545	13.544	13.630	13.620
	(0.009)	(0.012)	(0.013)	(0.013)
Black	0.240	0.207	0.247	0.249
	(0.004)	(0.005)	(0.005)	(0.005)
Hispanic	0.0516	0.0253	0.102	0.0721
	(0.002)	(0.003)	(0.003)	(0.003)
Free lunch eligible	0.237	0.140	0.284	0.243
	(0.004)	(0.005)	(0.005)	(0.005)
Parent's education	14.86	15.00	14.45	14.81
	(0.02)	(0.03)	(0.03)	(0.04)
Math score	55.92	59.63	56.29	61.41
	(0.285)	(0.37)	(0.347)	(0.336)
Reading score	56.5	59.49	54.65	60.1
	(0.278)	(0.362)	(0.336)	(0.325)
Number of students	10,544	6082	7191	7675
Panel B: School characteristics				
% Black	0.256	0.237	0.292	0.289
	(0.034)	(0.046)	(0.050)	(0.054)
% Hispanic	0.0443	0.0261	0.103	0.0714
	(0.006)	(0.009)	(0.016)	(0.017)
% Free lunch eligible	0.239	0.156	0.294	0.250
	(0.060)	(0.080)	(0.051)	(0.055)
Avg. parent education	14.90	14.89	14.61	14.68
	(0.238)	(0.322)	(0.364)	(0.420)
Pupil/teacher ratio	15.65	14.80	14.91	15.66
	(0.47)	(0.63)	(0.38)	(0.40)
Enrollment	958.5	1013.7	898.9	1096.4
	(87.6)	(118.5)	(49.0)	(52.4)
Avg. math score	56.14	58.48	55.88	61.03
	(2.59)	(3.51)	(4.22)	(4.51)
Avg. reading score	56.78	58.23	54.22	59.71
	(2.45)	(3.31)	(3.01)	(3.22)
Start time	7:32	8:01	7:30	8:14
	(0:01)	(0:02)	(0:01)	(0:01)
Schools	11	6	8	7

Tier I: 7:30–7:45. Tier II: 8:00–8:45. Standard deviations in parentheses. Panel B is unweighted.

Since the characteristics more prevalent among Tier I students are generally believed to be associated with lower test scores, the simple correlation between test scores is unlikely to represent a causal effect. Moreover, if students in early-starting schools have other characteristics that are correlated with test scores, controlling for the observed characteristics may not be sufficient to give an unbiased estimate of the effect of start times. The school characteristics in each tier are much more similar to each other than the student characteristics are.[15] In particular, schools in both tiers have similar enrollments and pupil to teacher ratios. For none of the variables are the differences between the two tiers statistically significant at conventional levels. Even for variables that are statistically different in panel A, the corresponding percentages in Panel B are not statistically different.

For the portion of my analysis that identifies the effect of later start times using only the variation in start times within schools, it is important to know the characteristics of schools that experience a change in start time. In particular, students in schools that change to an earlier start time may be different than students in schools that change to a later start time. Table 4 gives the means of selected demographic characteristics for students in schools that experienced a change in start time. Columns 1–3 give results for schools that changed from a later start time to an earlier one. Column 1 is the year before the change, column 2 the first year with the new start time, and column 3 the difference between columns 1 and 2. Columns 4–6 give results for schools that changed from an earlier start time to a later one. Column 4 is the year before the change, column 5 the first year with the new start time, and column 6 the difference between columns 4 and 5. Lastly, column 7 gives the difference between columns one and two. For both schools moving to an earlier start time and those moving to a later start time, students are similar before and after the change in terms of gender, race, eligibility for free lunch and parent's education. Students in schools that change to an earlier start time have higher test scores before the change, and students in schools that change to a later start time have higher test scores after the change, although the differences are not statistically significant. Students in schools that change to an earlier start time have similar characteristics to students in schools that change to a later start time in the year prior to the change. A comparison of Tables 3 and 4 indicates that schools which change start times differ from those that do not change start times. In particular, students in schools which change start times are less likely to be black, Hispanic or eligible for free lunch.

4. The relationship between start time and performance

The unconditional relationship between test scores and start times can be seen in Figs. 2 and 3. These figures present the empirical cumulative distribution function of test scores in schools with earlier (Tier I) and later (Tier II) start times. For both the reading and math tests, the cumulative distribution for late-starting schools first-order stochastically dominates the distribution for early-starting schools; for every percentile a greater proportion of students score at or below that percentile in the early-starting schools than in later starting schools.[16] For example, 45% of students in early-starting schools have math test scores at or below the 50th percentile, while only 36% of students at late-starting schools score at or below the 50th percentile.

These figures suggest that later start times positively impact student performance. Since there are other differences between early- and late-starting schools, the simple relationship between test scores and start times should not

[15] The differences in the percentage characteristics in panel B and the corresponding variables in panel A is mainly a result of Panel B being unweighted.

[16] In an alternate specification (not shown), I restricted the estimates to schools that started at either 7:30 or 8:15 (roughly 75% of the sample). Results were highly similar to those in Figs. 2 and 3.

Table 4
Summary statistics for students in schools which change start times.

	Schools changing to earlier start			Schools changing to later start			
	(1) Before	(2) After	(3) Difference (1)–(2)	(4) Before	(5) After	(6) Difference (4)–(5)	(7) Difference (1)–(4)
Female	0.486	0.497	−0.011	0.493	0.495	−0.002	−0.007
	(0.007)	(0.007)	(0.010)	(0.005)	(0.005)	(0.008)	(0.009)
Black	0.218	0.226	−0.008	0.221	0.219	0.002	−0.003
	(0.006)	(0.006)	(0.009)	(0.004)	(0.004)	(0.006)	(0.008)
Hispanic	0.034	0.038	−0.004	0.042	0.045	−0.003	−0.008
	(0.006)	(0.005)	(0.007)	(0.005)	(0.005)	(0.007)	(0.007)
Free lunch eligible	0.188	0.194	−0.006	0.182	0.184	−0.002	0.006
	(0.006)	(0.006)	(0.008)	(0.004)	(0.004)	(0.006)	(0.007)
Parent's education	14.900	14.840	0.060	14.930	15.130	−0.200	−0.030
	(0.204)	(0.294)	(0.358)	(0.196)	(0.212)	(0.289)	(0.283)
Math score	57.540	56.610	0.930	58.140	59.540	−1.400	−0.600
	(0.929)	(0.879)	(1.279)	(0.673)	(0.656)	(0.940)	(1.147)
Reading score	58.120	57.420	0.700	58.580	59.860	−1.280	−0.460
	(0.923)	(0.872)	(1.270)	(0.669)	(0.651)	(0.933)	(1.140)
N	4647	5224	−577	8684	8982	−298	–
School-years	5	5	–	9	9	–	–

be viewed as representing the impact of later start times. I estimate several regression models to account for observed characteristics. Without any fixed-effects, the specifications have the basic form:

$$y_{ijt} = \beta START_{jt} + X'_{ijt}\gamma + Z'_{jt}\delta + \epsilon_{ijt} \quad (1)$$

Here y_{ijt} is the test score of student *i* who attends school *j* in year *t*. $START_{jt}$ is the start time of school *j* in year *t*, measured in hours after midnight. X_{ijt} is a vector of individual characteristics (not all of which vary over time) including gender, race, age, economic status, grade and parental education. Z_{jt} is a vector of school characteristics including pupil-to-teacher ratio, racial and economic composition, and ϵ_{ijt} is the error term. In order to isolate different sources of variation, I use different fixed effects in some specifications. Most specifications include a year specific effect. Some specifications use a school fixed effect, using only variation in start times within the same school, over time. A student-school fixed effect uses only changes in start times for students who stay in the same school.

Table 5 presents results of nested models of math (Panel A) and reading (Panel B) test scores. These specifications use both variation in start times across and within schools. Each column adds additional explanatory variables to the previous specification. Column 1 gives a simple regression of test scores on start times, without any additional covariates. For both math and reading, the coefficient on start time is quite large, 9.5 percentile points per hour for both

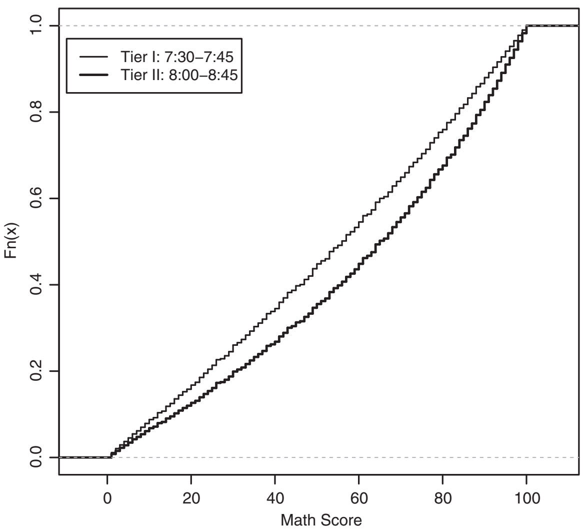

Fig. 2. Empirical cumulative distribution of math test scores by start time tier. *Note*: Data include all middle school students in non-magnet schools in the Wake County Public School System from 1999 to 2006. Test score is the percentile rank on the end of grade math test administered to all students in North Carolina.

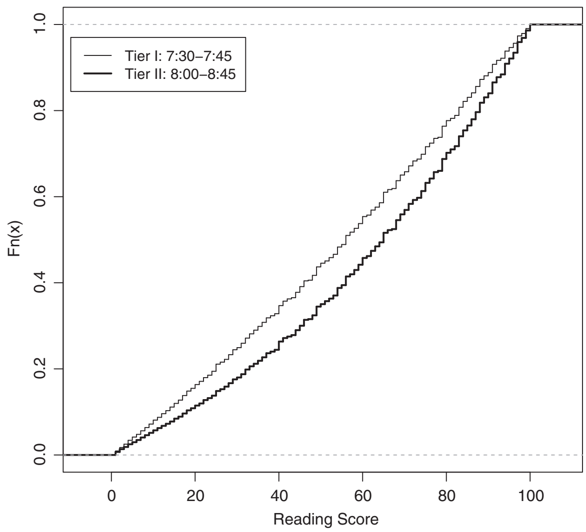

Fig. 3. Empirical cumulative distribution of reading test scores by start time tier. *Note*: Data includes all middle school students in non-magnet schools in the Wake County Public School System from 1999 to 2006. Test score is the percentile rank on the end of grade reading test administered to all students in North Carolina.

Table 5
Pooled OLS models.

	(1)	(2)	(3)	(4)
Panel A: Math test scores				
Start time	9.472***	4.484**	4.902*	2.896*
	(2.426)	(1.667)	(1.928)	(1.318)
Individual controls	No	Yes	Yes	Yes
School Controls	No	No	Yes	Yes
Year effect	No	No	No	Yes
Grade effect	No	No	No	Yes
Observations	102,506	102,506	97,877	97,877
R^2	0.013	0.452	0.459	0.487
Panel B: Reading test scores				
Start time	9.451***	4.807***	4.975***	3.357***
	(2.051)	(1.293)	(1.457)	(0.876)
Individual controls	No	Yes	Yes	Yes
School controls	No	No	Yes	Yes
Year effect	No	No	No	Yes
Grade effect	No	No	No	Yes
Observations	102,250	102,250	97,628	97,628
R^2	0.014	0.421	0.426	0.445

Data is from WCPSS grades 6–8 during the 1999–2006 school years. Dependent variable is score on end of grade math and reading test. Individual controls include gender, race, eligibility for free/reduced price lunch, academically gifted, learning disability, and limited English status, and years of parental education. School controls include total enrollment, pupil/teacher ratio, racial composition, percentage of students eligible for free lunch and percentage of returning students. Standard errors (robust to clustering at the school level) in parentheses.

* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

tests.[17] The effect drops substantially as I add additional covariates. Column 2 adds individual characteristics (the X_{ijt}). The coefficient on start time drops to 4.5 for math and 4.8 for reading. This result is not surprising since students at early-starting schools are more likely to belong to an ethnic minority, be eligible for free lunches and have less educated parents. The large effect found in column 1 reflects in part these differences. When school characteristics are added to the specification in column 3, the effect changes slightly to 4.9 for math and 5.0 for reading. Column 4 adds year and grade effects and obtains a result of 2.9 percentile points for math and 3.6 percentile points for reading. The start time coefficient decreases slightly as I add school characteristics, year and grade effects, but the drop is much less than from column 1 to column 2. In all cases the effect is statistically different from zero at the 1% level of significance.

As noted above, there are differences in observed demographic characteristics of students in early and late starting schools. If there are also unobserved differences that are correlated with start times, the results found in Table 5 may be biased estimates of the true effect of start times on academic performance. To avoid this potential bias, I estimate the effect of a later start time using only variation in start time within schools over time. Fig. 4 demonstrates the effect of later start times graphically. For schools that experience a change in start time, I plot the change in start time versus the changes in math test score, reading test score, total enrollment, and the percentage of students who were enrolled in the same school last year. A (unweighted) fitted line is plotted for each graph. A clear relationship between changes in start times and changes in test scores can be seen: a 1 h later start time is associated with an increase in average test scores of more than two percentile points on both math and reading tests. For both tests, the slope of the fitted line is significantly different from zero at the 10% confidence level. In contrast there is no significant relationship between changes in start times and changes in either per-grade enrollment or the percentage of students returning from last year.

In order to control for additional covariates, I estimate specifications of Eq. (1) that use school and student-school fixed effects. The results are presented in Table 6. The advantage of these specifications is that any unobserved characteristics that do not change over time will be captured by the relevant fixed effect. By disregarding the variation in start times across schools and identifying the effect of later start times using only the variation in start time within a given school, my estimates are less precise. Columns 1 and 2 in Table 6 estimate specifications using school fixed effects for the math and reading test, respectively. In this case the effect of a later start time is identified using only the variation in start times within schools over time. The effect of a 1 h later start time is a 2.2 percentile point increase in math test scores and a 1.6 percentile point increase in reading test scores. The school fixed effect controls for all school level characteristics that do not change over time. However, a remaining concern is that the student composition of schools may change over time. For example, high achieving students in a school which changed to an earlier start time might transfer to private schools, resulting in a spurious relationship between start times and test scores. To address this issue, columns 3 and 4 use student-school fixed effects, identifying the impact of later start times using only students who experience a change in start time while remaining in the same school. The effect of a 1 h later start time is 1.8 and 1.0 percentile points for math and reading, respectively. The results found using variation within students are generally lower than those found using both variation within and across schools, but are still significantly different from zero. The results for the reading test are lower than those from the math test.[18] The coefficients on the additional control variables put the magnitude of the effect of later start times in context. For math scores, the effect of a 1 h later start time is roughly 14% of the black–white test score gap, 40% of the gap between those eligible and those not eligible for free or reduced price lunch and 85% of the gain associated with an additional year of parental education.

To investigate how the effect of later start times varies across the distribution of test scores, I estimate the model given in equation 1 by quantile regression for each decile

[17] Throughout the paper, I measure differences in start times in hours to ease interpretation. The difference between early and late starting schools in the data is typically 45 min.

[18] A possible explanation for the larger effect on math scores compared to reading scores is that math skills may be dependent on the school environment than reading skills.

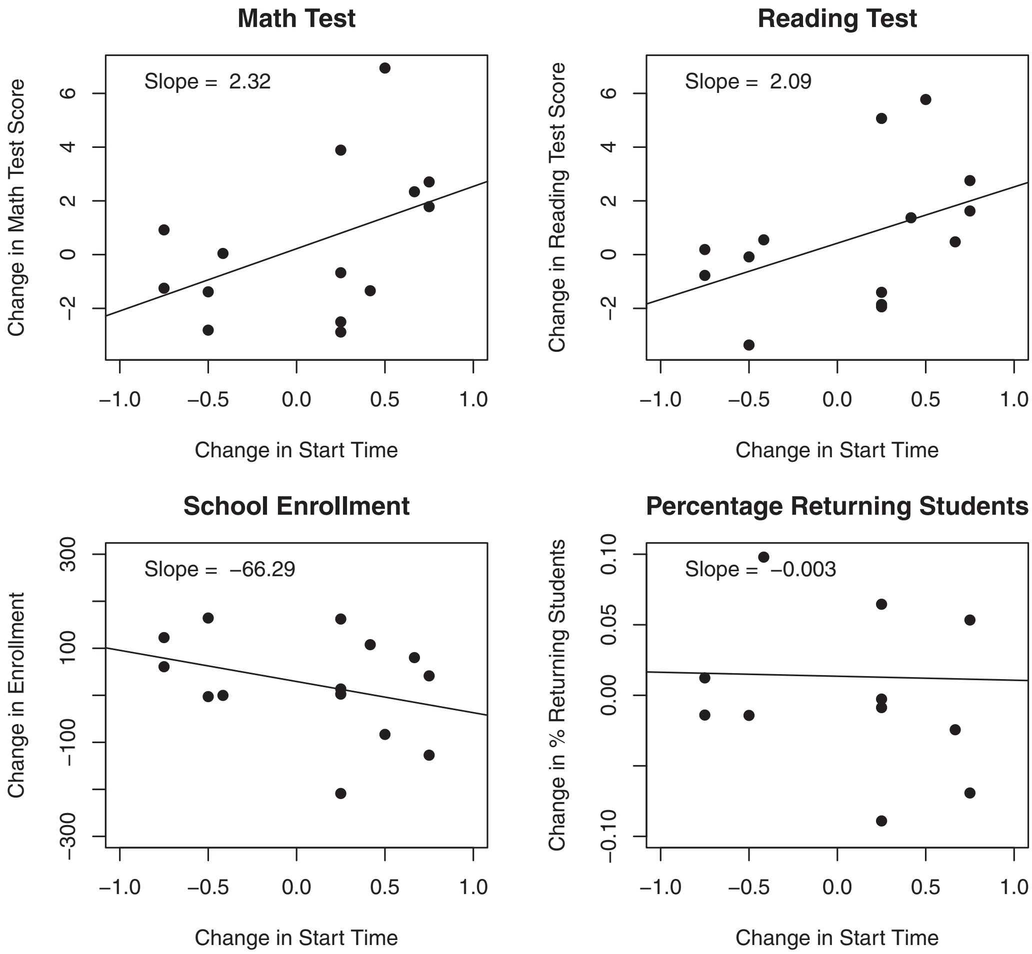

Fig. 4. Changes in start times vs. changes in other variables.

of the distribution. I include a full set of explanatory variables, as well as year and grade dummy variables and school fixed effects. Since the conditional quantile function is not a linear operator, the interpretation and estimation of fixed effects is different in quantile regression than in OLS regression. As a result there are several methods of including fixed effects in a quantile regression model. I employ the using the two step estimator proposed in Canay (2011), which models fixed effects as pure location shifts.[19] Figs. 5 and 6 present the conditional quantile effect of a 1 h later start time for the math and reading tests graphically. The shaded area represents 95% confidence intervals for the point estimates.[20] The solid horizontal line is the corresponding OLS coefficient (columns 1 and 3 of Table 6) and the dotted horizontal lines are the OLS 95% confidence interval. For both tests, the effect of later start times is much greater in the bottom half of the distribution, and is decreasing after the third decile. These results indicate that start times have a greater impact on the bottom half of students, which is one possible explanation as to why my results differ from those found in Hinrichs (2011). Hinrichs uses ACT test scores for students in the Minneapolis-St. Paul, MN metro area. For his sample years, between 59 and 66 percent of high school students in Minnesota took the ACT. If students in the bottom half of the distribution of test scores are less likely to take the ACT, the sample used in Hinrichs (2011) will under represent the portion of the population where I find the largest impact of start times. These results are consistent with the results of Cortes et al. (2012) who find that the effects of first period classes are particularly large for black students, but contrast with the results of Corak and Lauzon (2009) who find that smaller class sizes would increase the proportion of students performing in the lowest reading proficiency level. The larger effect of start times on the lower end of the grade distribution also suggests that delaying school start times may be an especially relevant policy change for school districts trying to meet minimum competency requirements (such as those mandated in the No Child Left Behind Act).

Table 7 gives estimation results for alternative samples and test score normalizations. Each entry is the coefficient

[19] In results not shown, I estimate a quantile regression model without school fixed effects. The point estimates of conditional quantile effects in this model have the same shape (decreasing after the third decile) as the results shown, but are generally larger.

[20] Standard errors are estimated using the Huber sandwich procedure described in Koenker (2005).

Table 6
School and student-school fixed effect models.

	(1) Math	(2) Reading	(3) Math	(4) Reading
Start time	2.236***	1.599***	1.778***	0.979**
	(0.436)	(0.445)	(0.300)	(0.357)
Female	−1.713***	2.534***	–	–
	(0.133)	(0.136)		
Age	−4.706***	−3.508***	−0.107	0.123
	(0.120)	(0.123)	(0.266)	(0.325)
Black	−15.80***	−15.29***	–	–
	(0.186)	(0.190)		
Hispanic	−5.441***	−5.794***	–	–
	(0.337)	(0.345)		
AG	23.28***	19.80***	1.128***	0.608
	(0.185)	(0.189)	(0.301)	(0.358)
Learning disability	−13.48***	−12.70***	0.417	0.237
	(0.194)	(0.198)	(0.264)	(0.315)
Limited English	−12.47***	−19.98***	−1.467	−1.104
	(0.472)	(0.496)	(0.818)	(0.977)
Parent education	2.638***	2.494***	−0.0299	0.101*
	(0.0424)	(0.0433)	(0.0360)	(0.0428)
Free lunch	−5.429***	−5.904***	0.0494	0.0763
	(0.195)	(0.200)	(0.232)	(0.277)
Pupil/teacher ratio	−0.282***	−0.129*	−0.0723	0.0465
	(0.0595)	(0.0607)	(0.0406)	(0.0483)
Enrollment	−0.00118	0.000113	−0.000422	−0.0000139
	(0.000994)	(0.00101)	(0.000729)	(0.000867)
% Returning students	8.994***	5.575**	6.083***	−4.764**
	(2.118)	(2.162)	(1.348)	(1.604)
% Black	9.033**	2.137	5.211*	−3.877
	(3.431)	(3.502)	(2.180)	(2.594)
% Hispanic	−38.47***	−21.49***	10.49*	−6.403
	(4.918)	(5.021)	(4.430)	(5.271)
% Free lunch	−3.803	1.814	−0.734	8.451***
	(2.041)	(2.080)	(1.479)	(1.760)
Year effect	Yes	Yes	Yes	Yes
Grade effect	Yes	Yes	Yes	Yes
Fixed effect	School	School	Student-school	Student-school
Observations	100,680	100,680	100,427	100,427
Adjusted R^2	0.37	0.89	0.33	0.83

Dependent variable is score on end of grade math or reading exam. All specifications include a constant term. Additional ethnic controls not reported. Standard errors (robust to clustering at the school level) in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

Table 7
Results for alternate specifications.

	Math	Math	Reading	Reading	Sample size
Full specification	2.236***	1.778***	1.599***	0.979**	97,877
	(0.436)	(0.300)	(0.445)	(0.357)	
Magnet included	1.300***	1.394***	1.134**	0.846**	162,449
	(0.371)	(0.247)	(0.379)	(0.296)	
Z Scores	0.0707***	0.0570***	0.0493**	0.0352**	97,877
	(0.0146)	(0.0103)	(0.0156)	(0.0126)	
School-specific time trend	1.703	1.224*	1.819	0.319	97,877
	(1.558)	(0.684)	(1.208)	(1.361)	
Start time & two-year lead	1.487**	1.952***	1.003*	0.994**	68,930
	(0.457)	(0.310)	(0.468)	(0.369)	
Fixed effect	School	Student-school	School	Student-school	

All columns include a full set of control variables, including year and grade effects. Sample size is for math test, sample size for reading test is typically slightly smaller. Standard errors (robust to clustering at the school level) in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

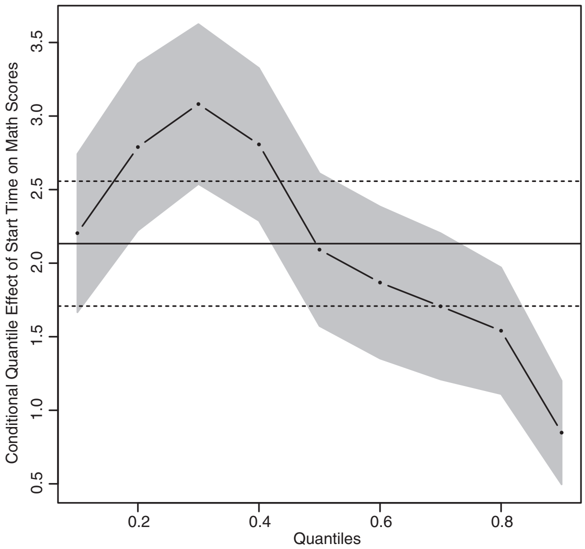

Fig. 5. Quantile regression of math test score on start time. *Note*: Each point represents the conditional quantile effect of a 1 h later start time on the percentile rank on the end of grade math exam. The shaded area represents 95% confidence intervals. The solid horizontal line is the corresponding OLS estimate, and the dotted horizontal lines bound a 95% confidence interval of the OLS estimate. See text for additional details.

on start time. Columns one and two show results for the math test, while columns three and four show results for the reading test. Columns one and three use both variation within and across schools, while columns two and four use only variation within schools for given students (a student-school fixed effect). The first row reproduces the results found in Table 6 for comparison purposes. The second row expands the sample to include magnet schools. Results are

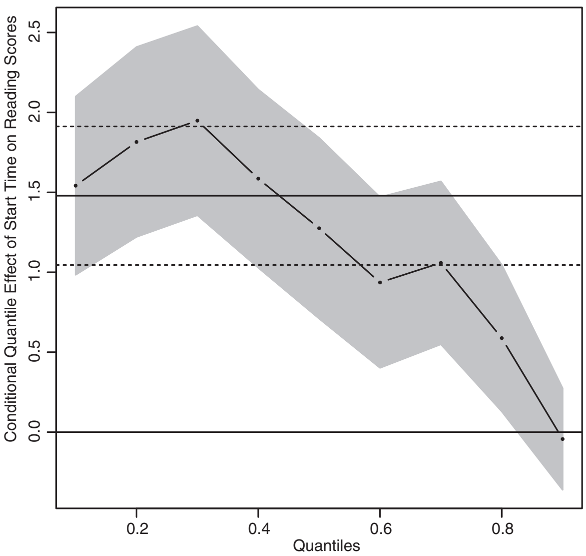

Fig. 6. Quantile regression of reading test score on start time. *Note*: Each point represents the conditional quantile effect of a 1 h later start time on the percentile rank on the end of grade reading exam. The shaded area represents 95% confidence intervals. The solid horizontal line is the corresponding OLS estimate, and the dotted horizontal lines bound a 95% confidence interval of the OLS estimate. See text for additional details.

similar, but slightly smaller than those found excluding magnet schools. Since students at magnet schools tend to have higher test scores, the quantile regression results found above may explain these differences. The third row uses an alternate normalization of test scores. Instead of constructing percentiles, I normalize test scores by subtracting the statewide mean and dividing by the statewide standard deviation for a given grade and year. Results are very similar to those using percentiles.[21] Row 4 includes school-specific time trends as additional explanatory variables. Results are generally similar to those without school-specific time trends but are imprecisely measured. Lastly, row five includes a two year lead of the school start time as an additional explanatory variable to ensure that the measured effect of later start times is not picking up a relationship with an unrelated, slow moving time trend. Results are similar to the main results.

5. Why do start times matter?

The typical explanation as to why later start times might increase academic achievement is that by starting school later, students will get more sleep. As students enter adolescence, hormonal changes make it difficult for them to compensate for early school start times by going to bed earlier. Since students enter adolescence during their middle school years, examining the effect of start times as students age provides a test of a novel fact predicted by the adolescent hormone explanation. If the adolescent hormone explanation is true, the effect of school start times should be larger for older students who are more likely to have begun puberty. I separate the students in my sample by quarter years of age and estimate the effect of start time on test scores separately for each group (including a full set of covariates and a student-school fixed effect). Figs. 7 and 8 present these results. The shaded area represents 95% confidence intervals. For both tests, the effect is relatively flat until age 13, when it begins to increase and continues to increase through the rest of the observed age range. These results provide evidence supporting the adolescent hormone hypothesis.[22]

To further investigate how the effect of later start times varies with age, I estimate the effect of start times on upper elementary (grades 3–5) students and high school students. If adolescent hormones are the mechanism through which start times affect academic performance, pre-adolescent elementary students should not be affected by early start times. Results for elementary students are shown in Table 8. Columns one and three present results using student-school fixed effects. Start times had no effect on elementary students regardless of the specification used. However, elementary schools start much later than middle schools (over half of elementary schools begin at 9:15, and almost all of the rest begin at 8:15). As a result,

[21] On average there are 30 percentiles to a standard deviation, so the 0.06 standard deviation effect for math scores is similar to a 1.8 percentile effect.

[22] If adolescent hormones disproportionally affect early starting students in other ways, this fact could also explain the pattern observed here.

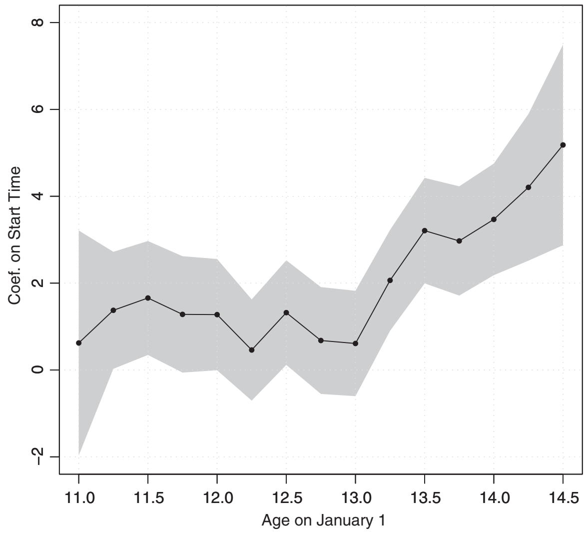

Fig. 7. Effect of start time on math test score by age. *Note*: Each point represents the effect of a 1 h later start time on the percentile rank on the end of grade math exam. The shaded area represents 95% confidence intervals. Age is measured in quarter years as of January 1. See text for additional details.

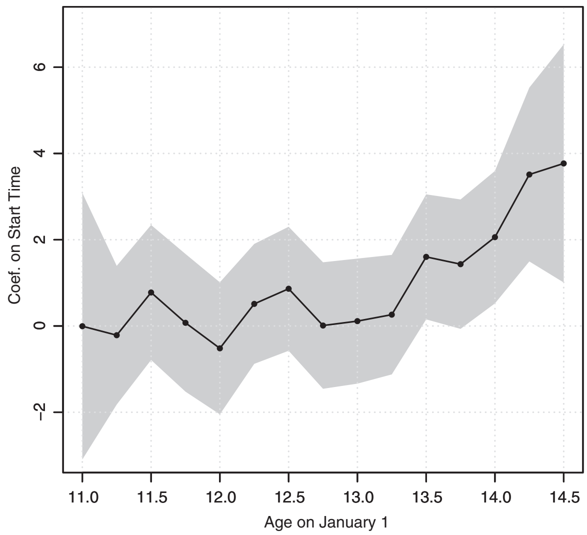

Fig. 8. Effect of start time on reading test score by age. *Note*: Each point represents the effect of a 1 h later start time on the percentile rank on the end of grade reading exam. The shaded area represents 95% confidence intervals. Age is measured in quarter years as of January 1. See text for additional details.

it is not clear if there is no effect because start times do not impact the academic performance of prepubescent students, or because the schools start much later and only early start times effect performance. Columns two and four present results using the (future) start times in grade 6.[23] These specifications serve as a falsification test. If students with higher test scores in elementary school tend to be

Table 8
Results for elementary schools.

	(1) Math	(2) Math	(3) Reading	(4) Reading
Start time	−0.217		0.118	
	(0.196)		(0.214)	
Start time		−0.404		−0.471
(grade 6)		(0.289)		(0.280)
Fixed effect	Student-school	School	Student-school	School
Observations	168,474	167,552	99,413	99,037

Dependent variable is score on end of course math or reading exam. All columns include a full set of control variables, including year and grade effects. All columns include Cluster robust standard errors in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

assigned to later starting schools, the coefficient on 6th grade start time would be positive. That is not the case here; the effect of grade 6 start times is statistically insignificant and has a negative sign.

High school students do not take a specific test at the end of each *grade*. Instead they are required to take an exam at the end of specific *courses*, such as geometry or physical science. There is no explicit requirement for the grade in which each course is to be taken, and the typical sequence often varies from school to school. For example, in some schools, students typically take Geometry before Algebra II, while in others they take Geometry after Algebra II. Furthermore some students take these end-of-course exams in eighth grade (particularly Algebra I). However, all students are required to take the "High School Comprehensive Exam" at the end of grade 10.[24] The comprehensive exam measures growth in reading and math since the end of grade eight and is in most respects similar to the end of grade tests taken in grades 3–8. An important distinction is that students taking the comprehensive exam may not be enrolled in a math class during their sophomore year. Table 9 presents estimation results for the high school comprehensive test. Columns 1 and 4 use the high school start time, columns 2 and 5 use the student's start time in 8th grade and columns 3 and 6 use both. Each specification uses a school fixed effect and so are comparable to the middle school results found in columns 1 and 3 of Table 6. The effect of a 1 h later start time in grade ten is slightly larger than the effect in middle school: 3.3 percentiles on the math test and 3.7 percentiles on the reading test. A 1 h later start time in eighth grade increases test scores on the high school comprehensive exam by 2.0 percentile points for math and 1.6 percentile points for reading. When both start times are included in the same specification, results are similar, although the results for high school are imprecise. In comparison, the effect was 2.1 percentile points for math and 1.5 percentiles for reading in grade eight. These results indicate that the negative impact of early start times persists over time.

Increased sleep is not the only possible reason why later starting students have higher test scores. Students

[23] Since there is no variation in grade six start times within individual students, these models use a school fixed effect.

[24] Due to changes in state policy, the High School Comprehensive Exam was not administered in the 2001–02, 2004–05 or 2005–06 school years.

Table 9
High school comprehensive.

	(1) Math	(2) Math	(3) Math	(4) Reading	(5) Reading	(6) Reading
Start time	3.326**		4.355	3.711**		6.999
	(1.327)		(5.612)	(1.482)		(6.885)
Start time (8th grade)		2.033***	1.983***		1.603**	1.553**
		(0.671)	(0.673)		(0.760)	(0.760)
Fixed effect	School	School	School	School	School	School
Observations	18,491	8512	8512	18,221	8338	8338

Dependent variable is score on high school comprehensive math or reading exam. All columns include a full set of control variables, including year, course and high school effects. Standard errors (robust to clustering at the school level) in parentheses.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

Table 10
Alternative mechanisms.

	(1) TV	(2) Homework	(3) Absences
Start time	−9.45***	12.48***	−1.33***
	(1.80)	(2.86)	(0.40)
Fixed effect	School	School	None
Observations	97,126	9080	30,896
Adjusted R^2	0.33	0.61	0.14

TV is measured in minutes per day, Homework in minutes per week, and Absences in days per year. All columns include a full set of control variables, including year and grade effects. Standard errors (robust to clustering at the school level) in parentheses. See text for additional details.
* $p < 0.10$.
** $p < 0.05$.
*** $p < 0.01$.

in early-starting schools could be more likely to skip breakfast. Since they also get out of school earlier, they could spend more (or less) time playing sports, watching television or doing homework. They could be more likely than other students to be absent, tardy or have behavioral problems in school. Other explanations are possible as well. While my data do not allow me to explore all possible mechanisms, I present evidence in favor or against some explanations below.

A unique aspect of the NCERDC data set is that it includes self-reported amounts of television watched per day and time spent doing homework per week. Column 1 of Table 10 shows the result of a regression of minutes of television per day against start time and a full set of covariates. Students who start school 1 h later watch 12 fewer minutes of television per day. In column 2, I regress minutes of homework per week against start time. Students who start school 1 h later spend 9 min more on homework per week.[25] These results could arise because students who start school earlier spend more time after school at home alone. Students who start school earlier also come home from school earlier and may go to bed earlier. As a result, early-starting students spend more time after school at home alone and less time at home with their parents, relative to late starting students. If students watch television when they are home alone and do their homework when their parents are home, this behavior could explain why students who start school later have higher test scores. In this explanation it is not early start times that matter, but early end times.

The existing start times literature tends to find that students in early-starting schools are both more likely to be tardy to school and to be absent than other students.[26] The data set used in this paper only includes information on absences for two years: the 2003–04 and 2004–05 school years. Since no schools had changes in start time between those two years (there was one new school), I can only consider variation in start times across schools. Column 3 of Table 10 presents the results of a regression of days absent per year on start time and a full set of covariates. Students who start school 1 h later have 1.3 fewer absences (the median student has five absences). Reduced absences may explain why later starting students have higher test scores. Students who have an early start time miss more school (which may or may not be a result of getting less sleep) and as a result perform worse on standardized tests.

6. Conclusion

Later school start times have been often cited in the popular press as a way to increase student performance. However, there has not been much empirical evidence supporting this claim or calculating how large of an effect later start times might have. Using variation in start times both within and across schools, I find that an increase in start times by 1 h would lead to a 3 percentile point gain in both math and reading test scores for the average student. Using only variation within schools the effect is 2 percentile points for math and 1.5 percentile point for reading. The impact of middle school start times on test scores persists into the tenth grade. The effect is larger for the lower end of the distribution of test grades. I find evidence supporting reduced sleep, in combination with the adolescent hormonal cycle as a mechanism through which later start times may affect test scores. I also find evidence supporting time at home with parents as a mechanism.

[25] For both television and homework, the results are similar if I use a school fixed effect, identifying the differences based only on the variation in start times within schools.

[26] See for example (Wahlstrom, 2000).

These results suggest that delaying start times may be a cost-effective method of increasing student performance. Since the effect of later start times is stronger for the lower end of the distribution of test scores, later start times may be particularly effective in meeting accountability standards that require a minimum level of competency. If elementary students are not affected by later start times (which cannot be definitively determined from my data), it may be possible to increase test scores for middle school students at zero cost by having elementary schools start first. Alternatively, the entire schedule could be shifted later into the day. However, these changes may be politically infeasible due to childcare constraints for younger students and jobs and after school activities for older students. A third option would be to eliminate tiered busing schedules and have all schools begin at the same time. A plausible estimate of the cost of moving start times later is the additional cost of running a single tier bus system. The WCPSS Transportation Department estimates that over a ten year period from 1993 to 2003, using a three-tiered bus system saved roughly $100 million in transportation costs (Wake County Public School System Department of Transportation, 2004). With approximately 100,000 students per year divided into three tiers, it would cost roughly $150 per student each year to move each of the 66,000 students in the two earliest start time tiers to the latest start time. In comparison, Krueger (1999) finds that reducing early elementary class size by one third increases test scores by 4 percentile ranks in the first year at a cost of $2151 per student per year. If the same class size effect holds for middle school students, and the effect is linear, it would cost seven times as much to obtain a two percentile point increase by reducing class size than by starting school 1 h later. While very rough, these calculations suggest that increased spending on bus transportation, in order to delay the beginning of the school day, may be substantially cheaper than reducing class size to gain a comparable improvement in test scores.

Acknowledgements

I am grateful to Darren Lubotsky for advice and suggestions, and to Ron Laschever, Daniel McMillen, Elizabeth Powers, and seminar participants at the University of Illinois for helpful comments. I owe a special thanks to Kara Bonneau and Clara Muschkin at the North Carolina Education Research Data Center (NCERDC) for facilitating use of their confidential data and to Ina Stringfellow at the Wake County Public School System Transportation Department for providing me with start time data. I am responsible for any errors.

References

Allen, R. P. (1992). Social factors associated with the amount of school week sleep lag for seniors in an early starting suburban high school. *Sleep Research*, 114.

Baroni, E., Naku, K., Spaulding, N., Gavin, M., Finalborgo, M., LeBourgeois, M. K., et al. (2004). Sleep habits and daytime functioning in students attending early versus late starting middle schools. *Sleep*, A396–A397.

Canay, I. A. (2011). A simple approach to quantile regression for panel data. *The Econometrics Journal*, *14*, 368–386.

Carrell, S. E., Maghakian, T., & West, J. E. (2011). A's from Zzzz's? The causal effect of school start time on the academic achievement of adolescents. *American Economic Journal: Economic Policy*, *3*, 62–81.

Corak, M., & Lauzon, D. (2009). Differences in the distribution of high school achievement: The role of class-size and time-in-term. *Economics of Education Review*, *28*, 189–198.

Cortes, K., Bricker, J., & Rohlfs, C. (2012). The role of specific subjects in education production functions: Evidence from morning classes in Chicago public high schools. *The B.E. Journal of Economic Analysis & Policy*, *12*(1). Article 27.

Dahl, R. E., & Lewin, D. S. (2002). Pathways to adolescent health sleep regulation and behavior. *Journal of Adolescent Health*, *31*, 175–184.

Danner, F., & Phillips, B. (2008). Adolescent sleep, school start times, and teen motor vehicle crashes. *Journal of Clinical Sleep Medicine*, *4*, 533–535.

Dills, A. K., & Hernandez-Julian, R. (2008). Course scheduling and academic performance. *Economics of Education Review*, *27*, 646–654.

Fugenschuh, A. (2009). Solving a school bus scheduling problem with integer programming. *European Journal of Operational Research*, *193*, 867–884.

Hinrichs, P. (2011). When the bell tolls: The effects of school starting times on academic achievement. *Education Finance and Policy*, *6*, 486–507.

Kalish, N. (2008). Early bird gets the bad grade. In *New York Times*. January 14.

Keller, H., & Muller, W. (1979). Optimierung des schulerverkehrs durch gemischt ganzzahlige programmierung. *Zeitschrift fur Operations Research B*, *23*, 105–122 (in German).

Koenker, R. (2005). *Quantile regression*. Cambridge University Press.

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

Meijer, A. M., Habekothe, H. T., & Van Den Wittenboer, G. L. (2000). Time in bed, quality of sleep and school functioning of children. *Journal of Sleep Research*, *9*, 145–153.

Pilcher, J. J., & Huffcutt, A. J. (1996). Effects of sleep deprivation on performance: A meta-analysis. *Sleep: Journal of Sleep Research & Sleep Medicine*, *19*, 318–326.

Trudeau, M. (2007). High schools starting later to help sleepy teens. In *Morning Edition*. January 18.

Wahlstrom, K. (2000). Changing start times: Findings from the first longitudinal study of later high school start times. *NASSP Bulletin*, *86*, 3–21.

Wake County Board of Education (2004). Open Minutes – Board Meeting. Minutes of Tuesday, April 13, 2004.

Wake County Public School System Department of Transportation. (2004). The three-tier bus transportation system. *Wheels of Education*, (Winter).

Wolfson, A., & Carskadon, M. (2003). Understanding adolescents' sleep patterns and school performance: A critical appraisal. *Sleep medicine Reviews*, *7*, 491–506.

Wolfson, A., & Carskadon, M. (2005). A survey of factors affecting high school start times. *NASSP Bulletin*, *89*, 47–66.

Wolfson, A. R., Spaulding, N. L., Dandrow, C., & Baroni, E. M. (2007). Middle school start times: The importance of a good night's sleep for young adolescents. *Behavioral Sleep Medicine*, *5*, 194–209.

Wong, J. (2011). Does school start too early for student learning? Working Paper.