Academic Achievement Across the Day: Evidence from Randomized Class Schedules

Teny Maghakian Shapiro * Kevin M. Williams†

April 5, 2016

Abstract

Students' biological rhythms influence their ability to focus and learn. This paper examines how the daily structure of classes could be reorganized to improve student achievement, all within the confines of a traditional school day schedule. Specifically, we identify the effects of course timing, student fatigue, and teacher schedules. Data consists of five cohorts of college freshman who face randomized scheduling. We find students perform 0.2 standard deviations better in the afternoon than in the morning, but fatigue from prior courses dampens net gains. Heterogeneous effects suggest that redistributing schedules could aid low-achievers, equivalent to improving teacher quality by 0.4 standard deviations.

JEL: I20, J24
Keywords: human capital, class schedules, learning, school start time, adolescents

*Santa Clara University; Email: tshapiro@scu.edu
†University of California, Davis; Email: kmwilliams@ucdavis.edu

The views expressed in this article are those of the authors and do not necessarily reflect the official policy or position of the USAF, DoD, or the U.S. Government. This article was completed with data that is managed in collaboration with Lt. Col. Scott Carrell and Jim West. We thank Kathryn Bruchmann, Scott Carrell, Seth Gershenson, Andrew Goodman-Bacon, Hilary Hoynes, Matthew Larsen, Lisa Schulkind, and Adam Shapiro for their valuable input as well as seminar participants at Santa Clara University, UC Davis, the Southern Economic Association Annual Meeting, the APPAM Annual Conference, and the SREE Annual Conference.

1. Introduction

Nearly sixty percent of high school students report feeling tired during class, making it hard for them to focus or stay awake (National Sleep Foundation (2006)). While some inattention during the school day is inevitable, it may also be preventable. A growing body of evidence from economics and other social sciences has shown students' alertness and academic achievement is strongly affected by their biological rhythms. Student's focus will be naturally different based on the time of day a class takes place and what they have done earlier that day. Thus far, the strongest evidence of the impact that class times have on achievement comes from showing that delaying school start times has a positive effect on teens (Carrell et al. (2011), Edwards (2012)). While the evidence on delaying start times has led to passionate discussions in a number of schools and districts across the United States, relatively few school have actually changed their start times, with opponents arguing that the challenges of delaying start times are too large to overcome.

In this paper, we extend upon the link between students' academic achievement and their biological rhythms and determine precisely how the organization of courses throughout the school day impacts performance. Specifically, we explore the independent roles of three aspects of the school day schedule: the time a class is held, student fatigue due to preceding courses, and the instructors' schedule. We also determine heterogeneity of these effects across course and student type and determine whether the order of classes and breaks can affect achievement. Finally, we show through simulations that improvements in average student achievement are possible by rescheduling students within the confines of existing scheduling constraints. By understanding precisely how these features affect academic achievement, school administrators and students may be able to improve outcomes without needing to alter the overall timing and structure of school schedules.

This is the first paper to separately identify the effects of the time a class is held, student fatigue, and the instructors' schedule. This would be difficult, if not impossible, to do in most school settings due to selection into courses/instructors and the subjectivity of grading. Further, in schools where students and teachers are assigned a class during each period, the effect of time of day can not be separately identified from the effect of fatigue. We are able to overcome these issues by utilizing data from the United States Air Force Academy (USAFA). The school day at USAFA is split into seven class periods, four before lunch and three after, a daily structure very similar to that of the average U.S. middle or high school. But there are a number of other institutional characteristics at USAFA that make it an ideal setting to assess the role of schedules on academic achievement. Schedule assignment is random, grading and instruction are standardized across all sections of a course, exams are taken during a common testing session and teachers regularly teach multiple sections of the same course. Students also alternate daily between two class schedules within the same semester. While total academic course load is similar across students, the alternating schedule creates variation in how much time students spend in class on a given schedule-day. It also allows us to assess how a student performs with one schedule relative to *themselves* with a different schedule.

We focus our analyses on fall-semester freshmen, as they are still in their teens and much of the focus of changes in school start times and schedules is on teens because of their distinct time preferences and its misalignment with traditional school schedules Crowley et al. (2007). We recognize that USAFA students are not the average student; they were high-achievers in high school and chose to attend a military service academy. Although we do not know for certain if school schedules affect high-achievers or military-types differently than the average student, we have no reason to believe that the students in our sample would be *more* adversely affected by components of their daily schedule than the average teen.

We find that, all else equal, the afternoon is the best time of day for student learning, but gains from having a class during the afternoon relative to the morning are mostly offset by fatigue. Specifically, if a student were taking their first class of the day at 2:00 p.m. rather than 7:30 a.m., they would perform about a fifth of a standard deviation better. However, fatigue is such that a student in a 2:00 class which follows a full schedule of classes is predicted to perform 0.13 standard deviations *worse* than in the 7:30 class. Even two students in the same class at the same time of day may have differences in expected grade as large as 0.1 standard deviations simply due to variation in fatigue from their prior schedules.

Subgroup analysis reveals that the negative effects of fatigue are more extreme for students in the bottom tercile of predicted aptitude. In fact, students in the top tercile are impacted by neither the time of day a course takes place nor schedule fatigue. This suggests that schools can raise mean performance by assigning struggling students to the schedules that are best matched with biological rhythms. Our simulations, which reassign the worst students to the best schedules, find that we can obtain gains for bottom tercile students that are equivalent to increasing teacher quality by 0.4 standard deviations in all courses.[1] We conclude with a discussion of policies, obstacles, costs, and benefits facing the implementation of a rescheduled school day and argue that, compared to many of the inputs commonly studied in the education production function, such as teacher quality and class-size, rescheduling classes to better align with students' optimal learning times is a cost-effective intervention that may be easier for schools to implement than a later start time.

2. Background

To fully understand how the organization of the school day schedule can influence academic achievement, it is important to have a basic understanding of the biology of sleep, wakefulness, and daily fluctuations in cognitive function. The biological rhythm that governs our sleep-wake cycle is called the circadian rhythm, a hard-wired "clock" in the brain that controls the production of the sleep-inducing hormone melatonin. During adolescence, there are major changes in one's circadian rhythm. More adult-like sleep patterns develop, there are increases in daytime sleepiness, and there is a shift in the circadian rhythm towards later bed and wake-up times (Crowley et al. (2007); Carskadon et al. (1993); Wolfson and Carskadon (1998)). There are also times of the day

[1]This is scaled relative to work from Chetty et al. (2014), Kane and Staiger (2008) and Carrell and West (2010) who find a one standard-deviation increase in teacher quality benefits students anywhere from .1-.2 of a standard deviation. We conservatively use a .1 standard deviation improvement for relating our predicted schedule impacts an increase in teacher quality.

when a person is more alert, independent of sleep, which is related to their circadian timing (Blake, 1967). For adolescents, alertness begins in the late morning, drops off mid-afternoon, and peaks again in the early evening (Cardinali, 2008). Research from various scientific fields, including neurobiology and cognitive science, finds that an individual's ability to learn fluctuates throughout the day based on both their biological rhythm (Schmidt et al., 2007) and the total amount of mental activity they have already engaged in (Persson et al., 2007). Goldstein et al. (2007) find that teens perform six points higher on IQ tests if tested during their preferred time of day. Standard academic schedules are quite "out of sync" with teens' circadian rhythms and require students to wake up earlier than their ideal wake time and have many of their classes at a time that is asynchronous with their optimal cognitive function.[2]

Our understanding of sleep, wakefulness, and cognitive function suggests that a student's daily schedule can affect their grades in two separate ways. The first is through the timing of the class; students may not perform as well if classes are scheduled when they're naturally less alert. We refer to this as the time-of-day effect. The second is through the cognitive load a student has experienced prior to the start of a class. We refer to this as the student fatigue effect. While we expect student fatigue to unambiguously hinder academic performance, the time-of-day effect may vary throughout the day. Because academic achievement is an interaction of both learning and teaching, we also estimate the effect of instructor fatigue. However, the expected effect of instructor fatigue is ambiguous. Unlike students, teachers are frequently assigned to teach the same class multiple times per day. Tiredness and mental fatigue could mean teachers are less effective as the day goes on, but learning-by-doing could lead to improvements later in the day. Separately identifying these three components of the daily school schedule allows us to suggest a number of strategies for improving student achievement.

[2]This is not to understate the importance of sleep, which itself is an important to cognitive function. Several studies find an inverse relationship between sleep and academic performance at both the secondary and post-secondary level (Curcio et al., 2006; Wolfson and Carskadon, 1998; Trocket et al., 2000).

A few of strands of literature have assessed the role of time and scheduling on academic and workplace outcomes. The impact of school start times on student achievement has been studied using natural variation across schools or cohorts for identification. The findings have been mixed. Edwards (2012) find positive effects from start time delays on standardized test scores and Carrell et al. (2011), who use the same data as this study, find that grades throughout the entire day benefit from later start times. Meanwhile, Wahlstrom (2002) and Hinrichs (2011) find no effect from the start time change within the Minneapolis Public School district.

Relatively few studies have looked at differential achievement across morning and afternoon classes. Pope (2015) concludes that learning actually decreases throughout the school day by comparing standardized test scores of students who had classes in the morning versus afternoon. Cortes et al. (2012) and Dills and Hernandez-Julian (2008) find the opposite – students perform better in classes that meet later in the day. In each of these studies, the time-of-day effect can not be separately identified from the effect of fatigue.

3. Data

Data for this study come from the United States Air Force Academy (USAFA). USAFA is a fully accredited post-secondary institution with annual enrollment of approximately 4,500 students, offering 32 majors within the humanities, social sciences, basic sciences, and engineering. Despite its military setting, USAFA is comparable to other selective colleges and universities in the United States. Like other selective post-secondary schools, USAFA faculty hold graduate degrees from high quality programs in their fields. Approximately 40 percent of classroom instructors have terminal degrees, similar to large universities where graduate students teach introductory courses. However, class size at USAFA is rarely larger than 25 students, and students are encouraged to interact with faculty members in and out of the classroom. Therefore, the learning environment at USAFA is similar to that of small liberal arts colleges. Students at USAFA are high achievers, with average math and verbal SAT scores at the 88^{th} and 85^{th} percentiles of the nationwide SAT distribution, respectively. Only 14 percent of applicants where admitted to USAFA in 2007. Students are drawn from each Congressional district in the US by a highly competitive admission process that ensures geographic diversity.

A number of USAFA's institutional characteristics make it ideal for addressing this research question. First, the school day at USAFA is very structured, which is atypical of most universities, but similar to a high school setting. Table 1 shows the class schedules for our sample period. There are four 53-minute class periods each morning and three each afternoon after an 85-minute lunch break.[3] All students are required to attend a mandatory breakfast 25 minutes before first period, thus, they must all wake up at around the same time regardless of the time of their first class. Second, students are randomly assigned to all of their courses and instructors. Prior to the start of freshman year, students take placement exams in mathematics, chemistry, and select foreign languages. Scores on these exams are used to place students into the appropriate starting courses (e.g., remedial math, Calculus I, Calculus II, etc.). Conditional on course placement, athlete status, and gender, the USAFA registrar randomly assigns students to required course sections. Students have no ability to choose the class period or their professors in the required core courses. Third, attendance in all classes is mandatory. Fourth, USAFA's grading structure for core courses allows for a consistent measure of student achievement; faculty members teaching the same course in each semester use an identical syllabus, give the same exams during a common testing period, and assign course grades jointly with other instructors, allowing for standardized grades within a course-semester. Finally, USAFA runs on an M/T schedule. On M days, students have one set of classes and on T days they have a different set of classes. The M/T schedules alternate days of the week. Thus, the same student has two different class schedules within the same semester.[4] These

[3]The class schedule changed twice during this time period. In our robustness analyses, we show that this does not affect our findings.

[4]Language courses are an exception and meet every schedule-day during the same period. Students are coded as in class for both M and T day of their language course, but only the grade and preceding courses from the M day are included in analysis.

institutional characteristics provide us with random variation in class schedules both across and within students, which, along with extensive background data on students, allow us to examine how course scheduling affects student achievement without worrying about confounding factors or self-selection issues. Athletes are dropped from primary analysis due to their course schedules being influenced by practice times.

Our dataset consists of 4,816 first-year students from the entering classes of 2004 to 2008 who collectively took 22,600 core courses. For each student we have pre-treatment demographic data and measures of their academic, athletic, and leadership aptitude. Academic aptitude is measured through SAT verbal and math scores and an academic composite computed by the USAFA admissions office, which is a weighted average of an individual's high school GPA, class rank, and the quality of the high school they attended. The measure of pre-treatment athletic aptitude is a score on a fitness test required by all applicants prior to entrance. The measure of pre-treatment leadership aptitude is a leadership composite also computed by the USAFA admissions office, which is a weighted average of high school and community activities. Other individual-level controls include indicators for whether a student is Black, Hispanic, Asian, female, a recruited athlete, whether they attended a military preparatory school, and the number of class credits students have on that schedule-day.

We measure academic performance using students' final percentage score earned in a course. To account for differences in course difficulty or grading across years, we normalize all scores to a mean of zero and a variance of one within a course-semester. We refer to this measure as the student's normalized grade. We also consider whether a student received an A or F in the course as an outcome to see the impacts on the extremes of the grade distribution. Students at USAFA are required to take a set of approximately 30 core courses in mathematics, basic sciences, social sciences, humanities, and engineering. In this study, we focus primarily on the mandatory introductory courses in mathematics, chemistry, engineering, computer sciences, English, foreign languages and history. We refer to these as the required freshman courses. Grades in the humanities courses (English and history) are mostly determined by papers and assignments done outside the classroom, whereas grades in STEM (science, technology, engineering, and math) courses are based on performance on common exams. Accordingly, we examine the effects of STEM and non-STEM course timing separately to see if the effects differ across course type.

Tables 2 and 3 show summary statistics for our sample. The data are at the student-course level. Column (1) of Table 2 shows the summary statistics at this level. Column (2) shows summary statistics at the student-level. Nineteen percent of the students are female, approximately four, eight, and nine percent are black, Hispanic, and Asian, respectively. The mean SAT math score was 669. Column (3) of Table 2 shows statistics for the freshman core courses that we focus our analysis on, while Column (4) shows the STEM core classes specifically. Students enrolled in STEM classes are very similar to those in all required courses. This makes us confident that there is no selection into STEM courses by higher achieving students. The final columns show the characteristics of the students by their tercile of academic composite scores; the "high" tercile are the highest achievers.

Table 3 shows summary statistics by class period. There are some differences across class periods. First, the number of observations for each class period differs, with the most for fifth period (4,600) and the fewest for seventh period (1,738). Student characteristics also vary, as do grades. The goal of this analysis is to determine how much of the variation in grades across the class periods is due to time of day and course schedules, abstracting from differences in student, instructor, and course characteristics.

4. Methodology and Results

Primary Analysis

We begin our analysis by verifying that assignment to different class periods is random with respect to student ability. To do so, we regress student background characteristics on periods of the day dummy variables and course-semester fixed effects to capture within-course deviations in characteristics. Figure 1 shows the results for the distribution of females, minorities, academic composite, SAT math and verbal and peer academic composite. The 90% confidence intervals are shown. All individual characteristics are clearly uncorrelated with class period. Peer academic composite is the one variable showing differences, with peer "quality" being lower in the morning and higher in the afternoon. This is due to the inclusion of athletes whose courses are disproportionately in the morning. Athletes are included when calculating other students' peer variables, but excluded from the sample we analyze. Carrell et al. (2010) further show that student assignment to required courses at USAFA is random with respect to peer characteristics and professor experience, academic rank, and terminal degree status. They also find no correlation between student characteristics and professor gender. Nonetheless, we are also careful to control for classroom-level peer characteristics to address differences in peers across classes and control for professor characteristics by including instructor-semester and course-by-day fixed effects.

To get a general sense of how grades fluctuate throughout the day, we regress the normalized grade on period dummy variables and course-semester fixed effects. The estimates for all students in our sample are shown in the top panel of Figure 2. The second panel shows grades for STEM and non-STEM courses separately. The third panel shows grades by the three terciles of academic aptitude. A few patterns emerge. First, grades rise and fall over the course of the day–grades dip during 1^{st}, 4^{th} and 7^{th} periods and a peak during 2^{nd} and 6^{th} periods. Second, performance in STEM courses is generally higher than in non-STEM classes, but they follow a very similar pattern across class periods. Finally, the general pattern is similar across ability groups, but appears to be more pronounced for the lower tercile students. Within these patterns are some interesting puzzles. Mean performance in 2^{nd} period is quite strong even though it is at a time asynchronous with adolescents' optimal learning times. Alternatively, 4^{th} period is at a time that is synchronous with adolescents' optimal learning times for learning; however, mean grades in those periods are quite low. While looking at means gives us some insight into patterns that may exist, especially at USAFA where courses and professors are randomly assigned, they also reflect differences in courses offered during different class periods and differences in professor quality. Using a regression framework, we are able to disentangle the effect of different components of the daily class schedule on student achievement from all other attributes of the student and their schedule. To do so, we estimate the following equation:

$$\begin{aligned} Grade_{icjtsp} =& \alpha + \psi_p + \beta Fatigue + \mu InstructorSchedule + \delta_1 X_{ict} + \delta_2 Peers_{cjtsp} \\ &+ \phi_{cts} + \gamma_{jt} + \rho_i + \epsilon_{icjtsp} \end{aligned} \tag{1}$$

where $Grade_{icjtsp}$ is the normalized grade for student i in course c with instructor j on schedule-day s in period p in year t. ψ_p are period-of-day dummies with 1^{st} period omitted, which measure the time-of-day effect. $Fatigue$ is a vector of the student fatigue characteristics, which we discuss in detail below and $InstructorSchedule$ is a vector of instructor schedule characteristics, also described below. The vector X_{ist} includes the following student characteristics: SAT math and SAT verbal test scores, academic and leadership composites, fitness score, race, gender, whether s/he attended a military preparatory school, and how many credit hours the student had on that schedule-day. To control for classroom peer effects, we include $Peers_{icjtsp}$, the average pre-treatment characteristics of all students in the class except for individual i.[5] ϕ_{cst} are course by year by schedule-day fixed effects, which control for unobserved mean differences in academic achievement or grading standards across courses, years, and schedule-days. Professor by year fixed effects, γ_{jt}, control for fixed differences in instructors within a given year. We also show specifications that include individual student fixed effects, ρ_i, to exploit the within-student variation in schedules across the M/T schedule-days. Standard errors are clustered by student.

---

[5]Formally, the $Peers$ variables are defined as follows: $\frac{\sum_{k \neq i} X_{kcjtsp}}{n_{cjtsp} - 1}$, where X represents the various observable student characteristics.

To assess the effect of student fatigue, we exploit the random variation in both the number of classes a student has had before a given class without a break (consecutive classes) and the number of total classes a student has had before a given class (cumulative classes). The number of consecutive and cumulative classes can vary both *across* students and *within* students because of the M/T schedule-days. For example, Student A may have classes during 2^{nd}, 4^{th}, and 6^{th} periods on one schedule-day, while Student B has classes during 1^{st}, 2^{nd}, 5^{th}, and 6^{th} periods. By 6^{th} period, Student A has had two cumulative classes, but zero consecutive classes (since he had 5^{th} period off), while Student B has had three cumulative classes and one consecutive class. If academic achievement is affected by having had to focus and learn earlier in the day, Student A and B's performance in 6^{th} period will be affected by the time the class is held *and* the number of classes they have had that day, both consecutive and cumulative.[6] Accordingly, we include the following variables in the $Fatigue$ vector: the number of consecutive and cumulative classes a student had before a class and the squares of these variables to account for non-linear fatigue effects.[7]

We include analogous variables in the $InstructorSchedule$ vector: the number of consecutive and cumulative classes an instructor has taught before a given class and the squares of these terms. As with students at USAFA, there is random variation both in the number of consecutive and cumulative classes a professor has taught before a given class. It is unclear, a priori, exactly how instructor schedules should affect student achievement. Teaching may not be cognitively-taxing as learning, but certainly leads to more physical fatigue. While instructors may grow tired as they teach more classes (reflected by a negative effect on student grades), they may also become better at teaching that specific content (reflected by a positive effect on student grades).

[6]We count lunch as a break, so 5^{th} period classes are always given a consecutive value of zero. We have explored alternate definitions of the variable where we do not consider lunch a break and results are quantitatively similar.

[7]We acknowledge that there are a number of ways to define student fatigue, and have explored a variety of alternate definitions. Results are qualitatively very similar and are available upon request.

Classmate Comparison

Our first analysis identifies the time-of-day and fatigue effects on learning by leveraging variation in the times a given course is offered. Our second approach considers only within-class differences in performance. Here, rather than comparing two students taking the same course at different times of the day, we are comparing students in the same class (classmates) who had different schedules earlier in the day. This is achieved by including section specific, rather than course specific, fixed effects. For a given section of a class, students have been randomly assigned to the section at hand and also their preceding schedules. In essence, a student's schedule immediately beforehand can be thought of as a "treatment" on their ability to learn at that time. By comparing students in the same section, we are holding teacher quality and time of day constant.

We refer to the student's schedule preceding a class as their $LeadUp$ scenario and estimate the following equation:

$$Grade_{icjtsp} = \alpha + \beta LeadUp_{icjtsp} * \psi_p + \delta_1 X_{ict} + \delta_2 Peers + \phi_{ctspj} + \gamma_{jt} + \rho_i + \epsilon_{icjtsp} \qquad (2)$$

The primary difference between Equation 1 and this one is the inclusion of the $LeadUp_{icjtsp}$ variables, of which there are four possibilities: Free Period, P.E., STEM Class, Non-STEM Class. Interacting the $LeadUp$ variables with period dummies, ψ_p allows for the effect of a student's prior classes to vary over the day. The section fixed effects, ϕ_{ctspj}, replace the course fixed effects from Equation 1.[8]

Results

The results from the primary analysis are shown in Table 4. The first two columns show the coefficients when excluding the fatigue measures from Equation 1. All columns include individual controls, peer controls and teacher fixed effects. Even-numbered columns also include individual

[8]As before, only core freshman courses are considered. 1^{st} period observations are dropped due to lack of variation.

fixed effects. The third and fourth column add measurements of student fatigue and the fifth and sixth add the instructor fatigue variables. The estimates show that the time of day a class is taken can have a large effect on achievement. For example, all else equal, a student taking their first class of the day at third period performs approximately 0.14 standard deviations better in a class than if they had taken the same class during first period. Interestingly, the period coefficients follow a similar pattern to adolescent sleep-wake cycles, where alertness increases throughout the morning, dips in the early afternoon and then rises again. For easier interpretation, we plot the period coefficients from Column (5), our preferred specification, and their 90% confidence intervals in the first panel of Figure 3.

The student fatigue estimates show consistently negative effects of consecutive classes– each consecutive class decreases performance in a course by about 0.06 standard deviations. The number of cumulative classes a student had before a given class also has negative effects on achievement, but the statistical significance of these estimates are sensitive to the econometric specification. These results suggest that achievement is certainly affected by the fatigue that students experience throughout the school day. However, student achievement is not affected by instructor fatigue. That is, an instructor's prior experience during the school day has no effect on a student's performance in a class. Aggregated coefficients that correspond to a regular daily class schedule are shown in Table 7. The first column assumes student has a full schedule with no breaks besides lunch. The second column shows predictions for a student who has one free period, which is assumed to be in the prior period for each estimate (i.e. number of consecutive classes is always assumed to be 0). Fatigue hinders students' performance as the school day progresses, offsetting the benefits of the later time. For example, a student taking a class during 3^{rd} period is estimated to perform 0.024 standard deviations *worse* than if they had taken the same class during 1^{st} period if the 3^{rd} period class is following two consecutive classes. However, they are estimated to do 0.106 standard deviations *better* if they have 2^{nd} period off. Pope (2015)'s conclusion that students perform better in the morning than in the afternoon is likely a result of accumulated fatigue in the afternoons, not because students learn better in the morning.

We next assess the heterogeneity of these effects across subsamples of the data. Doing so can help us understand how to optimize class schedules so that the classes and/or students that benefit the most from being during "prime" times are the ones given those times. We use our preferred specification, which includes all fatigue variables and instructor fixed effects, for these analyses.[9] Estimates are shown in Table 5. Column (1) shows estimates with our full sample for easy comparison. Columns (2) - (4) show estimates for students based on their predicted academic tercile upon entering USAFA. It is important to note that since USAFA is a highly selective institution, even the bottom tercile students are among the top 15 percent of students nationally.

We see no statistically significant effects for the top tercile students. Middle tercile students are negatively impacted by having consecutive classes, but their performance does not vary with the time of day. The bottom tercile students, on the other hand, are quite affected by both the time of day and fatigue. The time-of-day effect is striking for this group. The bottom tercile students perform a quarter of a standard deviation better in a 4^{th} period class than during 1^{st} period, with the effects of the afternoon classes being even larger.

These differences by subgroup have meaningful implications for how schools could use scheduling to improve mean achievement. Top students seem robust to their schedules. A likely explanation is that top students are both generally more focused throughout the day and also better able to learn material on their own outside of class. Lower ability students are more dependent on absorbing knowledge during lectures and thus are more adversely affected when unable to focus in class. Given that USAFA students are high achievers to begin with, the cognitive penalties could be even larger for other student populations. Some ill-timed classes are inevitable, but a school could target better schedules at struggling students.

[9]We do not show the estimates from the instructor fatigue variables in the table since they continue to be statistically insignificant; however, these variables are included in the regressions.

Columns (5) and (6) of Table 5 show estimates for STEM and non-STEM classes, respectively. Estimates for STEM classes are very similar to those from our full sample. While non-STEM courses hold a similar pattern, most of the estimates are not statistically different from zero. It's important to note that there are more observations of STEM classes, since a larger share of USAFA's core classes are STEM. Columns (7) and (8) include the full sample of students, but considers different outcomes to better understand the margins at which student achievement is affected. Ace considers whether a student earned an A or A- in the course and fail considers whether the student earned a D or F in the course. Time of day affects students at both these margins, but more significantly in the morning class periods.

Results from the classmate comparison are shown in Figure 4. This specification includes section fixed effects and variation comes from comparing differences in prior schedules among students in the same section. Each bar in the figure represents a single coefficient, β, from Equation 2 sorted by period of the day and color-coded by $LeadUp$ scenario. Having had a non-STEM course prior is the reference group. The second graph includes individual student fixed effects.[10] For two students in the same 3^{rd} period section, one who had a free period during 2^{nd} period and one who had a non-stem course, the student with the free period has an expected normalized grade of 0.15 standard deviations higher. P.E. is similarly beneficial in the morning, but it doesn't seem as though the physicality of P.E. causes it to have differential effects from a free period. A free period beforehand is a strong predictor of success in 2^{nd}, 3^{rd}, 4^{th} and 6^{th} periods; 7^{th} period is an interesting exception. Here, both P.E. and a free period beforehand lead to an expected decrease in performance. One explanation is that these students are mentally "checked out." Lunch, combined with either P.E. or no class means that students have had a nearly three and a half hour break from the classroom. After being in academic mode for some portion of the morning, it may be difficult for students to take an extended break and then re-focus for a single afternoon class. Students who

[10]P.E. is a two-period class, but only meets starting in periods 1, 3 ,and 5 so there are no estimates for a P.E. $LeadUp$ effect in periods 2 or 4.

have two or more classes in the afternoon (thus likely having had breaks in the morning) appear to be better able to perk-up for their classes after lunch.

Robustness Checks

We verify the robustness of our estimates to several changes in model specification with results shown in Table 6. Columns (1) excludes foreign language courses from the analysis. Since students select into their foreign language and classes meet on both schedule-days, these are the least subjective of all the core courses. Column (2) excludes chemistry courses from the analysis since it is a two-period long class. Next, we verify that the results are not purely driven by one of the start-time regimes. Column (3) shows the model restricted to academic year 2007, when first period started at 7:00 a.m. Column (4) is limited to academic years 2005 and 2006, when first period started at 7:30, and Column (5) is for academic years 2008 and 2009, with a 7:50 a.m. start time. The point estimates in the 7:00 start time are much larger than the other years; however, the time of day patterns are similar, as are the fatigue effects. Column (6) shows estimates with the inclusion of recruited athletes, a group whose class assignment may not be as random since they are generally not assigned afternoon courses. The estimates from our robustness specifications are qualitatively similar to those from our main specification, and provide strong evidence that our results are not driven by anomalies in the data.

Simulations

Our main results focus on the impact that time of day, student fatigue and instructor schedules have on grades in individual classes, but we are also interested schedules' aggregate impacts. To examine the overall impact course rescheduling could have on student achievement, we perform two simulations where we assess how achievement would differ if students were assigned schedules based on their academic aptitude. The simulations aim to estimate the extent to which course schedules could be used to equalize student outcomes while raising mean achievement.

We first calculate a predicted own-GPA score for each student, determined using only their background characteristics (e.g. gender, SAT scores). This is used to rank students by predicted ability. Then, every observation (i.e. a student-course) is assigned a schedule-GPA using the time of day and fatigue coefficients.[11] The schedule-GPA value is independent of the characteristics of the student and represents the average impact of each course in a student's schedule. We calculate the average of the schedule-GPA scores by student, yielding each schedule's average impact. In total, there are 4,536 student-schedules, representing over 1,900 different combinations of schedules.

We then re-assign schedules, controlling for the number of courses students take, such that students with low predicted GPAs are given the best-performing schedules and the high-ability students are given the worst ones. We limit ourselves to the set of existing schedules in our data to ensure that results would be feasible within USAFA's current scheduling constraints, such as faculty size and classroom availability, which most schools also face.

The first simulation assumes that the time-of-day and fatigue effects are homogenous across students. Results are shown in Table 8. Results show a narrowing of the overall grade distribution with no change in average performance across students. Specifically, the standard deviation of grades decreases from 0.415 to 0.397, an overall decrease in variance of around 8%. Bottom tercile students experience a 2% of a standard deviation increase in overall performance, but a similar loss is predicted for the top tercile.

The second simulation allows for heterogeneous effects of schedules based on students' predicted ability. The simulation is done in a similar fashion, but we use coefficients from Columns (2) - (4) of Table 5 to estimate the schedule-GPA of each course. Once again, students are assigned schedules in an inverse relationship to their predicted ability– the worst student (based on predicted

[11] Both own-GPA and schedule-GPA use coefficients taken from Column (5) of Table 4.

ability) is assigned the best schedule, while the best student is assigned the worst schedule.[12] By assuming that top, middle, and bottom tercile students experience differing effects of fatigue, there is an opportunity to both narrow the overall GPA distribution and also raise mean performance.

Results from the second simulation are shown in Table 9 and show that re-assigning schedules raises expected performance by 1.2% of a standard deviation for all students. Variance in student achievement is again reduced by 8%. These gains are concentrated in the bottom tercile of student ability – this group experiences an average GPA increase of 3.3% of a standard deviation. This gain is equivalent to increasing teacher quality in all their courses by 0.4 standard deviations. Middle- and top-tercile students experience very slight gains and losses, respectively. Since these students are less affected by course schedule than lower-ability students, being assigned sub-optimal course schedules does not greatly affect their predicted achievement.

5. Discussion and Conclusions

The goal of this study was to determine how the organization of classes throughout the school day affects academic achievement for adolescents. To do this, we consider the effects of three distinct components of course schedules: time of day, student fatigue, and teacher fatigue. We find clear results that both the time of day a class is held and the level of cognitive fatigue a student faces impacts their academic achievement in the class. Two similar students taking the same classes with the same teachers, but with different schedules could be expected to get grades as different as two-tenths of a standard deviation (approximately a grade difference of a B- to a B+). These findings support the notion that the way in which school schedules are currently organized hinders student performance. Adolescents learn better in the late morning and afternoon– times that are better aligned with their circadian rhythms. These results are consistent with Goldstein et al. (2007) who find, that for adolescents, scores on intelligence tests are significantly lower

[12] For each student's schedule, three scores are calculated: one each for the schedule's predicted impact on top, middle and bottom tercile students. A schedule's predicted impact is, again, independent of student characteristics and is only based on the time of day the course takes place and timing of courses a student took earlier on that day.

during the early morning hours. The course and grading structure at USAFA is ideal for this study. Assignment to classes and professors is random, attendance in all classes is mandatory, and all students enrolled in a course in a given semester take the exams during a common testing period and are graded on a collective curve. Because of these features we can be certain that the effects we find reflect differences in learning/understanding of class material and not differences in grading standards. Lower ability students at USAFA, the population in our sample most likely to be similar to the average student nationwide, are most affected by class timing. This research extends our understanding of what outside factors affect academic achievement and provides an opportunity to increase achievement, and, presumably human capital, by rescheduling the times that classes are held.

There are several recommended policies, or rules-of-thumbs, administrators or students could follow, based on our results.[13] First, our findings consistently and strongly support start times in the 9 a.m.-10 a.m. range. However, shifting a school's entire schedule may be expensive to implement or an unpopular policy among parents, teachers, and coaches.[14] Morning P.E. classes, however, can be an effective way to mitigate some of the negative effects of early start times. We also show a clear penalty of consecutive classes, especially for the lowest-performing students. Thus, scheduling free periods and P.E. so they provide breaks throughout the day is beneficial to students.[15]

Subgroup results show that average ability students and STEM classes are most susceptible to the effects of timing and fatigue. The STEM effect could be due to the nature of STEM classes

[13]We recognize that schedules are often difficult to create, because of the multitude of constraints facing specific schools and districts. These include factors such as busing and transportation schedules, after-school programs, classroom availability, athletic schedules, field availability, and teacher loads, among others.

[14]See Jacob and Rockoff (2011) for a full discussion. The authors find that moving back school start time may cost anywhere from $0-$1,900 per student.

[15]We show that an hour-long lunch is akin to a free period. Thus, a break immediately before or after lunch does not provide as much benefit. Free periods during the last period of the day are also wasteful– teens learn well in the afternoon and breaks are best used to offset accumulating fatigue. Sports commonly dictate that students have their last period free because of scheduling conflicts, but our evidence suggests that giving students their last period free should be avoided whenever possible.

(often more lecture based versus discussion based non-STEM courses), but also may simply be due to the limited non-STEM courses in our data. These results suggest that a student's weakest classes should be scheduled at the best times of day, either in the afternoon or following a break. In general, targeting one or two classes per student for optimal timing may be more feasible than restructuring their entire schedule. Unfortunately, our data do not allow us to determine *why* differences across ability groups exist. There are a number of hypotheses as to what explains this difference (high achieving students may be better able to learn when tired, teach themselves material they missed in class, or are more likely to be morning-oriented individuals), and would be an interesting area for further research within the social or biological sciences.

References

Blake, M. (1967), 'Time of day effects on performance in a range of tasks', *Psychonomic science* **9**(6), 349–350.

Cardinali, D. (2008), Chronoeducation: How the biological clock influences the learning process, *in* A. Battro, K. Fischer and P. Lena, eds, 'The Educated Brain', Cambridge University Press.

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

Carrell, S. E., Page, M. E. and West, J. E. (2010), 'Sex and science: How professor gender perpetuates the gender gap', *The Quarterly Journal of Economics* **125**(3), 1101–1144.

Carrell, S. E. and West, J. E. (2010), 'Does professor quality matter? evidence from random assignment of students to professors', *Journal of Political Economy* **118**(1).

Carskadon, M., Vieira, C. and Acebo, C. (1993), 'Association between puberty and delayed phase preference', *Sleep* **16**(3).

Chetty, R., Friedman, J. N. and Rockoff, J. E. (2014), 'Measuring the impacts of teachers i: Evaluating bias in teacher value-added estimates', *American Economic Review* **104**(9), 2593–2632.

Cortes, K., Bricker, J. and 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**.

Crowley, S., Acebo, C. and Carskadon, M. (2007), 'Sleep, circadian rhythms, and delayed phase in adolescents', *Sleep Medicine* **8**.

Curcio, G., Ferrara, M. and Gennaro, L. D. (2006), 'Sleep loss, learning capacity, and academic performance', *Sleep Medicine Reviews* **10**, 323–337.

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

Edwards, F. (2012), 'Early to rise? the effect of daily start times on academic performance', *Economics of Education Review* **31**(6), 970–983.

Goldstein, D., Hahn, C., Hasher, L., Wiprzycka, U. and Zelazo, P. D. (2007), 'Time of day, intellectual performance, and behavioral problems in morning versus evening type adolescents: Is there a synchrony effect?', *Personality and Individual Differences* **42**(3).

Hinrichs, P. (2011), 'When the bell tolls: The effects of school starting times on academic achievement', *Education Finance and Policy* **6**(4), 486–507.

Jacob, B. A. and Rockoff, J. E. (2011), *Organizing schools to improve student achievement: start times, grade configurations, and teacher assignments*, Brookings Institution, Hamilton Project.

Kane, T. J. and Staiger, D. O. (2008), Estimating teacher impacts on student achievement: An experimental evaluation, Working Paper 14607, National Bureau of Economic Research.

National Sleep Foundation (2006), Sleep in america poll: Highlights and key findings, Technical report, National Sleep Foundation.

Persson, J., Welsh, K. M., Jonides, J. and Reuter-Lorenz, P. A. (2007), 'Cognitive fatigue of executive processes: Interaction between interference resolution tasks', *Neuropsychologia* **45**(7), 1571–1579.

Pope, N. (2015), 'How time of day affects productivity: Evidence from school schedules', *The Review of Economics and Statistics* **98**(1).

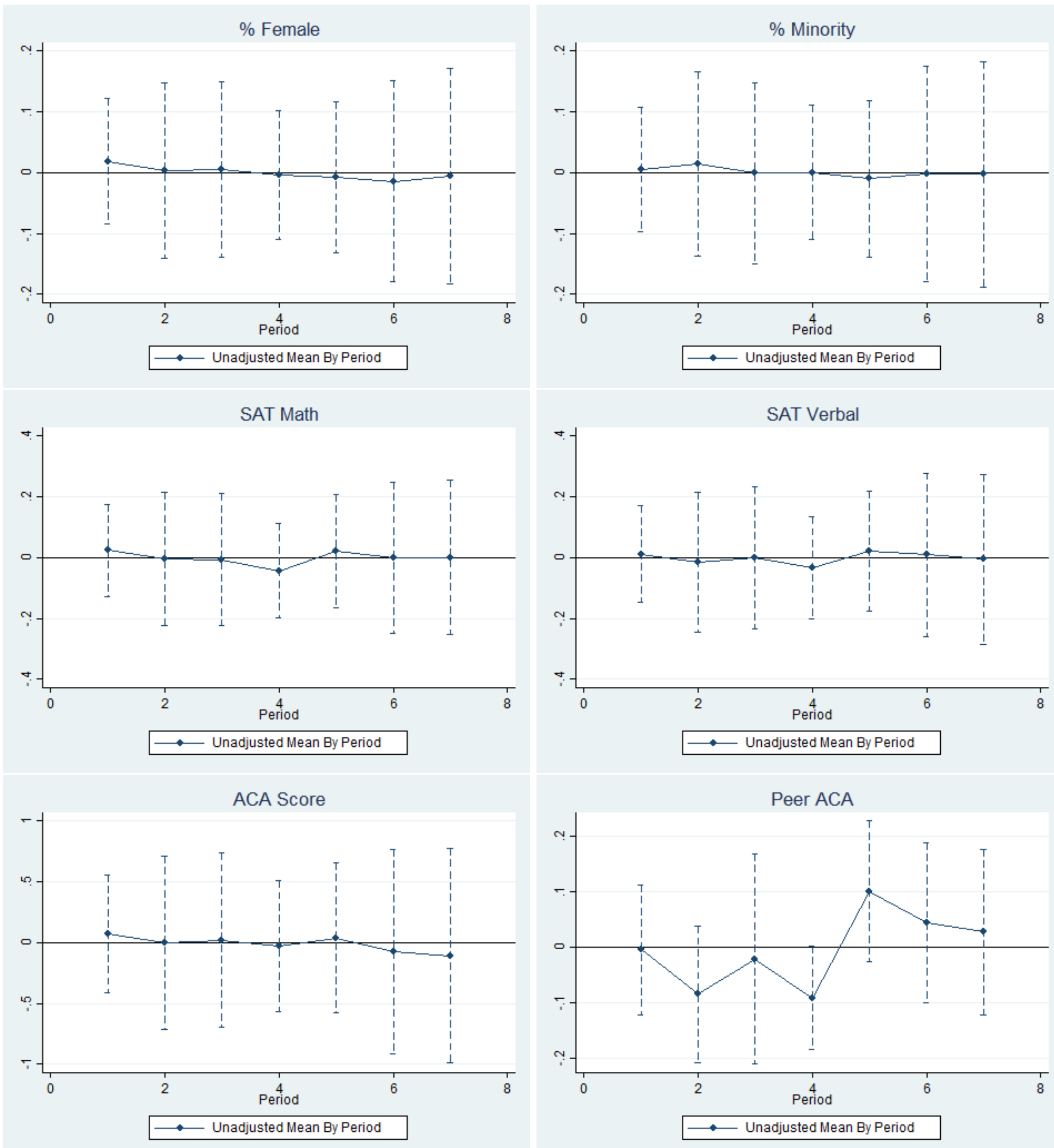
Schmidt, C., Collette, F., Cajochen, C. and Peigneux, P. (2007), 'A time to think: circadian rhythms in human cognition', *Cognitive Neuropsychology* **24**(7), 755–789.

Trocket, M., Barnes, M. and Egget, D. (2000), 'Health-related varibales and academic performance among first-year college students: Implications for sleep and other behaviors', *Journal of American College Health* **49**.

Wahlstrom, K. (2002), 'Changing times: Findings from the first longitudinal study of later high school start times', *NASSP Bulletin* **86**(633).

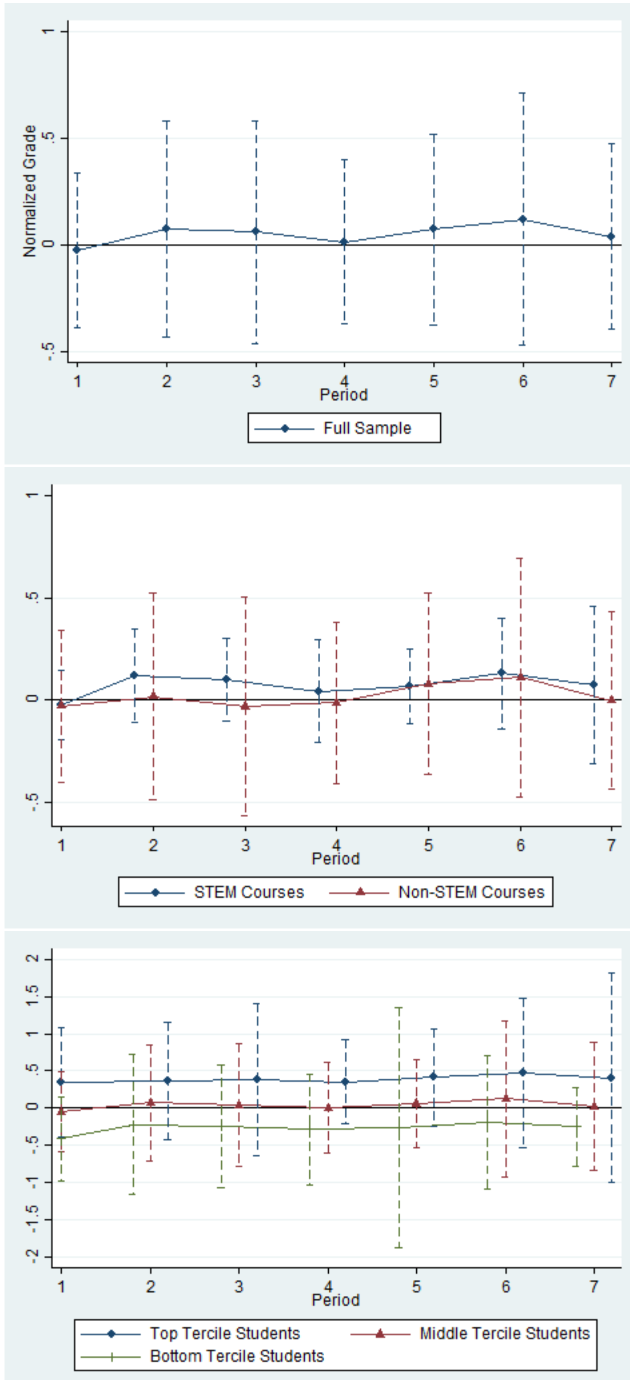
Wolfson, A. and Carskadon, M. (1998), 'Sleep schedules and daily functioning in adolescents', *Child Development* **69**(4).

Figure 1: Randomness Checks



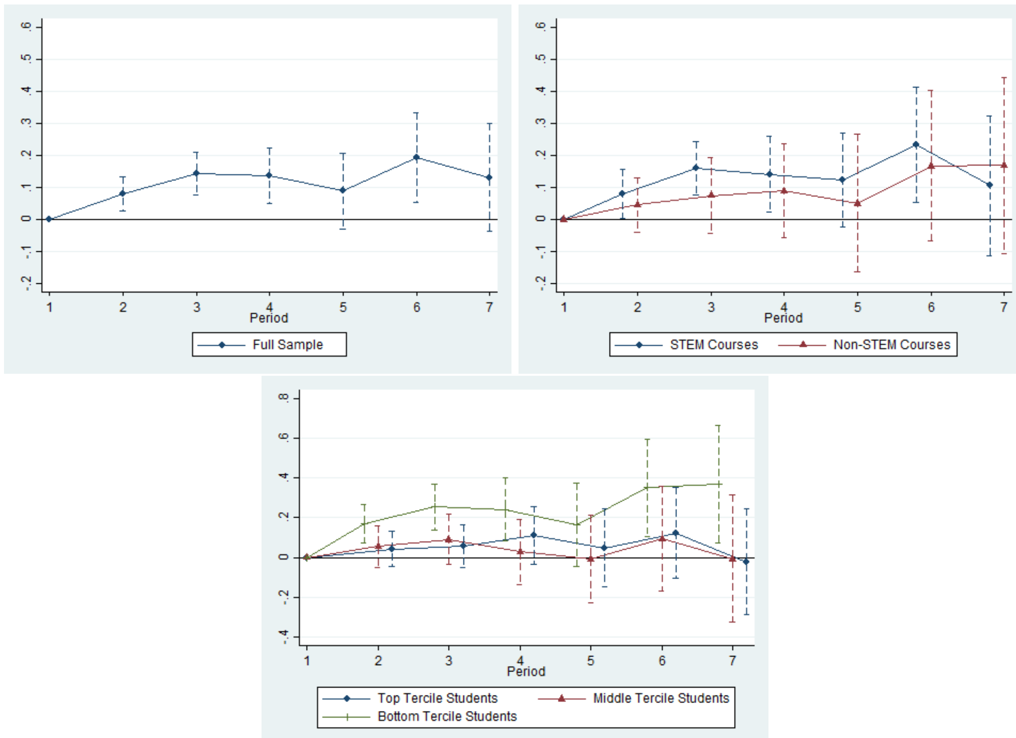
Note:The figures above show the regression coefficients on the class period dummy variables when regressing each of the above background characteristics on class period dummy variables and course-semester fixed effects. The 90% confidence intervals are shown.

Figure 2: Mean Normalized Grades Across Class Periods



Note:The figures above show the regression coefficients on the class period dummy variables when regressing each of the above background characteristics on class period dummy variables and course-semester fixed effects. The 90% confidence intervals are shown.

Figure 3: Plotted Regression Coefficients
Outcome: Normalized Grade



Note:The figures above show the regression coefficients on the class period dummy variables from Equation 1. These estimates are also shown in Table 4. The 90% confidence intervals are shown.

Figure 4: Effects of Preceding Class

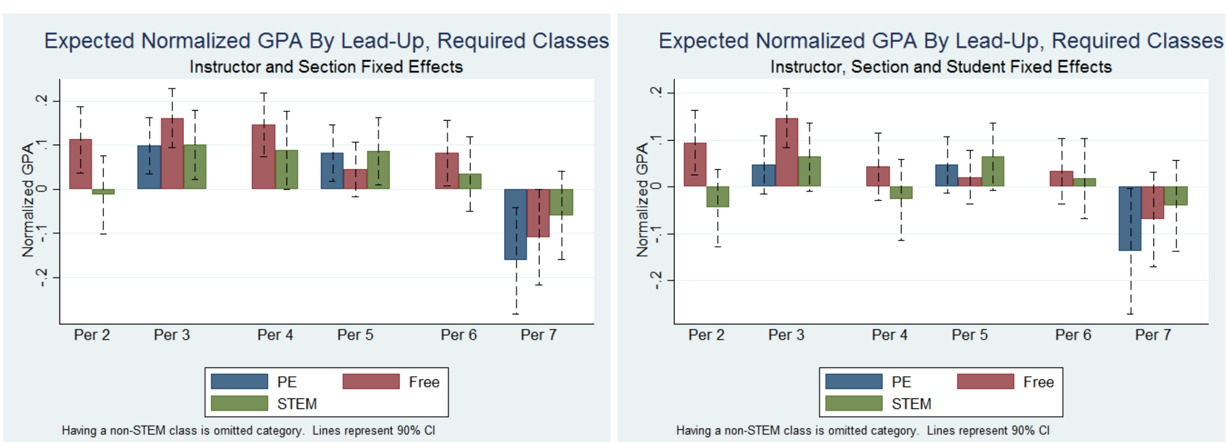


Figure 5: GPA distribution before and after homogenous simulation

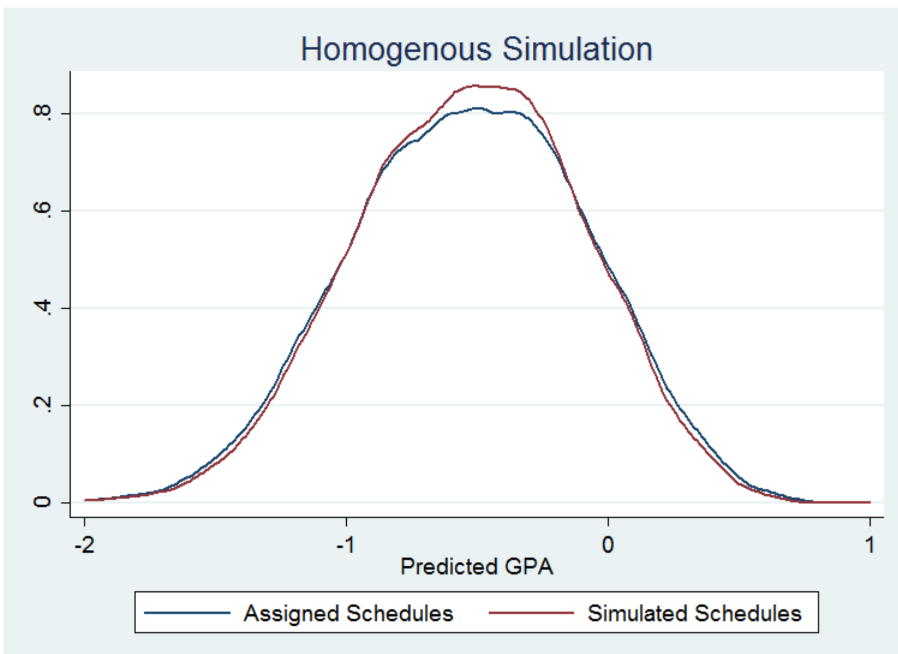


Table 1: **Daily Class Schedule at the U.S. Air Force Academy**

Period	AY1996 - AY2005	AY2006	AY2007 - AY2009
1	7:30	7:00	7:50
2	8:30	8:05	8:50
3	9:30	9:10	9:50
4	10:30	10:15	10:50
5	13:00	13:00	13:30
6	14:00	14:05	14:30
7	15:00	15:10	15:30

Table 2: Student Summary Statistics by Subgroup

	Student-Course		Student		Core		STEM		High		Middle		Low	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Black	0.0363	(0.187)	0.0372	(0.189)	0.0374	(0.190)	0.0340	(0.181)	0.0318	(0.176)	0.0407	(0.198)	0.0394	(0.195)
Hispanic	0.0816	(0.274)	0.0822	(0.275)	0.0816	(0.274)	0.0778	(0.268)	0.0829	(0.276)	0.0722	(0.259)	0.0897	(0.286)
Asian	0.0939	(0.292)	0.0939	(0.292)	0.0941	(0.292)	0.0958	(0.294)	0.0958	(0.294)	0.0888	(0.284)	0.0977	(0.297)
Female	0.190	(0.393)	0.192	(0.394)	0.193	(0.394)	0.189	(0.391)	0.196	(0.397)	0.217	(0.413)	0.165	(0.371)
Prep School	0.167	(0.373)	0.171	(0.376)	0.171	(0.376)	0.164	(0.370)	0.190	(0.393)	0.140	(0.347)	0.183	(0.386)
Fitness Level	4.040	(0.902)	4.045	(0.905)	4.040	(0.901)	4.076	(0.910)	4.073	(0.876)	4.012	(0.890)	4.036	(0.936)
Academic Comp.	13.18	(2.013)	13.16	(2.031)	13.16	(2.007)	13.25	(1.974)	15.28	(0.861)	13.41	(0.585)	10.87	(1.105)
Leadership Score	17.34	(1.796)	17.33	(1.797)	17.35	(1.793)	17.38	(1.798)	17.51	(1.763)	17.34	(1.819)	17.19	(1.783)
Sat Verbal	6.460	(0.650)	6.450	(0.656)	6.450	(0.647)	6.478	(0.643)	6.494	(0.710)	6.460	(0.633)	6.397	(0.589)
Sat Math	6.695	(0.633)	6.689	(0.639)	6.676	(0.627)	6.722	(0.624)	6.778	(0.688)	6.660	(0.608)	6.594	(0.569)
Credits/Day	8.602	(2.292)	8.498	(1.129)	8.854	(2.203)	8.782	(2.233)	8.883	(2.147)	8.896	(2.191)	8.786	(2.267)
Consecutive Classes	0.586	(0.880)	0.574	(0.344)	0.631	(0.902)	0.586	(0.870)	0.650	(0.922)	0.633	(0.899)	0.610	(0.884)
Cumulative Classes	1.524	(1.337)	1.499	(0.321)	1.663	(1.338)	1.508	(1.311)	1.684	(1.348)	1.674	(1.340)	1.633	(1.325)
Cumulative Taught	0.735	(0.888)	0.737	(0.406)	0.760	(0.897)	0.668	(0.869)	0.702	(0.859)	0.758	(0.892)	0.817	(0.932)
Consecutive Taught	0.326	(0.584)	0.325	(0.239)	0.368	(0.599)	0.289	(0.546)	0.346	(0.584)	0.359	(0.587)	0.398	(0.623)
Grade	0.0399	(0.993)	0.0292	(0.696)	0.0515	(0.996)	0.0683	(0.994)	0.390	(0.895)	0.0410	(0.971)	-0.263	(1.008)
Ace	0.216	(0.412)	0.213	(0.260)	0.228	(0.419)	0.253	(0.435)	0.360	(0.480)	0.209	(0.407)	0.119	(0.324)
Failed	0.0626	(0.242)	0.0644	(0.142)	0.0745	(0.263)	0.0947	(0.293)	0.0277	(0.164)	0.0686	(0.253)	0.125	(0.331)
Observations	29736		4816		24264		13210		7891		8175		8198	

Note: The mean and standard deviation of each variable are shown in the table above. The observations are at the student-course level, as shown in the first column. The third column shows the statistics when aggregating the data to the student level. The subsequent columns show summary statistics by course and student characteristics.

Table 3: Student Summary Statistics by Class Period

	Period 1		Period 2		Period 3		Period 4		Period 5		Period 6		Period 7	
	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd	mean	sd
Black	0.0412	(0.199)	0.0473	(0.212)	0.0397	(0.195)	0.0401	(0.196)	0.0311	(0.174)	0.0340	(0.181)	0.0236	(0.152)
Hispanic	0.0787	(0.269)	0.0818	(0.274)	0.0837	(0.277)	0.0905	(0.287)	0.0770	(0.267)	0.0814	(0.274)	0.0771	(0.267)
Asian	0.0982	(0.298)	0.0979	(0.297)	0.0890	(0.285)	0.0826	(0.275)	0.0948	(0.293)	0.0953	(0.294)	0.109	(0.312)
Female	0.212	(0.408)	0.196	(0.397)	0.198	(0.399)	0.189	(0.391)	0.185	(0.389)	0.178	(0.383)	0.187	(0.390)
Prep School	0.170	(0.376)	0.180	(0.384)	0.177	(0.381)	0.190	(0.392)	0.157	(0.364)	0.160	(0.367)	0.159	(0.366)
Fitness Level	4.035	(0.888)	4.044	(0.879)	4.074	(0.916)	4.061	(0.902)	4.034	(0.904)	3.998	(0.898)	4.011	(0.924)
Academic Comp.	13.23	(1.996)	13.16	(2.032)	13.18	(2.026)	13.13	(2.031)	13.20	(1.971)	13.09	(1.994)	13.05	(1.997)
Leadership Score	17.38	(1.798)	17.28	(1.774)	17.38	(1.815)	17.36	(1.831)	17.32	(1.802)	17.35	(1.758)	17.34	(1.729)
Sat Verbal	6.461	(0.645)	6.436	(0.660)	6.448	(0.658)	6.416	(0.659)	6.470	(0.629)	6.459	(0.637)	6.445	(0.636)
Sat Math	6.701	(0.645)	6.672	(0.649)	6.669	(0.641)	6.633	(0.622)	6.698	(0.609)	6.676	(0.611)	6.676	(0.599)
Credits/Day	9.050	(2.176)	9.109	(2.129)	8.853	(2.165)	8.906	(2.171)	8.613	(2.305)	8.684	(2.203)	8.864	(2.185)
Consecutive Classes	0	(0)	0.467	(0.499)	1.352	(0.849)	1.275	(1.361)	0	(0)	0.437	(0.496)	1.199	(0.771)
Cumulative Classes	0	(0)	0.467	(0.499)	1.415	(0.782)	1.890	(1.020)	2.473	(1.056)	2.766	(1.007)	3.223	(0.952)
Cumulative Taught	0	(0)	0.510	(0.500)	0.581	(0.680)	0.799	(0.756)	0.954	(0.951)	1.312	(1.010)	1.605	(1.144)
Consecutive Taught	0	(0)	0.516	(0.500)	0.406	(0.686)	0.670	(0.734)	0	(0)	0.464	(0.499)	1.002	(0.752)
Grade	-0.0259	(1.011)	0.0738	(0.981)	0.0586	(1.009)	0.0146	(1.006)	0.0722	(0.987)	0.120	(0.986)	0.0375	(0.974)
Ace	0.216	(0.412)	0.238	(0.426)	0.233	(0.422)	0.221	(0.415)	0.223	(0.417)	0.243	(0.429)	0.219	(0.414)
Failed	0.110	(0.313)	0.0551	(0.228)	0.0855	(0.280)	0.0646	(0.246)	0.0834	(0.277)	0.0503	(0.219)	0.0475	(0.213)
STEM_course	0.669	(0.471)	0.500	(0.500)	0.627	(0.484)	0.440	(0.496)	0.588	(0.492)	0.405	(0.491)	0.503	(0.500)
Observations	3645		3105		4529		3269		4600		3378		1738	

Note: The table above shows the mean and standard deviation of each variable for each class period. The observations are at the student-course level.

Table 4: Primary Analysis

	(1)	(2)	(3)	(4)	(5)	(6)
Period 2	0.0552**	0.0806***	0.0906***	0.123***	0.0797**	0.105***
	(0.0262)	(0.0253)	(0.0284)	(0.0270)	(0.0329)	(0.0308)
Period 3	0.0680***	0.0958***	0.147***	0.181***	0.143***	0.144***
	(0.0246)	(0.0233)	(0.0319)	(0.0299)	(0.0400)	(0.0378)
Period 4	0.0880***	0.117***	0.151***	0.193***	0.136**	0.133***
	(0.0314)	(0.0292)	(0.0393)	(0.0370)	(0.0532)	(0.0503)
Period 5	0.0586	0.0792**	0.100*	0.140***	0.0881	0.0403
	(0.0374)	(0.0352)	(0.0514)	(0.0487)	(0.0720)	(0.0674)
Period 6	0.164***	0.165***	0.228***	0.249***	0.194**	0.125
	(0.0439)	(0.0420)	(0.0562)	(0.0537)	(0.0843)	(0.0793)
Period 7	0.0933*	0.117**	0.180***	0.219***	0.131	0.0704
	(0.0499)	(0.0478)	(0.0623)	(0.0600)	(0.101)	(0.0945)
Consecutive Classes			-0.0655**	-0.0530*	-0.0675**	-0.0553**
			(0.0280)	(0.0271)	(0.0281)	(0.0274)
Cumulative Classes			-0.0326	-0.0469**	-0.0338	-0.0463**
			(0.0236)	(0.0224)	(0.0237)	(0.0225)
Consecutive Squared			0.0154	0.0127	0.0168	0.0143
			(0.0104)	(0.0103)	(0.0105)	(0.0104)
Cumulative Squared			0.00607	0.00944**	0.00612	0.00888**
			(0.00463)	(0.00437)	(0.00464)	(0.00437)
Cumulative Taught					-0.0140	0.0468
					(0.0370)	(0.0346)
Consecutive Taught					0.0618	-0.00242
					(0.0441)	(0.0421)
Teach Cumul Squared					0.0104	0.00199
					(0.00958)	(0.00897)
Teach Consec Squared					-0.0346	-0.0145
					(0.0222)	(0.0213)
Teacher FEs	Y	Y	Y	Y	Y	Y
Individual FEs	N	Y	N	Y	N	Y
N	22600	22600	22600	22600	22445	22445
R2	0.253	0.680	0.254	0.681	0.255	0.682

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Note: The table above shows the estimates from Equation 1 when including the variables listed above. All regressions include controls for student characteristics and classroom peer effects as well as course by year by schedule-day fixed effects. Standard errors are clustered by student.

Table 5: Subgroup Analysis

	(1) All	(2) Top Tercile	(3) Mid Tercile	(4) Low Tercile	(5) STEM	(6) Non-STEM	(7) Ace	(8) Fail
Period 2	0.0797**	0.0425	0.0555	0.168***	0.0801*	0.0456	0.0193	-0.0344***
	(0.0329)	(0.0532)	(0.0634)	(0.0585)	(0.0463)	(0.0519)	(0.0142)	(0.00992)
Period 3	0.143***	0.0561	0.0908	0.254***	0.159***	0.0745	0.0404**	-0.0502***
	(0.0400)	(0.0650)	(0.0769)	(0.0699)	(0.0497)	(0.0723)	(0.0171)	(0.0131)
Period 4	0.136**	0.111	0.0280	0.243**	0.140**	0.0889	0.0598***	-0.0413**
	(0.0532)	(0.0893)	(0.0994)	(0.0944)	(0.0713)	(0.0889)	(0.0227)	(0.0166)
Period 5	0.0881	0.0472	-0.00713	0.164	0.122	0.0502	0.0397	-0.0306
	(0.0720)	(0.119)	(0.134)	(0.128)	(0.0884)	(0.130)	(0.0314)	(0.0236)
Period 6	0.194**	0.123	0.0930	0.351**	0.233**	0.167	0.0693*	-0.0410
	(0.0843)	(0.138)	(0.159)	(0.148)	(0.109)	(0.143)	(0.0365)	(0.0266)
Period 7	0.131	-0.0220	-0.00584	0.368**	0.105	0.168	0.0400	-0.0441
	(0.101)	(0.161)	(0.193)	(0.180)	(0.132)	(0.167)	(0.0430)	(0.0310)
Consecutive Classes	-0.0675**	0.0163	-0.0982*	-0.110**	-0.0541	-0.0751*	0.00476	0.0147*
	(0.0281)	(0.0477)	(0.0517)	(0.0525)	(0.0392)	(0.0411)	(0.0122)	(0.00752)
Cumulative Classes	-0.0338	-0.0590	0.00942	-0.0493	-0.0525*	0.0110	-0.0168*	0.00391
	(0.0237)	(0.0402)	(0.0417)	(0.0430)	(0.0299)	(0.0347)	(0.00987)	(0.00619)
Consecutive Squared	0.0168	-0.00277	0.0183	0.0355*	0.0106	0.0216	-0.00384	-0.00284
	(0.0105)	(0.0169)	(0.0198)	(0.0202)	(0.0149)	(0.0154)	(0.00446)	(0.00284)
Cumulative Squared	0.00612	0.00939	-0.00147	0.0129	0.00947	-0.00270	0.00396**	-0.000913
	(0.00464)	(0.00772)	(0.00817)	(0.00851)	(0.00597)	(0.00685)	(0.00196)	(0.00116)
Teacher FEs	Y	Y	Y	Y	Y	Y	Y	Y
Teacher Fatigue	Y	Y	Y	Y	Y	Y	Y	Y
Individual FEs	N	N	N	N	N	N	N	N
N	22445	7291	7550	7604	13080	9365	22445	22445
R2	0.255	0.324	0.242	0.238	0.282	0.267	0.220	0.148

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Note: The table above shows the estimates from Equation 1 when including the variables listed above. The outcomes for the regressions shown in columns (1) - (6) is the normalized grade in the course. The outcome for Column (7) is whether the student earned an A or A- in the course while the outcome for Column (8) is whether the student earned a D or F in the course. All regressions include controls for student characteristics and classroom peer effects as well as course by year by schedule-day fixed effects. A full set of teacher schedule variables are included in all regressions, but not shown because they remain statistically insignificant. Standard errors are clustered by student.

Table 6: Robustness Checks

	(1) No Languages	(2) No Chem	(3) 7:00am	(4) 7:30am	(5) 7:50am	(6) Athletes
Period 2	0.0904***	0.0791**	0.119	0.0235	0.109**	0.0589**
	(0.0349)	(0.0371)	(0.0756)	(0.0585)	(0.0513)	(0.0276)
Period 3	0.158***	0.140***	0.291***	0.124**	0.107	0.117***
	(0.0409)	(0.0525)	(0.105)	(0.0587)	(0.0717)	(0.0344)
Period 4	0.150***	0.135**	0.357**	0.124	0.0630	0.127***
	(0.0549)	(0.0666)	(0.151)	(0.0754)	(0.0961)	(0.0460)
Period 5	0.122*	0.0878	0.363	0.109	-0.0184	0.0764
	(0.0735)	(0.0964)	(0.239)	(0.0981)	(0.131)	(0.0643)
Period 6	0.250***	0.193*	0.560**	0.190*	0.106	0.175**
	(0.0870)	(0.107)	(0.277)	(0.114)	(0.155)	(0.0741)
Period 7	0.186*	0.124	0.460	0.155	0.0228	0.116
	(0.104)	(0.124)	(0.320)	(0.137)	(0.186)	(0.0887)
Consecutive Classes	-0.0655**	-0.0684**	-0.108*	0.00663	-0.124***	-0.0851***
	(0.0293)	(0.0292)	(0.0626)	(0.0457)	(0.0434)	(0.0242)
Cumulative Classes	-0.0409*	-0.0500*	-0.0798	-0.0608	0.0123	-0.0364*
	(0.0242)	(0.0260)	(0.0502)	(0.0377)	(0.0384)	(0.0220)
Consecutive Squared	0.0164	0.0185*	0.0324	-0.00454	0.0324**	0.0228***
	(0.0111)	(0.0108)	(0.0236)	(0.0171)	(0.0162)	(0.00884)
Cumulative Squared	0.00765	0.00815	0.0141	0.00872	-0.000950	0.00621
	(0.00479)	(0.00500)	(0.0102)	(0.00746)	(0.00734)	(0.00437)
Teacher FEs	Y	Y	Y	Y	Y	Y
Teacher Fatigue	Y	Y	Y	Y	Y	Y
Individual FEs	N	N	N	N	N	N
N	20655	18643	4503	8963	8979	28825
R2	0.256	0.250	0.273	0.253	0.256	0.256

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Note: The table above shows the estimates from Equation 1 when including the variables listed above. All regressions include controls for student characteristics and classroom peer effects as well as course by year by schedule-day fixed effects. Columns (1) and (2) exclude foreign language courses. Columns (3) and (4) exclude student athletes. Columns (5) - (7) limit the sample to each of the start time regimes. Standard errors are clustered by student.

Table 7: Aggregated Coefficients

	Full Schedule	One Break
Period		
1	0	0
2	0.012	0.104
3	-0.024	0.106
4	-0.092	0.075
5	-0.043	-0.058
6	-0.065	-0.043
7	-0.069	-0.009

Estimates are taken from Column (6) of Table 4 and represent predicted impacts of course time and schedule fatigue on student GPA. The earliest class is normalized to 0. *The first column assumes student has a full schedule with no breaks. The second column shows predictions for a student who has one free period, which is assumed to be in the prior period for each estimate (i.e. number of consecutive classes is always assumed to be 0)*

Table 8: Simulation Results: Homogenous Schedules

	All			Bottom Tercile		
	Mean	SD	N	Mean	SD	N
$ActualGPA$	0.0	0.451	4,536	-0.585	0.226	1525
$SimulatedGPA^{homo}$	0.0	0.435	4,536	-0.566	0.219	1525
Difference	**0.0**			**0.019**		

Note: The table above shows the estimates from simulations where students were inversely re-assigned schedules based on predicted own-GPA and predicted overall schedule impact. Schedules were assumed to have homogenous impacts across all students using results from Column (5) of Table 4.

Table 9: Simulation Results: Heterogenous Schedules

	All			Bottom		Middle		Top	
	Mean	SD	N	Mean	SD	Mean	SD	Mean	SD
$ActualGPA$	0.0	0.415	4,536	-0.429	0.245	-0.013	0.121	0.467	0.207
$SimulatedGPA^{het}$	0.013	0.397	4,536	-0.396	0.226	-0.007	0.113	0.464	0.199
Difference	**0.013**			**0.033**		**0.006**		**-0.003**	

Note: The table above shows the estimates from simulations where students were inversely re-assigned schedules based on predicted own-GPA and predicted overall schedule impact. Schedule impacts were predicted separately by ability tercile using results from Columns (2)-(4) of Table 5 .