6-1-2015

Cognitive Function and Human Capital Accumulation Across the Day: Evidence from Randomized School Schedules

Teny Maghakian Shapiro
Santa Clara University, tshapiro@scu.edu

Kevin M. Williams

James E. West

COGNITIVE FUNCTION AND HUMAN CAPITAL ACCUMULATION ACROSS THE DAY: EVIDENCE FROM RANDOMIZED CLASS SCHEDULES

Teny Maghakian Shapiro *   Kevin M. Williams†

May 22, 2015

This study examines how variation of within-day cognitive function affects human capital accumulation. Cognitive function, which neurobiologists have found varies widely across the day, has thus far been an important omission in the economics literature. We quantify its role on human capital accumulation using data from five cohorts of college freshman at the United States Air Force Academy, where students face randomized scheduling and a common set of classes and exams. We find clear evidence that daily fluctuations in cognitive function affects academic achievement–a student does 0.25 standard deviations better at her highest observed ability than at her worst. Cognitive function is affected by the time of day that learning takes place, but also importantly, by the context of a student's schedule and the degree of cognitive fatigue at that time of day– students perform 0.05 standard deviations worse if they have back-to-back classes than if they just had a break. Differences in effects along the ability distribution suggest that overall efficiency gains are possible. Prioritizing the schedules of those most impacted by cognitive fatigue would be equivalent to improving their teacher quality by a standard deviation in 40% of offered classes. Findings suggest that a re-organization of students' daily school schedules is a promising and potentially low-cost educational intervention.

JEL: I23, J13, J24
Keywords: human capital, education production function, school start time, course schedule, adolescents

*Santa Clara University; Email: tshapiro@scu.edu; Phone: 408-554-4052;
Address: Lucas Hall; Santa Clara University; 500 El Camino Real; Santa Clara, CA 95053

†University of California, Davis; 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

Under a standard education production function, both human capital accumulation and cognitive function are assumed constant throughout the day. That is, all else equal, a person's ability to learn is assumed to be the same in the morning as in the afternoon or evening. This, however, conflicts with research by biologists and sleep experts who have established that an individual's cognitive function, and therefore ability to learn, varies widely throughout the day (Goldstein et al., 2007). The goal of this study is to determine whether *within-day* fluctuations in individual-specific cognitive function have a significant impact on human capital accumulation. We do so using data from the United States Air Force Academy (USAFA), where students are randomly assigned class schedules– this removes any selection bias that would otherwise make this a difficult question to address.[1] By understanding the importance of these factors, administrators may have the opportunity to improve student outcomes at a very low cost– by reorganizing the time in which courses are offered.

The randomized component of our data allow us to separately identify time-of-day effects from cognitive fatigue effects. This is an important aspect of our study as neurobiologists have shown that cognitive ability is determined by an individual's internal-clock as well as the total amount of mental activity one has already done that day (Persson et al., 2007). We find that, all else equal, the afternoon is the best time of day for student learning, but gains from having a class during a more optimal time of day 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 quarter of a standard deviation better; however, a student in a 2:00 class that follows a full schedule of classes can end

[1]The other reasons comparing student performance at different times of the day is challenging include: the inconsistent measures of student academic achievement within a school, students select teachers, students choose classes at times that are optimal for them, teachers of the same course may not cover the same material or give the same tests, grades may be curved within each section of a course, students in morning classes may help friends cheat on tests in the afternoon and teacher quality may vary over the day. Further, standardized exams, such at the SAT or annual state tests, are often used as a measure of academic outcomes, they do not necessarily reflect learning that happened in a specific class, thus can't be used in this setting.

up performing 0.13 standard deviations *worse* than in the 7:30 class. Even two students in the same class may have differences in expected grade as large as 0.1 standard deviations simply due to their prior schedules. Subgroup analysis reveals that the negative effects of fatigue are more extreme for students in the bottom tercile of predicted aptitude and for STEM classes. Our simulations, which reassign students to schedules, find that we can obtain gains for bottom tercile students that are equivalent to a one standard deviation increase in teacher quality in 40% of courses. We conclude with a discussion of policies, obstacles, costs and benefits facing the implementation of a re-scheduled school day. We 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 relatively low cost, yet highly beneficial, intervention.

In addition to the random assignment of course schedules, there are a number of other institutional characteristics at USAFA that make it an ideal setting to assess the role of cognitive function on academic achievement. First, the grading structure at USAFA is standardized– all professors teaching the same course use the same syllabus, give the same exams during a common testing session, and grade collectively. Second, teachers regularly teach multiple sections of the same course. This allows us isolate to time-of-day effects from factors related to teaching. Finally, the school day at USAFA is split into seven class periods, four before lunch and three after, and runs on an alternating schedule. Thus, the same student has two different class schedules within the same semester. This allows us to assess how a student performs with one schedule relative to *themselves* with a different schedule.

Despite our use of university-level data, we believe the findings from this study are generalizable to other populations as well. Not only is the daily structure of the USAFA schedule very similar to that of the average U.S. middle or high school, but we focus our analysis on freshmen since they are still in their teens. Teens have been the focus of many proposed changes to school scheduling because of their distinct time preferences and its misalignment with traditional school schedules. However, there is no reason to believe that the relationship we establish between cognitive function and learning is unique to teens, thus our findings may be applied more broadly. Finally, 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 cognitive function of the students in our sample would be *more* adversely affected by time of day or fatigue.

2. Background

Cognitive Function and the Education Production Function

The idea that multiple inputs affect these outcomes has been formalized as the education production function. In addition to their innate aptitude, a student's academic outcomes have been shown to also be determined by their teachers, classmates, parental investments, and health, among others.[2] Research from various scientific fields, including neurobiology and cognitive science, finds that an individual's cognitive function, and, therefore, ability to learn, fluctuates throughout the day based on their biological rhythm (Schmidt et al., 2007). Thus, the implicit assumption that the education production function is constant throughout the day may be flawed.

To fully understand how time of day can influence academic achievement, it is important to have a basic understanding of the biology of sleep and wakefulness. 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)).

[2] See (Cawley et al., 2001; Heckman et al., 2006; Rivkin et al., 2005; Carrell et al., 2009; Hoxby, 2000; Krueger and Whitmore, 2001; Cunha and Heckman, 2007; Currie and Stabile, 2006; Currie, 2009) for examples of this work.

The adolescent body does not begin producing melatonin until around 11 p.m. and continues in peak production until about 7 a.m., then stops at about 8 a.m. In contrast, for the average adult, melatonin levels peak at 4 a.m. There are times of the day, independent of sleep, when a person is more and less alert, 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). 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.[3]

We apply these concepts to the education production function as follows: a student's learning/grade in a given class is a function of all the known contributors to the eduction production function and now also the student's cognitive function. Cognitive function is determined of both the time of day the class takes places and the level of cognitive fatigue. We initially assume that cognitive function is included additively to the education production function, but our data allow us to test whether there are any interactive effects between cognitive function and a few of other the inputs.

Related Literature

There are a couple of strands of literature that have assessed the role of time on either academic or workplace outcomes. The first is the literature exploring the effects of class times. A number of studies have assessed the impact of school start times on student achievement. The findings have been mixed. Wahlstrom (2002) and Hinrichs (2011) find no effect from the start time change within the Minneapolis Public School district. Edwards (2012) and Carrell et al. (2011) find positive effects from start time delays on standardized test scores and course grades, respectively. The most

[3]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).

compelling evidence comes from Carrell et al. (2011), who use the same data as this study, and find that grades in classes throughout the entire day benefit from later start times. Another set of studies has looked differential achievement across morning and afternoon classes. Cortes et al. (2012) and Dills and Hernandez-Julian (2008) find that students perform better in classes that meet later in the day, while Pope (2014) concludes that learning actually *decreases* throughout the school day.

The second strand of related literature is about productivity in the workplace. Despite the fact that the cognitive skills used in a work setting may be different than those used when learning, there have been a number of studies assessing differences in productivity across the day (Smith et al., 1994). Folkard and Tucker (2003) find that productivity and safety declines during the night shift and is relatively constant for day shifts and that the likelihood of sustaining an injury is higher at night. Additionally, many studies have found that sleep deprivation in medical residents decreases their performance (Veasey et al., 2002). Economic research on workplace productivity has explored the effects of wages (Charness and Kuhn, 2007), multi-tasking (Coviello et al., 2014), and telecommuting (Dutcher, 2012), among others, but no work has been done on how productivity varies within a given shift.

USAFA's structured and regimented academic environment allows us to contribute a better understanding of how two components of cognitive function–time of day and cognitive fatigue–affect academic achievement. Previous studies have suggested changes to school schedules based on knowledge about adolescents' circadian rhythm and time preferences, but no study has been able to cleanly estimate this relationship.[4] In addition to informing the above literatures, this study adds to the rich literature on education production functions and the inputs into academic achievement. Improving achievement and human capital accumulation has large benefits for both the individual and society (Hanushek and Kimko, 2000; Barro, 2001; Murnane and Willett, 1995). Causal estimates of within-day cognitive function on achievement add another avenue to achieve performance gains.

[4] A full list of difficulties in causal estimations can be found in footnote 1.

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. Students are required to graduate within four years and typically serve a minimum five year commitment as a commissioned officer in the United States Air Force following graduation. 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 USFAFA 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 than ensures geographic diversity.

A number of institutional characteristics at the United States Air Force Academy (USAFA) 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. Figure 2 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.[5] All students are required to attend a mandatory breakfast 25 minutes before first period. Second, students are randomly assigned to all of their courses and instructors. Prior to the start of freshman year, students take placement

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

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. Thus, 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 schedule runs every other day. Thus, the same student has two different class schedules within the same semester.[6] These 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 depending on their practices.

Our dataset consists of 4,816 first-year students from the entering classes of 2004 to 2008. 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, an 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

[6]Language courses are an exception and meet every schedule day at 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.

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 core set of approximately 30 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, and history. We refer to these as the required freshman courses. Because 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, we examine the effects of STEM and non-STEM course timing separately to see if the effects differ across course type.

Tables 1 through 3 show summary statistics for our sample. Our data are at the student-course level. Column (1) of 1 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) 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 2 shows summary statistics by class period. There are some differences across class periods. First, the number of observations for each class period differ, 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. Table 3 shows the number of each core course taught in each class period across our sample years. Chemistry and physical education are both two-period long classes, and thus only offered first, third, and fifth periods.

4. 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 dummies and course-semester fixed effects to capture within-course deviations in characteristics. Figure 4 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. The randomness of student assignment across different class periods allows us to utilize this variation to determine the causal impact of cognitive function – and its distinct components: time of day and number of previous classes– on achievement. 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.

We next look at unconditional raw normalized grades across class periods for all students in our sample in the top panel of Figure 3. These numbers are also summarized in Table 2. The second panel shows unconditional grades for STEM and non-STEM courses separately. Third panel shows grades by the three terciles of academic aptitude. A few patterns emerge. First, grades in first period are among the lowest in the day. Second, there appears to be a dip in grades during 4^{th} and 7^{th} periods and a peak during 2^{nd} and 6^{th} periods. Finally, average performance in STEM courses are generally higher than in non-STEM classes, but they follow a very similar pattern across class periods. These patterns hold across ability groups. Within these patterns are some interesting puzzles. Mean performance in second period is quite strong even though it is at a time asynchronous with adolescents' optimal learning times. Alternatively, fourth period is at a time that is synchronous with adolescents' optimal learning times for learning; however, mean grades in those period 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 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:

$$Grade_{icjtsp} = \alpha + \psi_p + \beta Attribute + \delta_1 X_{ict} + \delta_2 Peers_{cjtsp} + \phi_{cts} + \gamma_{jt} + \rho_i + \epsilon_{icjtsp} \quad (1)$$

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. $Attribute$ is a vector of the schedule characteristics, which we discuss in detail in the following sections. The vector X_{ist} includes the following student characteristics: SAT math and SAT verbal test scores, academic and leadership composites, fitness score, race, gender, and whether s/he attended a military preparatory school. To control for classroom peer effects, we include $Peers_{icjtsp}$, the average pre-treatment characteristics of all students in the section except for individual i.[7] 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 also 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.

Time of day

We first estimate Equation 1 focusing on ψ_p and setting $Attribute = 0$ to measure the "time-of-day effect." The estimates (shown in Columns (1) and (2) of Table 4 and the first panel of Figure 5) tell us how students perform in that class period relative to first period. Student performance is statistically significantly higher in every class period than it is in 1^{st} period. Students taking a class during 7^{th} period, for example, perform about one-tenth of a standard deviations better than in a 1^{st} period class, controlling for all attributes of the student, class, instructor, and peers. While we can't reject that periods 2, 3, 4, 5 & 7 are statistically different from each other, student performance is statistically different in period 6. Including individual fixed effects in Column (2) both increases the magnitude of the point estimates and improves their precision.

Estimates from the time of day analysis show that students perform better in later period classes; however, these coefficients are capturing three things. The first two are the components of cognitive function discussed earlier in the paper: the effect of having to learn at that specific time of day and the accumulating effects of student fatigue from having been in other classes previous to that one. The third is the professor's fatigue or experience up to that point. For instance, in later periods professors may be fatigued from having taught a number of classes prior to that or they

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

may be improving their teaching methods with the increased practice. The institutional setting at USAFA allows us to disentangle these three effects, because neither students nor professors are assigned to classes during all periods on each schedule day. We move next to exploring the student fatigue effect.

Student Fatigue

To assess the effect of student fatigue, we exploit the random variation both in the number of classes a student has had before a given class without a break (consecutive classes) as well as the number of total classes a student has had before a given class period (cumulative classes) at USAFA. 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, in fact, 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.[8] Because of the way classes are organized, students may also face a different "course-load" on each schedule day. Some students have their classes spaced equally across both days while others have a majority of their classes on one day. Since overall course-load can also fatigue students and affect their achievement in each of their classes, we also use the number of course credits the students are taking on a given schedule-day as one of our fatigue measures.

We determine the magnitude of the "student fatigue effect" and the relative importance of cumulative and consecutive classes and course-load by including the $Attributes$ vector in Equation 1. Since there are a number of ways to measure student fatigue, we explore different definitions

[8]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.

of fatigue in each of our specifications. In Column (3) of Table 4 we control for the number of credit hours a student is taking on that schedule day (Credits/Day) and the number of consecutive and cumulative classes they have had up to that specific class. All three fatigue variables are negative, supporting the theory that student performance in classes later in the day is hampered by fatigue. In Column (4) we also add the number of cumulative classes squared and the number of consecutive classes squared to determine whether there are non-linear effects of fatigue. The squared terms are positive, but not statistically different from zero. Since we find no accumulating effects of having more than one consecutive class, we replace the consecutive variables with a dummy variable, "back-to-back," indicating whether that specific class is immediately following another class, to our specification in Column (5). This is our preferred specification.[9] While all three of our fatigue measures (credits per day, cumulative classes, and back-to-back) are negative, only the back-to-back estimate is statistically significant. Having a class immediately following another one decreases performance in the second class by nearly six-hundredths of a standard deviation. Individual fixed-effects are included in Column (6) and again lead to more precisely measured and generally larger effects, indicating that the within-student differences between their classes and different schedule-days are significant.

Once fatigue is accounted for in the regressions, the time of day estimates increase dramatically, suggesting that the original estimates were biased downward by student fatigue. Performance generally increases throughout the school day with a dip during fifth and seventh periods. This is in line with sleep research which finds that adolescent alertness improves throughout the morning, but dips at 1:30 and 3:30 in the afternoon (Carskadon et al., 1993). The sixth period estimate of 0.235 from Column (5) can be interpreted as follows: if a student was taking their first class of their day during sixth period, they would perform nearly a quarter of a standard deviation better than if they were taking the same class during first period. This time-of-day effect is quite striking.

[9]We have tested the non-linearity of the fatigue effect in alternate specifications by including separate dummies for each possible value of consecutive and cumulative classes. Estimates supported that having a back-to-back class hurt student performance but there was no support for the effect increasing linearly.

However, we must interpret this alongside the fatigue estimates to understand how students would perform in a more typical academic setting. Assuming a reasonable schedule where sixth period is a student's second class in a row and sixth of the day, our model in Column (5) actually predicts that a student performs 0.0215 standard deviations *worse* than they would in first period. This explains why Pope (2014) finds that students perform worse in their afternoon classes. Table 12 shows the expected grade in each class period relative to first based on the number of cumulative classes and whether the class is immediately following another one.

Instructor Schedule

The last aspect of school schedules we explore is the "instructor schedule effect." 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.[10] Because of this, we have measures for the number of total cumulative and consecutive classes for the instructors as well. It is unclear, a priori, exactly how instructor schedules should affect student achievement. Teaching is not as cognitively-taxing as learning, but certainly leads to more physical fatigue. While instructors may grow tired as they teach more classes (reflected in a negative effect on student grades), they may also become better at teaching that specific content (reflected in a positive effect on student grades).

To assess the instructor schedule effect on its own, we use the number of consecutive and cumulative classes an instructor has taught as the $Attribute$ vector in Equation 1 as well as cumulative classes taught squared. Table 5 show the estimates from this analysis. We progressively add student fatigue controls and class period fixed effects in Columns (2) and (3), respectively. In Column (4) we instead use indicator variables for the number of cumulative and consecutive classes an instructor has taught. Column (5) includes cumulative taught, the square of cumulative taught, and an indicator variable for whether the instructor has taught back-to-back classes. This

[10] Another interesting aspect of instructors schedules is that some teach multiple courses in the same semester while others teach only one. While we would like to explore this further, instructors at USAFA tend to teach multiple sections of the same course each semester, leaving very little variation for us to exploit.

is our preferred specification. Column (6) also includes individual fixed effects. We find no effect of instructor schedule on student performance, indicating that fluctuations in student performance across the school day is not driven by instructor fatigue or experience.

4.1. Combined Results and Subgroup Analysis

The previous analyses established the schedule characteristics that affect student achievement. In this section we combine our preferred specifications of student fatigue and teacher schedules into the same $Attribute$ vector from Equation 1. Columns (1) and (2) of Table 6 shows the point estimates for the time-of-day and fatigue effects when also controlling for instructor schedule.[11] The time-of-day effect stays qualitatively the same as before including the instructor variables with students performing better later in the day, although a few of the point estimates are no longer statistically significant. Having back-to-back classes continues to have a negative effect on students' grades, as does the number of credits the student is taking on that schedule day.

We next assess whether there exists any heterogeneity of these effects across subgroups of our student population. 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. Generally speaking, stratifying our analyses by subgroups means losing statistical power since we are utilizing less of the data. Columns (3) - (5) of 6 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 only negatively impacted by having back-to-back classes. The bottom tercile students, on the other hand, are quite affected by these attributes. The time-of-day effect for this group is striking. The bottom tercile students perform a quarter of a standard deviation better in a fourth period class than during first period with the effects of the afternoon classes

[11] While included, teacher schedule variables remain insigificant in further specifications and so are not reported

being even larger. The last two columns of the table shows estimates for STEM and non-STEM classes, respectively. It's important to note that there are more observations of STEM classes, since a larger share of USAFA's core classes are STEM. Nonetheless, the time-of-day effects appear to matter more for STEM classes and the point estimates are very similar to that from our preferred specification in Column (1).

To better understand the margins at which student achievement is affected, we assess how daily schedules affect students' likelihood of acing or failing a class. Table 7 shows estimates when repeating the same analysis with the outcome variable being an indicator equal to one if a student earned an A or A- in the course. Table 8 shows the analogous analysis with the outcome variable being an indicator equal to one if a student earned a D or F in the course. We see that both the time-of-day and fatigue effects affect students at both extremes of the grade distribution. Students in a 3^{rd} period class are 4.2 percentage points more likely to earn an A or A- and 4.9 percentage points less likely to earn a D or F compared their counterparts who take the same class during 1^{st} period, all else equal. Columns (3) - (5) show the estimates when stratifying by academic ability. The top tercile students are affected on the A margin – their likelihood of earning an A improves as the day progresses (except 7^{th} period) and that fatigue from having classes throughout the day decreases their likelihood of earning an A. On the failing margin, we find statistically significant time-of-day effects for the middle and bottom tercile students, but only in the morning classes. For example, a bottom tercile student is about ten percentage points less likely to fail their 3^{rd} period class than their 1^{st} period class.

Heterogeneity Across the Day

Next we take a deeper look into the heterogenous effects of course characteristics across the school day. That is, we assess whether there are non-linearities in educational inputs across the school day. To do so, we estimate the following equation, where we interact the class period indicator variables with other course characteristics:

$$Grade_{icjtsp} = \alpha + \psi_p + \beta_h \sum_{p=1}^{7} \psi_p * (Char) + \delta_1 X_{ict} + \delta_2 Peers + \phi_{cts} + \gamma_{jt} + \rho_i + \epsilon_{icjtsp} \quad (2)$$

The estimates of β_h explain the differential effect of that class characteristic across each of the seven class periods. The first class characteristic we assess is peer effects. In this case, the variable $Char$ is the mean academic composite of one's classmates. We also assess the effects of class size and course load across the day. Estimates are plotted along with their 90% confidence intervals in Figure 6. Each variable tends to exhibit the expected sign with peer ability aiding performance and class size and course load harming. However, while a few of the estimates are statistically significant at the ten percent level, we find no strong patterns and cannot conclude that classroom characteristics have varying impact across the day.

Section Fixed Effects

Our main specification identifies the impact of the time of day on learning by leveraging the variation in times the same course is offered. A second approach is to consider only within-section differences in performance. Here, rather than compare two students taking the same class at different times of the day, we only compare students within the same section, but with varying experiences earlier in the day. This is achieved by including section specific fixed effects, rather than course specific ones. For a given section of a class, students have been randomly assigned to the section at hand, but also their preceding schedules. In essence, a student's schedule immediately beforehand can be thought of as a "treatment" on their cognitive function. 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 $LeadUp$ situation and estimate:

$$Grade_{icjtsp} = \alpha + \beta LeadUp_{icjtsp} * \lambda_p + \delta_1 X_{ict} + \delta_2 Peers + \phi_{ctspj} + \gamma_{jt} + \rho_i + \epsilon_{icjtsp} \quad (3)$$

$\beta LeadUp_{icjtsp} * \psi_p$ allows for the effect of a student's prior classes to vary over the day. The section fixed effects, ϕ_{ctspj} replace ϕ_{cts}. We have multiple sets of $LeadUp$ variables, each set is estimated in a separate regression.[12]

The first set categorizes what was on a student's schedule in the period immediately prior. It has four possibilities: Free Period, P.E., STEM Class, Non-STEM Class. Results can be seen in Figure 7. While neither involve an academic course, we distinguish between free periods and physical education to capture differences that physical activity may have on performance later in the day. Each bar in the figure represents a single coefficient β from Equation 3 sorted by period of the day and color-coded by $LeadUp$ scenario with having had a non-STEM course prior as the reference group. The second graph includes individual student fixed effects[13]. Results show just why the coefficient on the back-to-back variable from earlier was consistently negative. Looking at 3^{rd} period, take two students in the same 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 .15 standard deviations higher. P.E. is similarly beneficial to have had 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. This is discussed in tandem with the second set of results below.

The second set of $LeadUp$ variables looks at the prior *two* periods. It has five possibilities: Two Free Periods, P.E., Free then Class, Class then Free and Two Classes with results in 8. Having two classes prior is the omitted category. In 3^{rd} and 4^{th} period courses, having two free periods or P.E. beforehand leads to expected gains of around 0.13sd, even compared to those who only had one free period. Impacts are smaller when looking at 5^{th} period due to all students having lunch immediately beforehand. The Class-Free combination is consistently better than Free-Class

[12]As before, only core freshman courses are considered. Depending on the specification, 1^{st} and 2^{nd} period observations may be dropped due to lack of variation in students schedules prior to these courses

[13]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.

(excepting 5^{th} period) which reinforces the performance hit that comes with having back-to-back courses.

In both sets of graphs, Period 7 is an interesting exception. The free period students from the first set of $LeadUp$ estimations can be seen as a combination of the two free period and class then free group of the second set. Students who have had either two free periods or a P.E. immediately prior to 7^th period perform around 0.15 standard deviations worse than students who had at least one class in periods 5 or 6. One explanation is that these students are mentally "checked out." Lunch, combined with two periods of either P.E. or no class means that students in the P.E. or two-free periods category have had a nearly 3.5 hour break from a classroom. Potentially, after being in academic mode for some portion of the morning, it is difficult for them to take such an extended break and then re-focus for a single afternoon class. Students who have two or more classes in the afternoon (thus likely having had breaks in the morning) are able to perk-up for their classes after lunch.

The inclusion of student fixed effects absorbs variation across students and reduces the magnitude of virtually coefficient in both sets of $LeadUp$ estimates. Having period 2 free remains a significant benefit to period 3 classes. While statistical significance is reduced, patterns are similar, with the effect of P.E. and free periods again flipping for the end of the day.

Figure 9 shows to more, distinct sets of $LeadUp$ possibilities. The top graph compares students in the same sections based on the total cumulative classes up to that point. The bottom graph compares them based on the number of consecutive classes. Here, the possible values of $LeadUp_{icjtsp}$ change depending on the period.[14] For cumulative classes, the omitted group is those having their first class of the day. The patterns before lunch are as expected, with students having their first class outperforming their peers. By the afternoon, those who have not had any classes yet (only

[14]For example, looking at cumulative classes, for 2^{nd} period $LeadUp \in (0, 1)$ since having a 1^{st} period class was the only chance to accumulate. For 4^{th} period $LeadUp \in (0, 1, 2, 3)$. For the consecutive graph, 5^{th} period is empty since lunch is considered a break for everyone.

3% of students) seem to perform worse, but estimates are insignificant. The lack of significance suggests that by the afternoon, students' immediate history (one or two periods prior) matters more than the collection of fatigue over the whole day.

The bottom graph of Figure 9 defines $LeadUp$ as the number of consecutive classes up to that point. Estimates show that, up until 7^{th} period, students who are not in a back-to-back class are consistently earning grades approximately .1sd higher than those who are in a back-to-back class. In 7^{th} period, having a back-to-back class becomes a positive, but insignificant, likely for reasons discussed above.

5. Robustness Checks

We verify the robustness of our estimates to several changes in model specification with results shown in Table 10. Columns (1) and (2) exclude 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. Columns (3) and (4) shows our model with the inclusion of recruited athletes, a group whose class assignment may not be as random since they are generally not assigned afternoon courses. Finally, we verify that the results are not purely driven by one of the start-time regimes. Column (5) shows the model restricted to academic year 2007, when first period started at 7:00 a.m. Column (6) is limited to academic years 2005 and 2006, when first period started at 7:30, and the last column 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. Table 11 shows estimates from our model as we sequentially add controls and fixed effects. 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 or from our choice of model specification.

Simulations

Our main results focus on the impact timing and fatigue have on student grades in individual classes. In Appendix A, we present simulations using a student's whole schedule as the unit of observation. The simulations allow us to estimate overall gains in performance that re-scheduling students could yield. Results from Section 4 are used to measure a schedule's predicted impact on a student's grade across all their classes. Two simulations are performed, one where schedule impacts are assumed to be homogenous across students and the other where we allow for heterogenous impacts based on student ability. In both cases, we re-assign schedules such that students with low predicted GPAs are given the best-performing schedules and the top predicted 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 USAFAs current needs of factors such as faculty size and classroom availability over the day. When schedule impacts are assumed homogenous, re-assigning students to schedules represents a tightening of the overall GPA distribution with no change in average performance. We estimate that, were schedules and predicted student GPA perfectly inversely related, overall variance in grades would decrease by around 8%. Bottom quartile students experience a 2% of a standard deviation increase in overall performance, but a similar loss is predicted for the top quartile. If we allow for heterogeneous effects of schedules based on predicted terciles then there is a chance to both tighten the overall GPA distribution, thus increasing equality, and also raise mean performance, increasing overall aptitude. Simulations show that expected performance rises by 1.2% of a standard deviation for all students. These gains are concentrated in the bottom tercile where students see and average GPA increase of 3.8%. The median student is enrolled in five core academic courses. If we assume the benefits of schedule reassignment is focused in two of their five courses, the gain is equivalent to increasing teacher quality by a standard deviation in 40% of courses. These estimates are conservative in most respects, but bullish in others. For example, they assume the distribution of classes throughout the day at USAFA is fixed. We know that period 7, an ideal time for learning, is the least used period of the day. If we allowed for the distribution of class times to change we could expect larger predicted performance gains. On the other hand, we are assuming peer effects are constant in these simulations. Giving the predicted bottom tercile students the best schedules may create more overlap in their classes and change the peer composition.

6. Discussion and Conclusions

The goal of this study was to determine whether *within-day* fluctuations in individual-specific cognitive function have a significant impact on human capital accumulation. We find clear results that both the time of day a class is held and the level of cognitive a student faces at the time of their class 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 is hindering 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 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 performing students, the population in our sample most likely to be similar to the average student, see larger effects of scheduling in their grades, especially when faced with multiple classes in a row. This research extends our understanding of the education production function and provides an opportunity to increase academic 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.[15] First, our findings consistently and strongly support start times in the 9am-10am range. However, shifting a school's entire schedule may be expensive to implement or an unpopular policy among parents, teachers, and coaches.[16] 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. [17]

Subgroup results suggest that achievement in classes students struggle in is most vulnerable to the penalties of cognitive function. 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. STEM classes also appear more susceptible to timing and fatigue. This could be due to the nature of STEM classes (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.

While our data is from an academic setting, there are also implications for optimal scheduling outside of the classroom environment. There is a lot of recent interest in understanding the implications of alternative work environments (e.g. telecommuting, multi-tasking, distributed teams). Since each person has their own sleep-wake cycle and optimal times of day, the findings from this study suggest that in more flexible and cognitively-challenging work environments, firms may see

[15]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.

[16]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.

[17]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.

productivity gains from allowing employees to adjust their work schedules to their personal time preference. These relatively low-cost changes for firms and schools can lead to great improvements in productivity and success.

Teny M. Shapiro
Assistant Professor of Economics
Leavey School of Business
Santa Clara University

Kevin M. Williams
PhD. Candidate
Department of Economics
University of California, Davis

A. APPENDIX: Aggregate Simulations

Results from Section 4 quantify the expected impact of various schedule features on an individual student's grade in a single course. Using these results, we are interested in quantifying the aggregate impact of schedules. We perform two simulations that re-assign students to schedules in order to maximize equality, one where schedule impacts are assumed to be homogenous, the other allowing for heterogenous schedule effects. These simulations hold the structure of USAFA's schedule constant and provide conservative estimates of the effects of optimal scheduling.[18]

To prioritize which students get which schedules, we use coefficient estimates from our preferred specification of Equation 1, whose results are shown in Table 6. Using only coefficients from student background characteristics from the specification in Column 1 (our preferred specification), we predict students' expected grade for each of the 22,445 individual course observations, $Stud\hat{G}PA_i$. $Stud\hat{G}PA_i$ does not depend on a student's particular schedule, teacher or choice of classes, it is a measure of predicted aptitude.

Next we predict the impact of scheduling on all core classes of all students. Again using results from Table 6.[19] There are two sets of predictions: homogenous and heterogenous. The homogemous prediction uses coefficients from Column 1 and assumes equal impact of class schedules across all students, denoted $SchedAll_{if}$, where f represents our full set of fatigue variables. Heterogenous impacts use coefficients from Columns 3,4 and 5 to construct $SchedTop_{if}$, $SchedMid_{if}$ and $SchedBot_{if}$ which represent the predicted impact of a course's schedule characteristics on a top, middle and bottom tercile student, respectively. These constructed variables are independent of the actual student who was assigned to a given course. The heterogenous set of predicted values were obtained from regressions on a subset of the data, but the predicted values are made for all

[18] Presumably, altering the distribution of how many classes are offered each period (i.e. shifting a greater percentage of classes to the afternoon) could lead to results of a larger magnitude. However, we do not have knowledge of teacher contracts, classroom availability and other factors involved in scheduling so we limit ourselves to re-assigning students to the existing set of USAFA schedules.

[19] Coefficients include period dummies for periods 2-6, back-to-back class, cumulative classes and cumulative squared.

22,445 observations.

While all analysis and results from 4 was done at the individual-by-course level, we are interested in aggregate schedule impacts. Hence, we convert the 22,445 individual course observations into schedules. After dropping students that take abnormally few core classes or have an unusual number of free periods, we are left with 4,536 schedules.[20] Each schedule contains values constructed from the prior predictions. $Stud\hat{G}PA_i$, the predicted average gpa of student i is unchanged. $\overline{SchedAll_{if}}$ represents the average predicted impact of i's schedule across all their core courses if students are treated homogenously. Figure 10 shows there is a weak, insignificant relationship between predicted student GPA and the predicted average schedule impact. $\overline{SchedTop_{if}}$, $\overline{SchedMid_{if}}$ and $\overline{SchedBot_{if}}$ indicate how we would expect each schedule to impact a top, middle or bottom tercile student.

A.1. Homogeneous Impacts

In the homogenous simulation we assign the best schedule (highest value of $\overline{SimAll}$) to the student with the lowest predicted performance, the second-best schedule to the second-lowest student and so forth. We compare $G\hat{P}A_{if}$ the sum of student predicted GPA, $\overline{StudGPA_i}$, with their actual schedule $\overline{SimAll_{if}}$ to $Sim\hat{G}PA_{if}^{homo}$, the sum of a student's predicted gpa to their assigned schedule based on rank.

This simulation results are reported in Table 13 and represents a zero-sum tightening of the distribution. The predicted lowest students are helped and the predicted best students are assigned the worst schedules. The tightening creates a %4.0 decrease in the sd or a %7.9 reduction in total

[20]At USAFA each student has an "M" and a "T" day schedule, each containing seven instructional periods so a schedule consists of fourteen periods. Core academic classes are all treated equally a single period can be in one of three states: *Core Class* (where the grade is used analysis and the class counts for purposes of student fatigue), *Non-Core Class* (grade is not used, but class counts for fatigue) or Free/PE (grade not used, not counted for fatigue). For our simulation we choose to limit ourselves to the set of existing schedules in our data rather than modeling the millions of theoretically possible USAFA schedules that are not observed. Limiting analyses to the set of observed schedules allows us to more easily look at hypothetical rearranging of schedules while still working within teaching and classroom constraints. 97% of students have between 5-8 total free periods, the rest are dropped. 96% of remaining schedules contain 4,5 or 6 core freshman courses. Those taking fewer have tested in to multiple upper-level courses and are dropped.

variance. Mean performance is unchanged. Figure 11 shows density plots.

Since much of the emphasis on in educational research is focused on raising the bottom of the distribution, we can also look at the impact on only the bottom quartile of students. For these students, the difference between$G\hat{P}A_{if}$ and $Sim\hat{G}PA_{if}^{homo}$ is 0.022. The median student is taking five core courses and the outcome was standardized at the course level. Thus the difference in the two numbers could be thought of as a small increase in all of a students courses or perhaps as a 0.11 standard deviation increase in one course, about equivalent to increasing one of the student's teacher's quality by a standard deviation.[21]

Heterogenous Impacts

Here we use the schedule impacts that were predicted from Columns 3,4 and 5 of Table 6. Bottom tercile students are inversely assigned the top schedules based on $\overline{SchedBot_{if}}^r$ across all 4,536 schedules. Middle tercile students are then inversely assigned from the remaining schedules based on $\overline{SchedMid_{if}}^r$. Top tercile students are inversely assigned the remaining third of schedules (which may actually be above average for them, despite being poor options for bottom or middle tercile students) based on $\overline{SchedTop_{if}}^r$. For each schedule, $Sim\hat{G}PA_{if}^{hetro}$ is the combinations of student's GPA based on own characteristics plus their assigned schedule.

Results are reported in Table 14. Allowing schedules to impact student heterogeneously leads to larger gains from schedule re-assignment. The overall variance grades decreases by 11.8%. Among all students, there is a 0.012 gain in average grades. This benefit is concentrated among the bottom tercile students, who see a 0.038 increase. For a median student with five core courses, this is equivalent to a .4 standard-deviation increase in teacher quality across all classes, or a 1 standard devation increase in two out of five classes. Middle tercile students receive a small increase of, on average, 0.01 standard deviations per class. Top tercile students, whose grade were shown in

[21] Chetty et al. (2014), Kane and Staiger (2008) and Carrell and West (2010) find a 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

Table 6 to be more robust to scheduling, are assigned the worst schedules, but only experience an average decline in grades of -0.013, equivalent to reducing teacher quality by two-thirds of a standard deviation in one out of five classes. Thus the gains for the bottom students are around three times as large as the costs for the top ones.

References

Barro, R. J. (2001), 'Human capital and growth', *American Economic Review* pp. 12–17.

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., Fullerton, R. L. and West, J. E. (2009), 'Does your cohort matter? measuring peer effects in college achievement', *Journal of Labor Economics* **27**(3), 439–464.

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).

Cawley, J., Heckman, J. and Vytlacil, E. (2001), 'Three observations on wages and measured cognitive ability', *Labour Economics* **8**(4), 419–442.

Charness, G. and Kuhn, P. (2007), 'Does pay inequality affect worker effort? experimental evidence', *Journal of Labor Economics* **25**(4), 693–723.

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**.

Coviello, D., Ichino, A. and Persico, N. (2014), 'Time allocation and task juggling', *The American Economic Review* **104**(2), 609–623.

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

Cunha, F. and Heckman, J. (2007), 'The technology of skill formation', *AEA Papers and Proceedings* **97**(2), 31–47.

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

Currie, J. (2009), 'Healthy, wealthy, and wise: Socioeconomic status, poor health in childhood, and human capital development', *Journal of Economic Literature* **47**(1), 87–122.

Currie, J. and Stabile, M. (2006), 'Child mental health and human capital accumulation: the case of adhd', *Journal of health economics* **25**(6), 1094–1118.

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

Dutcher, E. G. (2012), 'The effects of telecommuting on productivity: An experimental examination. the role of dull and creative tasks', *Journal of Economic Behavior & Organization* **84**(1), 355 – 363.

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

Folkard, S. and Tucker, P. (2003), ‘Shift work, safety and productivity’, *Occupational medicine* **53**(2), 95–101.

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).

Hanushek, E. A. and Kimko, D. D. (2000), ‘Schooling, labor-force quality, and the growth of nations’, *American economic review* pp. 1184–1208.

Heckman, J. J., Stixrud, J. and Urzua, S. (2006), ‘The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior’, *Journal of Labor Economics* **24**(3), 411–482.

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

Hoxby, C. M. (2000), ‘The effects of class size on student achievement: New evidence from population variation’, *Quarterly Journal of economics* pp. 1239–1285.

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.

Krueger, A. B. and Whitmore, D. M. (2001), ‘The effect of attending a small class in the early grades on college-test taking and middle school test results: Evidence from project star’, *The Economic Journal* **111**(468), 1–28.

Murnane, R. and Willett, J. (1995), 'The growing importance of cognitive skills in wage determination.', *Review of Economics & Statistics* **77**(2), 251–266.

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. (2014), 'How time of day affects productivity: Evidence from schoolschedules', *Working Paper* .

Rivkin, S. G., Hanushek, E. A. and Kain, J. F. (2005), 'Teachers, schools, and academic achievement', *Econometrica* **73**(2), 417–458.

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.

Smith, L., Folkard, S. and Poole, C. (1994), 'Increased injuries on night shift', *The Lancet* **344**(8930), 1137–1139.

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**.

Veasey, S., Rosen, R., Barzansky, B., Rosen, I. and Owens, J. (2002), 'Sleep loss and fatigue in residency training: a reappraisal', *Jama* **288**(9), 1116–1124.

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).

Figures and Tables

Figure 1: Example Student Course Schedule at the U.S. Air Force Academy

Day	Course	Description	Hours	Period Section	Room	Instructor
M	History 100	INTRO TO MILITARY HISTORY	3.00	M1B	5D37	[redacted]
	Math 141	CALCULUS I	3.00	M2D	5D12	[redacted]
	ReadSkls 103	READING ENHANC./4TH CL	2.00	M4A	1A78	[redacted]
	Russian 131	BASIC RUSSIAN	3.00	M5A	4H18	[redacted]
	English 111	INTRO/COMPOSITION & RESEARCH	3.00	M6B	4D6	[redacted]
	ExtProg 917	INTRAMURALS/GROUPS 1/2	0.00	M7A		
T	Chem 100	APPLICATIONS OF CHEMISTRY I	3.00	T1B T2B	2M117	[redacted]
	PhyEd 110D	BOXING	0.50	T3A T4A		
	Russian 131	BASIC RUSSIAN	3.00	T5A	4H18	[redacted]
U	FYE 101B	FIRST YEAR EXPERIENCE	1.00	U1T		[redacted]
	Cmsng Edu 100	4CL COMMISSIONING EDUCATION	0.00	U1A		

Figure 2: **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

Figure 3: Mean Normalized Grades Across Class Periods

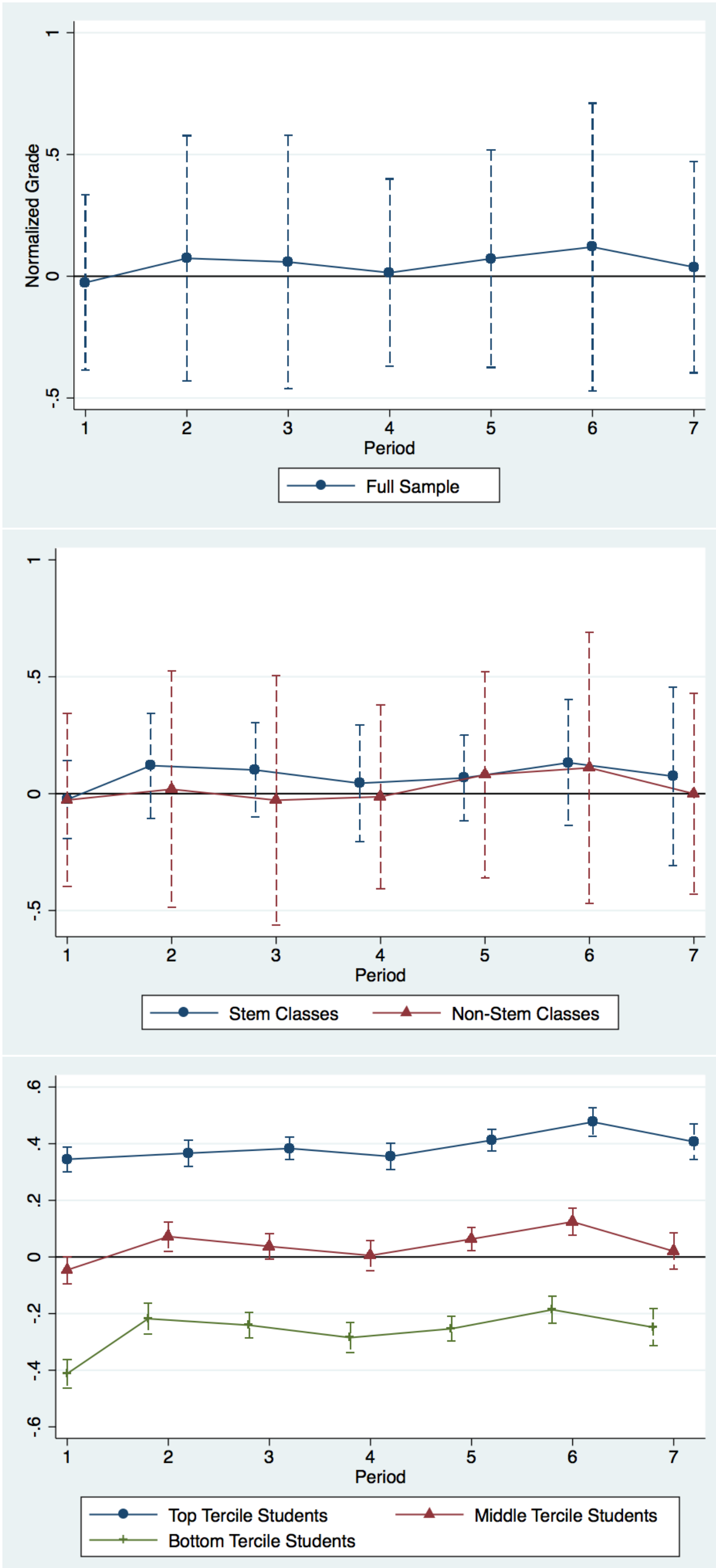
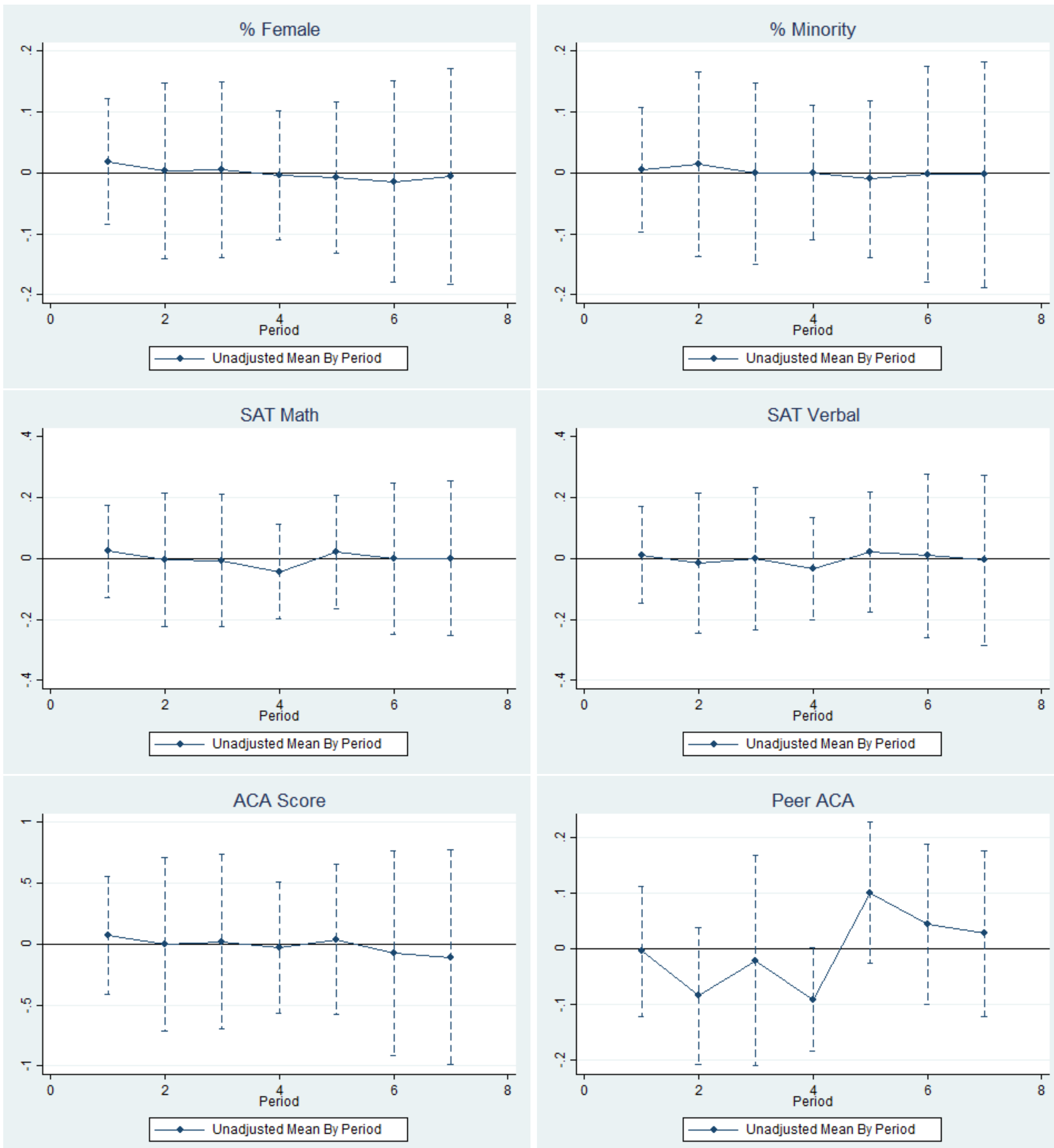
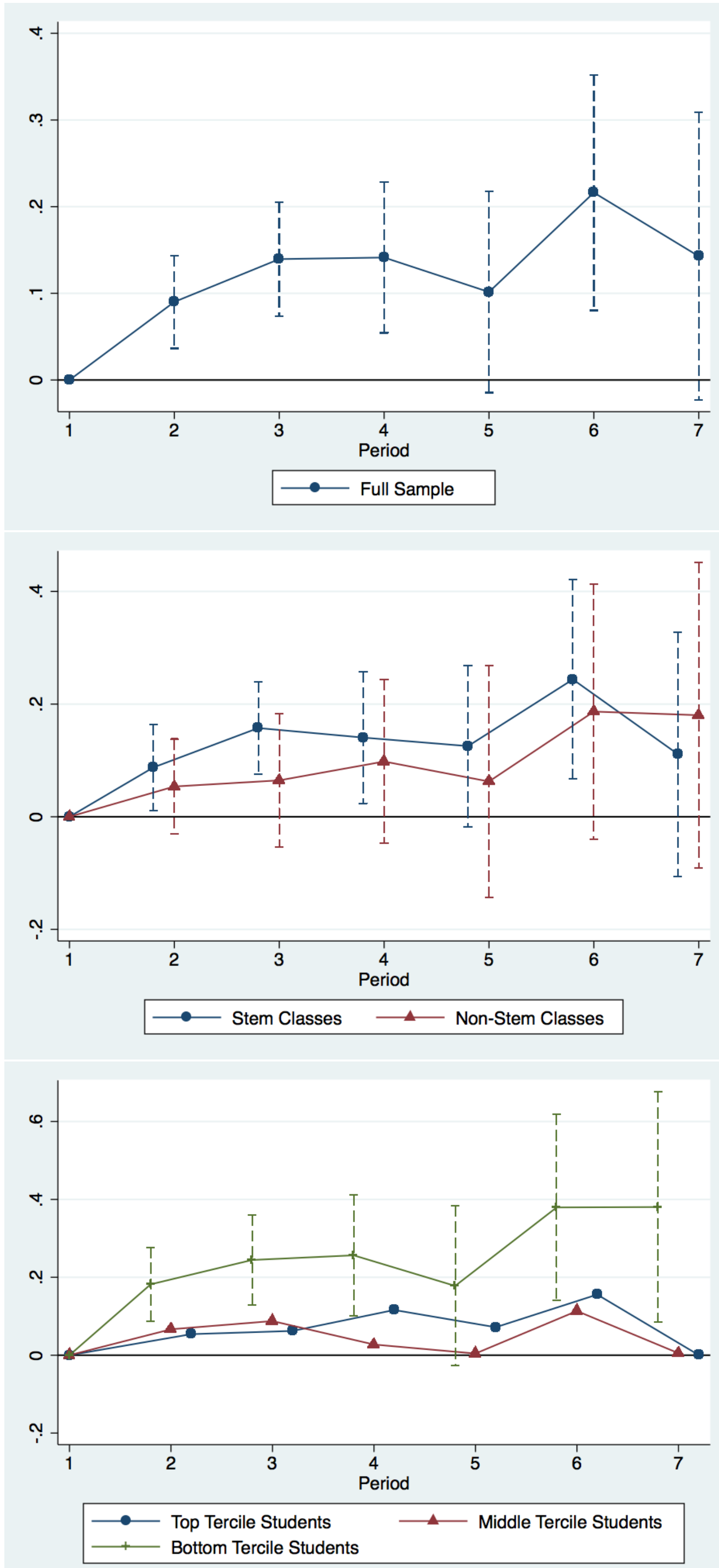


Figure 4: Randomness Checks



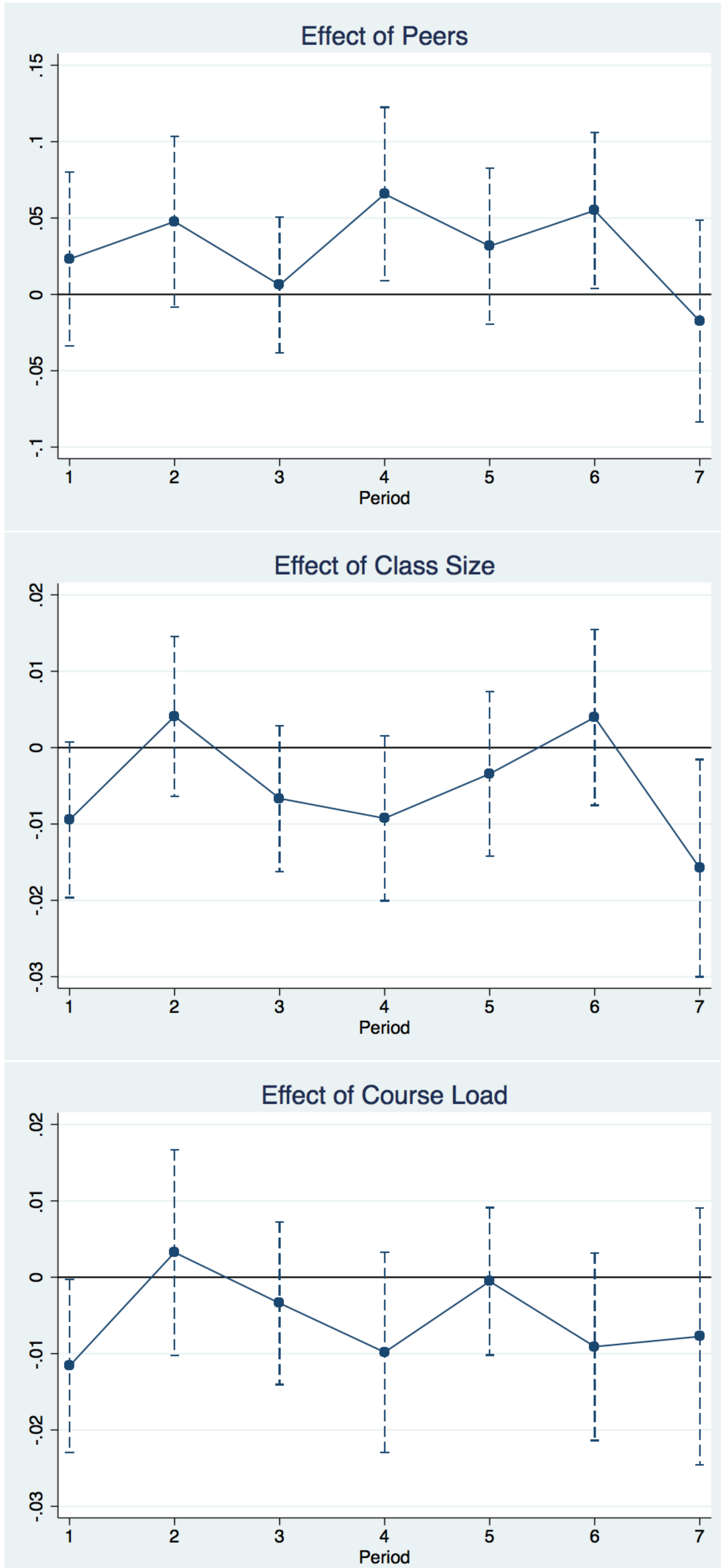
Note:The figures above show the regression coefficients on the class period dummy variables when regressing each of the above background characteristics on class period dummy variables and course-semester fixed effects. The 90% confidence intervals are shown.

Figure 5: Plotted Regression Coefficients
Outcome: Normalized Grade



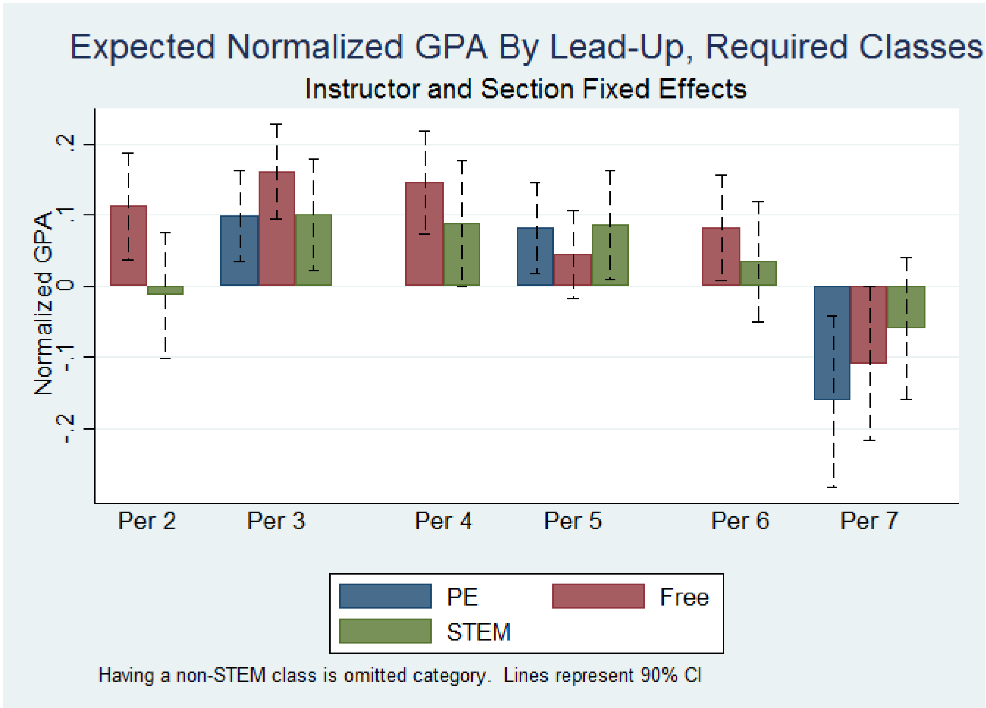
Note:The figures above show the regression coefficients on the class period dummy variables from Equation 1. These estimates are also shown in Table 6. The 90% confidence intervals are shown.

Figure 6: The Effect of Peers, Class Size, and Course Load Across the Day
Outcome: Normalized Grade



Note:The figures above show the regression coefficients on the class period dummy variable and course characteristic interaction, β_h, from Equation 2. The 90% confidence intervals are shown.

Figure 7: Effects of Preceding Class



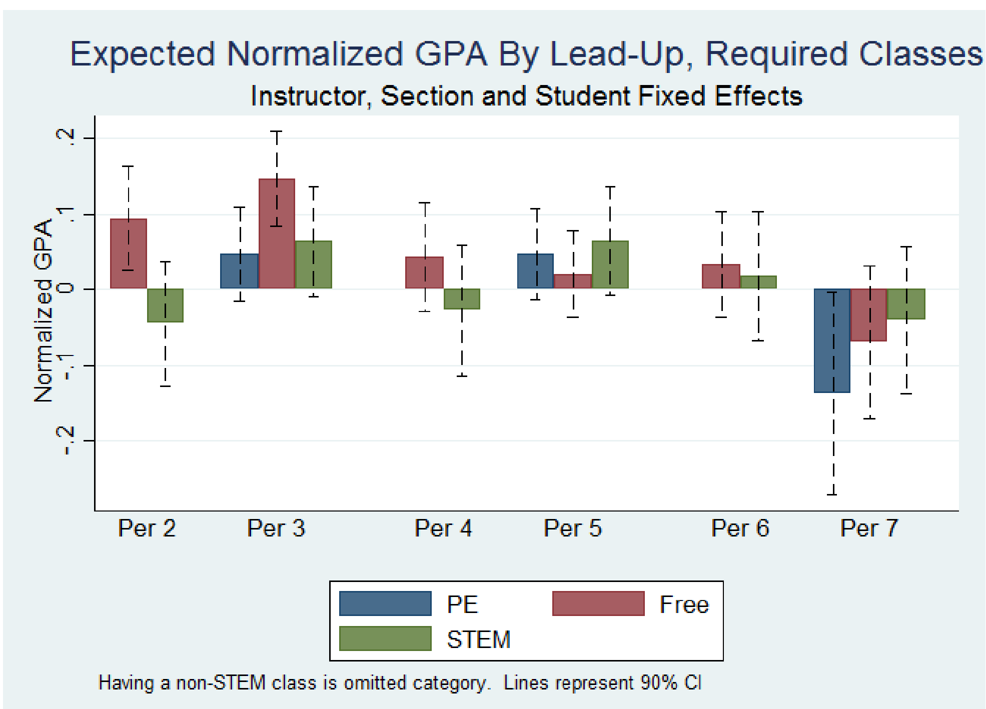
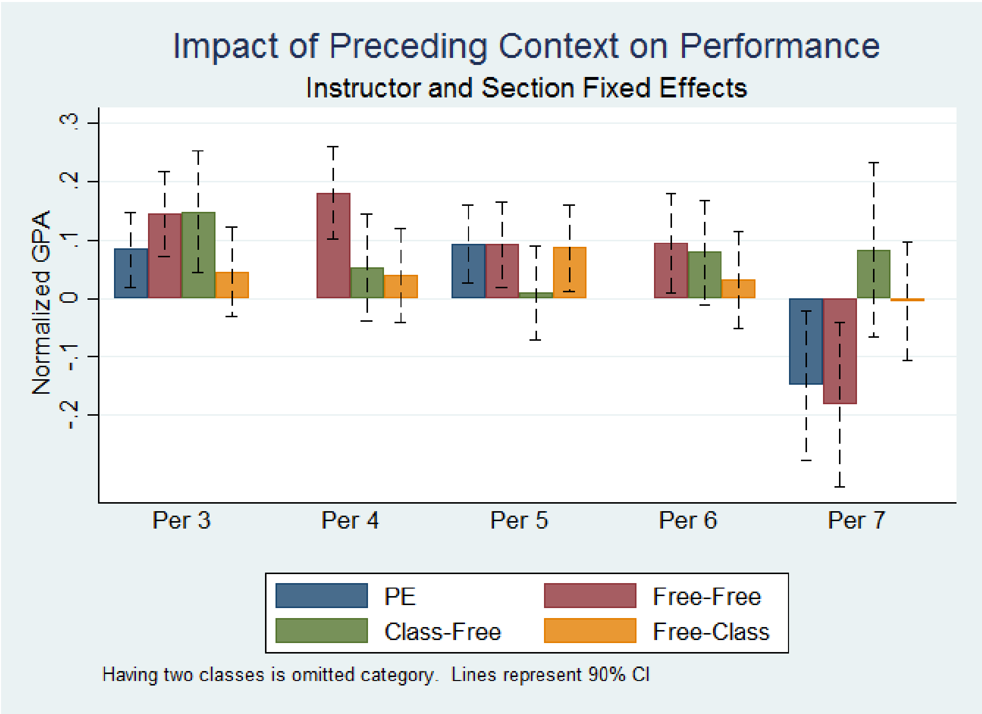


Figure 8: Effects of Preceding Two Classes



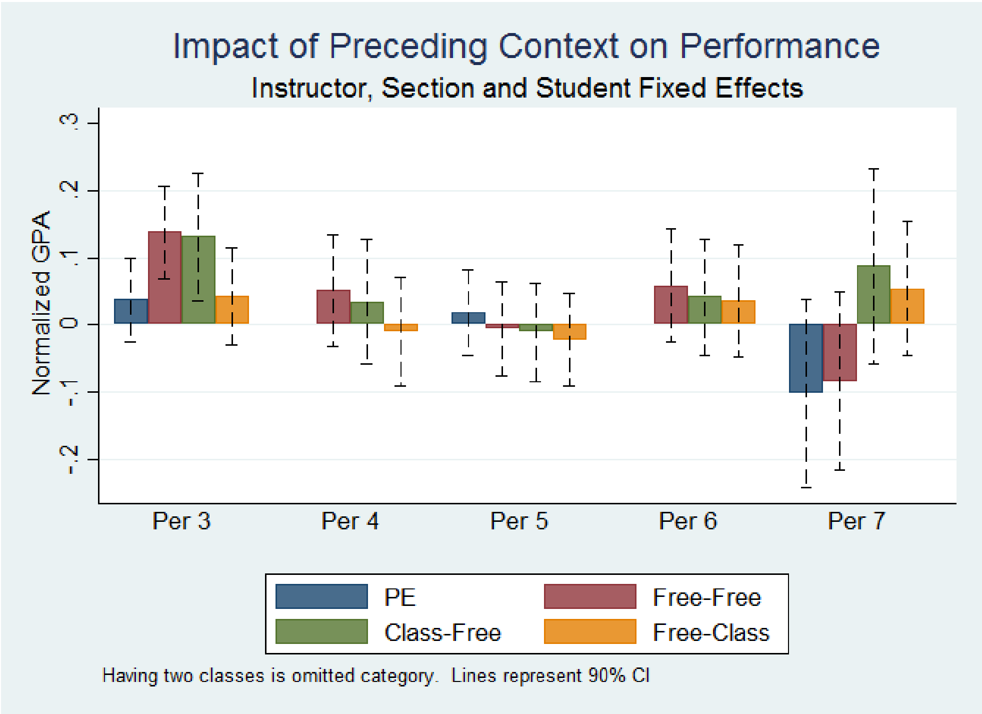
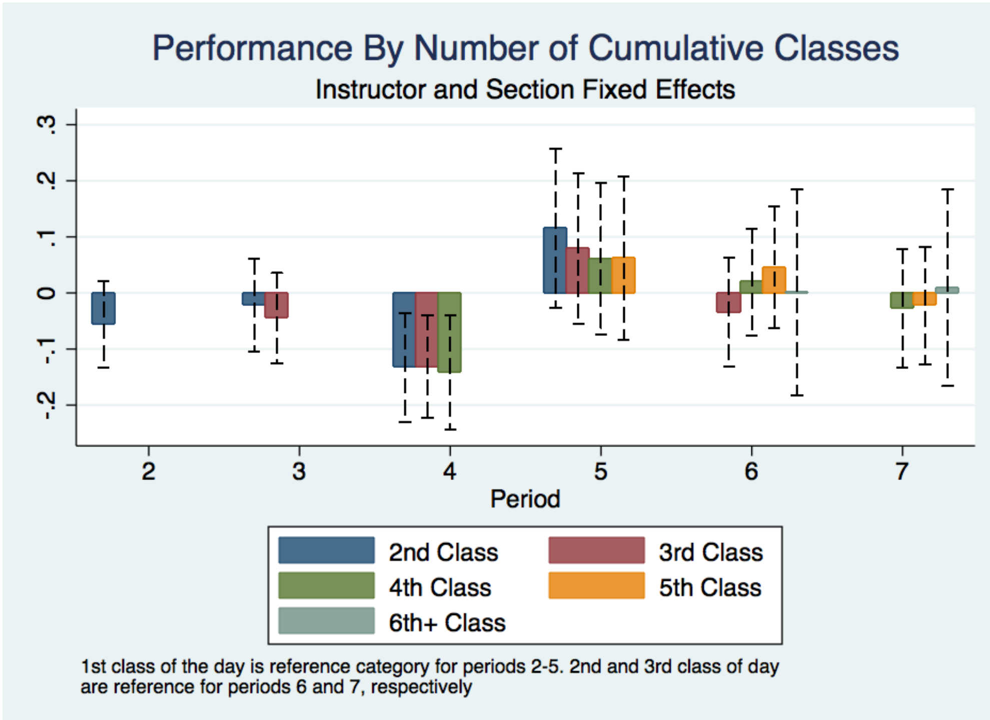


Figure 9: Effects of Cumulative vs Consecutive Classes



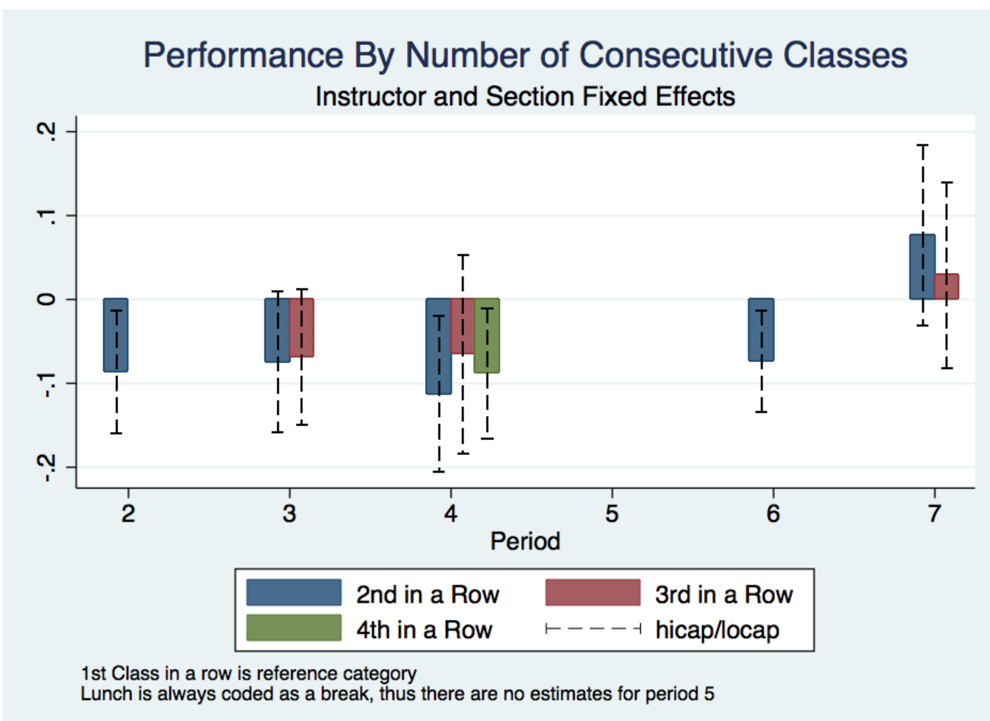


Figure 10: Plot of predicted student GPA vs predicted schedule impact

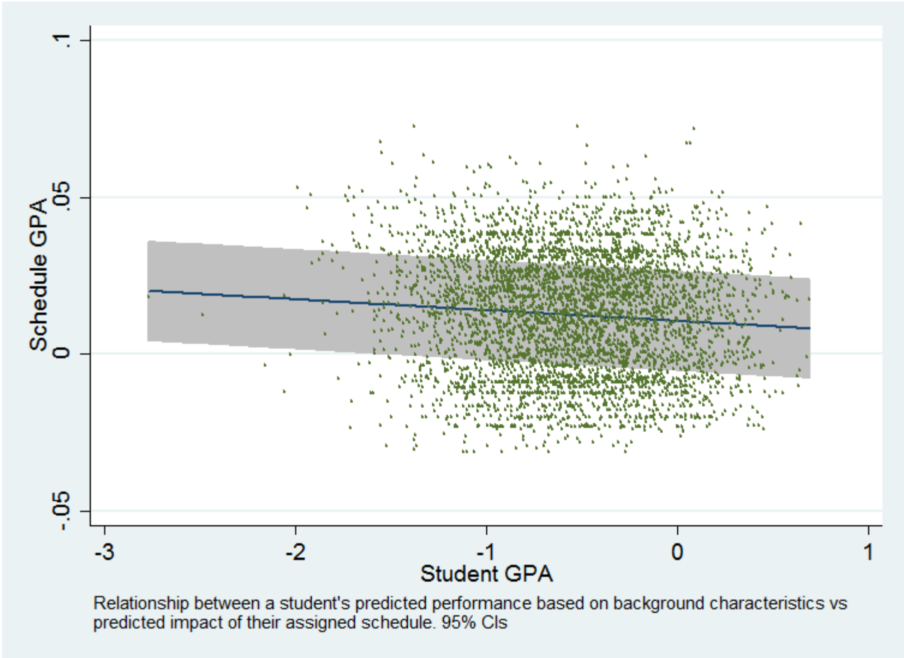


Figure 11: and GPA distribution before and after homogenous simulation

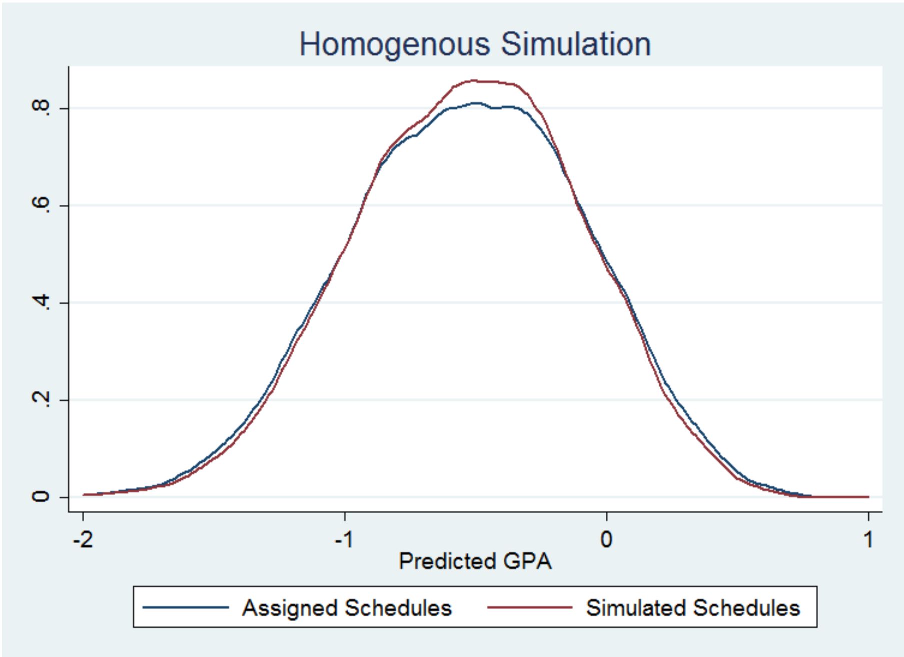


Table 1: Student Summary Statistics by Subgroup

	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 2: 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 3: Number of Course Sections By Period

	Per. 1 mean	Per. 2 mean	Per. 3 mean	Per. 4 mean	Per. 5 mean	Per. 6 mean	Per. 7 mean
Chemistry	82	0	79	0	67	0	0
Comp. Sci.	23	29	26	26	19	18	11
English	22	28	23	31	26	22	16
Engineering	18	19	19	19	16	16	13
History	16	23	29	29	29	28	16
Language	35	48	47	47	23	37	4
Math	32	50	56	51	42	44	26
P.E.	133	0	127	0	101	0	0

Average Section size is 19.3 and course size is 296

Note: The table above shows the number of sections of each subject by class period.

Table 4: The Time of Day and Student Fatigue Effects
Outcome: Normalized Grade

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Period 2	0.0564**	0.0812***	0.0729***	0.0906***	0.0984***	0.0955***	0.126***
	(0.0263)	(0.0254)	(0.0272)	(0.0284)	(0.0294)	(0.0285)	(0.0270)
Period 3	0.0694***	0.0972***	0.121***	0.147***	0.147***	0.144***	0.178***
	(0.0246)	(0.0233)	(0.0293)	(0.0319)	(0.0337)	(0.0318)	(0.0298)
Period 4	0.0907***	0.121***	0.141***	0.151***	0.164***	0.154***	0.197***
	(0.0313)	(0.0292)	(0.0360)	(0.0393)	(0.0412)	(0.0389)	(0.0363)
Period 5	0.0631*	0.0848**	0.0753	0.100*	0.109**	0.104**	0.149***
	(0.0373)	(0.0352)	(0.0459)	(0.0514)	(0.0529)	(0.0503)	(0.0479)
Period 6	0.169***	0.171***	0.197***	0.228***	0.238***	0.235***	0.258***
	(0.0439)	(0.0420)	(0.0529)	(0.0562)	(0.0572)	(0.0550)	(0.0528)
Period 7	0.0967*	0.120**	0.154**	0.180***	0.183***	0.182***	0.222***
	(0.0499)	(0.0478)	(0.0610)	(0.0623)	(0.0629)	(0.0617)	(0.0599)
Credits/Day			-0.00557*	-0.00543*	-0.00518	-0.00525	-0.00970***
			(0.00323)	(0.00324)	(0.00328)	(0.00325)	(0.00281)
Consecutive Classes			-0.0344***	-0.0655**			
			(0.0113)	(0.0280)			
Cumulative Classes			-0.00503	-0.0326		-0.0348	-0.0532**
			(0.00989)	(0.0236)		(0.0225)	(0.0213)
Consecutive Squared				0.0154			
				(0.0104)			
Cumulative Squared				0.00607		0.00634	0.0104**
				(0.00463)		(0.00453)	(0.00426)
Back-to-Back Classes						-0.0575***	-0.0388**
						(0.0200)	(0.0189)
Teacher FEs	Y	Y	Y	Y		Y	Y
Individual FEs	N	Y	N	N	N	N	Y
N	22600	22600	22600	22600	22600	22600	22600
R2	0.253	0.680	0.254	0.254	0.254	0.254	0.681

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: The Instructor Schedule Effect
Outcome: Normalized Grade

	(1)	(2)	(3)	(4)	(5)	(6)
Cumulative Taught	0.0327	0.0480*	0.00488		-0.0155	0.0418
	(0.0246)	(0.0275)	(0.0359)		(0.0365)	(0.0342)
Consecutive Taught	0.000573	0.00315	0.000539			
	(0.0167)	(0.0182)	(0.0224)			
Teach Cumul Squared	-0.00201	-0.00324	0.00377		0.00781	0.0000280
	(0.00786)	(0.00806)	(0.00891)		(0.00956)	(0.00890)
Tchr's 2nd Class of Day				-0.0124		
				(0.0329)		
Tchr's 3rd Class of Day				0.0185		
				(0.0543)		
Tchr's 4th+ Class of Day				0.0360		
				(0.0761)		
Tchr's 2nd in a Row				0.0292		
				(0.0287)		
Tchr's 3rd+ in a Row				-0.0110		
				(0.0485)		
Taught Back-to-Back					0.0233	-0.0181
					(0.0275)	(0.0262)
Fatigue Controls	N	Y	Y	Y	Y	Y
Periods	N	N	Y	Y	Y	Y
Indv FEs	N	N	N	N	N	Y
N	22445	22445	22445	22445	22445	22445
R2	0.253	0.254	0.255	0.255	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. The student fatigue control variables include credits per day, number of cumulative classes, number of cumulative classes squared, and whether the class was immediately following another one. Standard errors are clustered by student.

Table 6: Combined Effects and Subgroup Analysis
Outcome: Normalized Grade

	(1) All	(2) All	(3) High	(4) Middle	(5) Low	(6) STEM	(7) Non STEM
Period 2	0.0901***	0.113***	0.0540	0.0666	0.182***	0.0875*	0.0537
	(0.0324)	(0.0305)	(0.0521)	(0.0631)	(0.0577)	(0.0461)	(0.0513)
Period 3	0.139***	0.141***	0.0624	0.0877	0.244***	0.158***	0.0646
	(0.0399)	(0.0377)	(0.0647)	(0.0766)	(0.0701)	(0.0496)	(0.0719)
Period 4	0.141***	0.141***	0.116	0.0274	0.257***	0.141**	0.0981
	(0.0526)	(0.0497)	(0.0878)	(0.0983)	(0.0938)	(0.0707)	(0.0880)
Period 5	0.102	0.0634	0.0717	0.00424	0.178	0.125	0.0626
	(0.0703)	(0.0655)	(0.116)	(0.133)	(0.124)	(0.0869)	(0.125)
Period 6	0.216***	0.153**	0.156	0.113	0.379***	0.244**	0.187
	(0.0823)	(0.0770)	(0.134)	(0.156)	(0.145)	(0.107)	(0.138)
Period 7	0.143	0.0861	0.000411	0.00520	0.380**	0.111	0.180
	(0.101)	(0.0937)	(0.160)	(0.192)	(0.179)	(0.131)	(0.164)
Credits/Day	-0.00517	-0.00952***	-0.00794	-0.00559	0.00326	-0.00360	-0.00741
	(0.00325)	(0.00281)	(0.00511)	(0.00555)	(0.00634)	(0.00379)	(0.00502)
Back-to-Back Classes	-0.0565***	-0.0379**	0.00802	-0.0979***	-0.0681*	-0.0546**	-0.0535*
	(0.0201)	(0.0191)	(0.0333)	(0.0366)	(0.0376)	(0.0275)	(0.0284)
Cumulative Classes	-0.0362	-0.0525**	-0.0536	0.00233	-0.0494	-0.0553*	0.0108
	(0.0226)	(0.0214)	(0.0388)	(0.0402)	(0.0403)	(0.0285)	(0.0335)
Cumulative Squared	0.00654	0.00992**	0.00875	-0.000803	0.0135	0.00970*	-0.00247
	(0.00453)	(0.00426)	(0.00759)	(0.00803)	(0.00825)	(0.00586)	(0.00673)
Teacher Controls	Y	Y	Y	Y	Y	Y	Y
Individual FEs	N	Y	N	N	N	N	N
N	22445	22445	7291	7550	7604	13080	9365
R2	0.255	0.682	0.324	0.242	0.237	0.282	0.267

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. A full set of teacher schedule variables are included in all regressions, but not shown because they remain statistically insignificant. Columns (3) - (5) limits the sample to specific subgroups. High, Middle, and Low denote students in the top, middle, and bottom tercile of pre-college academic achievement, respectively. Standard errors are clustered by student.

Table 7: Combined Effects and Subgroup Analysis
Outcome: A in Class

	(1) All	(2) All	(3) High	(4) Middle	(5) Low	(6) STEM	(7) Non STEM
Period 2	0.0240*	0.0244	0.0311	0.0221	0.0177	0.0283	0.00419
	(0.0140)	(0.0152)	(0.0294)	(0.0252)	(0.0202)	(0.0200)	(0.0215)
Period 3	0.0421**	0.0306	0.0804**	0.0287	0.0147	0.0495**	0.00374
	(0.0171)	(0.0187)	(0.0354)	(0.0318)	(0.0241)	(0.0216)	(0.0302)
Period 4	0.0569**	0.0486**	0.122***	0.0238	0.0289	0.0450	0.0280
	(0.0224)	(0.0244)	(0.0474)	(0.0415)	(0.0315)	(0.0306)	(0.0379)
Period 5	0.0456	0.0343	0.129**	-0.0144	0.0204	0.0685*	-0.0213
	(0.0306)	(0.0331)	(0.0629)	(0.0591)	(0.0424)	(0.0381)	(0.0534)
Period 6	0.0794**	0.0515	0.207***	0.0113	0.0392	0.0980**	0.0198
	(0.0355)	(0.0388)	(0.0734)	(0.0677)	(0.0488)	(0.0465)	(0.0589)
Period 7	0.0480	0.0294	0.165*	-0.0365	0.0333	0.0355	0.00811
	(0.0426)	(0.0470)	(0.0884)	(0.0810)	(0.0590)	(0.0563)	(0.0702)
Credits/Day	-0.00222*	-0.00188	-0.00411	-0.00318	0.00170	-0.00161	-0.00323*
	(0.00128)	(0.00143)	(0.00282)	(0.00228)	(0.00176)	(0.00166)	(0.00193)
Back-to-Back Classes	-0.00521	-0.00351	0.0254	-0.0170	-0.0143	-0.00791	-0.000488
	(0.00854)	(0.00934)	(0.0178)	(0.0159)	(0.0120)	(0.0120)	(0.0120)
Cumulative Classes	-0.0174*	-0.0139	-0.0417**	-0.00299	-0.00979	-0.0239*	-0.00314
	(0.00948)	(0.0106)	(0.0206)	(0.0167)	(0.0132)	(0.0126)	(0.0134)
Cumulative Squared	0.00387**	0.00325	0.00755*	0.00140	0.00371	0.00517*	0.00106
	(0.00193)	(0.00215)	(0.00412)	(0.00330)	(0.00283)	(0.00269)	(0.00269)
Teacher Controls	Y	Y	Y	Y	Y	Y	Y
Individual FEs	N	Y	N	N	N	N	N
N	22445	22445	7291	7550	7604	13080	9365
R2	0.220	0.558	0.292	0.212	0.192	0.242	0.198

Standard errors in parentheses

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

Note: The outcome is whether the student earned an A or A- in the class. 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. A full set of teacher schedule variables are included in all regressions, but not shown because they remain statistically insignificant. Columns (3) - (5) limits the sample to specific subgroups. High, Middle, and Low denote students in the top, middle, and bottom tercile of pre-college academic achievement, respectively. Standard errors are clustered by student.

Table 8: Combined Effects and Subgroup Analysis
Outcome: Failed Class

	(1) All	(2) All	(3) High	(4) Middle	(5) Low	(6) STEM	(7) Non STEM
Period 2	-0.0353***	-0.0387***	-0.0195*	-0.0476***	-0.0513**	-0.0367**	-0.0201
	(0.00979)	(0.0108)	(0.0106)	(0.0183)	(0.0208)	(0.0147)	(0.0127)
Period 3	-0.0492***	-0.0527***	-0.000910	-0.0495**	-0.0996***	-0.0544***	-0.0192
	(0.0130)	(0.0148)	(0.0155)	(0.0243)	(0.0264)	(0.0180)	(0.0170)
Period 4	-0.0414**	-0.0414**	0.00372	-0.0415	-0.0955***	-0.0444*	-0.0176
	(0.0165)	(0.0191)	(0.0196)	(0.0302)	(0.0340)	(0.0246)	(0.0208)
Period 5	-0.0326	-0.0231	0.00701	-0.0469	-0.0601	-0.0170	-0.0337
	(0.0231)	(0.0262)	(0.0270)	(0.0421)	(0.0472)	(0.0315)	(0.0307)
Period 6	-0.0434*	-0.0325	0.00460	-0.0581	-0.0848	-0.0271	-0.0505
	(0.0259)	(0.0297)	(0.0313)	(0.0482)	(0.0524)	(0.0370)	(0.0338)
Period 7	-0.0451	-0.0325	0.00775	-0.0705	-0.0740	-0.0140	-0.0569
	(0.0306)	(0.0351)	(0.0363)	(0.0574)	(0.0618)	(0.0443)	(0.0399)
Credits/Day	-0.0000855	0.000935	0.00113	-0.00158	0.000496	-0.000407	0.000454
	(0.000897)	(0.00100)	(0.000947)	(0.00154)	(0.00199)	(0.00121)	(0.00113)
Back-to-Back Classes	0.0128**	0.00844	0.00120	0.0294***	0.00778	0.0253***	-0.00197
	(0.00525)	(0.00619)	(0.00614)	(0.00936)	(0.0114)	(0.00804)	(0.00642)
Cumulative Classes	0.00579	0.0117	-0.00275	-0.0123	0.0306**	0.000778	0.00576
	(0.00597)	(0.00761)	(0.00734)	(0.0110)	(0.0122)	(0.00861)	(0.00724)
Cumulative Squared	-0.00114	-0.00257*	0.000723	0.00222	-0.00708***	-0.000422	-0.000903
	(0.00116)	(0.00144)	(0.00141)	(0.00211)	(0.00242)	(0.00172)	(0.00140)
Teacher Controls	Y	Y	Y	Y	Y	Y	Y
Individual FEs	N	Y	N	N	N	N	N
N	22445	22445	7291	7550	7604	13080	9365
R2	0.148	0.475	0.148	0.187	0.228	0.166	0.107

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < .01$

Note: The outcome is whether the student earned a D or F in the class. 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. A full set of teacher schedule variables are included in all regressions, but not shown because they remain statistically insignificant. Columns (3) - (5) limits the sample to specific subgroups. High, Middle, and Low denote students in the top, middle, and bottom tercile of pre-college academic achievement, respectively. Standard errors are clustered by student.

Table 9: Combined Effects and Subgroup Analysis
Outcome: Raw Score

	(1) All	(2) All	(3) High	(4) Middle	(5) Low	(6) STEM	(7) Non STEM
Period 2	0.765***	0.839***	0.473	0.536	1.595***	0.653	0.472
	(0.278)	(0.261)	(0.424)	(0.547)	(0.500)	(0.408)	(0.392)
Period 3	1.295***	1.131***	0.472	0.760	2.458***	1.458***	0.488
	(0.360)	(0.339)	(0.564)	(0.696)	(0.637)	(0.474)	(0.554)
Period 4	1.229***	1.067**	0.862	0.241	2.460***	1.141*	0.780
	(0.466)	(0.442)	(0.754)	(0.875)	(0.836)	(0.658)	(0.678)
Period 5	1.024	0.479	0.714	0.237	1.863	1.137	0.680
	(0.636)	(0.594)	(1.023)	(1.204)	(1.133)	(0.829)	(0.967)
Period 6	1.824**	1.001	1.208	0.928	3.463***	1.921*	1.577
	(0.734)	(0.684)	(1.166)	(1.403)	(1.304)	(1.007)	(1.070)
Period 7	1.422	0.620	0.142	0.235	3.655**	1.100	1.575
	(0.893)	(0.827)	(1.383)	(1.710)	(1.600)	(1.228)	(1.282)
Credits/Day	-0.0497*	-0.0824***	-0.0674*	-0.0498	0.000704	-0.0413	-0.0641*
	(0.0270)	(0.0244)	(0.0408)	(0.0451)	(0.0537)	(0.0333)	(0.0381)
Back-to-Back Classes	-0.417***	-0.269*	0.129	-0.759**	-0.499	-0.436*	-0.360*
	(0.161)	(0.157)	(0.259)	(0.297)	(0.304)	(0.229)	(0.214)
Cumulative Classes	-0.295	-0.349*	-0.417	0.0680	-0.499	-0.417*	0.0639
	(0.183)	(0.183)	(0.306)	(0.332)	(0.326)	(0.241)	(0.254)
Cumulative Squared	0.0556	0.0747**	0.0726	-0.0111	0.133**	0.0766	-0.0135
	(0.0359)	(0.0355)	(0.0585)	(0.0647)	(0.0652)	(0.0481)	(0.0511)
Teacher Controls	Y	Y	Y	Y	Y	Y	Y
Individual FEs	N	Y	N	N	N	N	N
N	22445	22445	7291	7550	7604	13080	9365
R2	0.486	0.774	0.525	0.478	0.491	0.547	0.304

Standard errors in parentheses

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

Note: The outcome is the percent score the student earned in the class. 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. A full set of teacher schedule variables are included in all regressions, but not shown because they remain statistically insignificant. Columns (3) - (5) limits the sample to specific subgroups. High, Middle, and Low denote students in the top, middle, and bottom tercile of pre-college academic achievement, respectively. Standard errors are clustered by student.

Table 10: Subgroup Robustness Specifications
Outcome: Normalized Grade

	(1) No Lang	(2) No Lang	(3) Athletes	(4) Athletes	(5) 7:00	(6) 7:30	(7) 7:50
Period 2	0.0996***	0.119***	0.0633**	0.0820***	0.115	0.0393	0.117**
	(0.0345)	(0.0328)	(0.0269)	(0.0257)	(0.0736)	(0.0579)	(0.0505)
Period 3	0.155***	0.154***	0.106***	0.105***	0.279***	0.122**	0.0982
	(0.0408)	(0.0390)	(0.0344)	(0.0330)	(0.103)	(0.0587)	(0.0715)
Period 4	0.155***	0.154***	0.127***	0.119***	0.362**	0.113	0.0738
	(0.0544)	(0.0515)	(0.0456)	(0.0437)	(0.146)	(0.0750)	(0.0953)
Period 5	0.131*	0.106	0.0725	0.0347	0.334	0.123	-0.00670
	(0.0717)	(0.0681)	(0.0622)	(0.0588)	(0.223)	(0.0955)	(0.127)
Period 6	0.269***	0.215***	0.178**	0.119*	0.527**	0.218**	0.122
	(0.0848)	(0.0811)	(0.0718)	(0.0688)	(0.259)	(0.111)	(0.151)
Period 7	0.196*	0.146	0.107	0.0664	0.425	0.161	0.0302
	(0.104)	(0.0983)	(0.0878)	(0.0837)	(0.308)	(0.136)	(0.183)
Credits/Day	-0.00567*	-0.0104***	-0.00389	-0.00986***	-0.00992	-0.00604	-0.000467
	(0.00338)	(0.00294)	(0.00294)	(0.00255)	(0.00746)	(0.00554)	(0.00475)
Back-to-Back Classes	-0.0553***	-0.0378*	-0.0630***	-0.0403**	-0.0749	-0.00664	-0.0952***
	(0.0208)	(0.0200)	(0.0178)	(0.0170)	(0.0455)	(0.0318)	(0.0312)
Cumulative Classes	-0.0427*	-0.0594***	-0.0349*	-0.0454**	-0.0796*	-0.0603*	0.00793
	(0.0230)	(0.0220)	(0.0208)	(0.0198)	(0.0478)	(0.0355)	(0.0368)
Cumulative Squared	0.00797*	0.0113**	0.00608	0.00881**	0.0145	0.00871	-0.000174
	(0.00468)	(0.00443)	(0.00426)	(0.00401)	(0.0100)	(0.00721)	(0.00720)
Cumulative Taught	-0.0169	0.0426	-0.00168	0.0380	-0.0720	0.00225	-0.00651
	(0.0373)	(0.0355)	(0.0331)	(0.0318)	(0.117)	(0.0511)	(0.0652)
Taught Back-to-Back	0.0205	-0.0112	0.0272	-0.0171	0.0593	-0.00678	0.0160
	(0.0283)	(0.0277)	(0.0243)	(0.0241)	(0.0804)	(0.0397)	(0.0481)
Teach Cumul Squared	0.00585	-0.00326	0.00367	-0.00214	0.0109	0.00256	0.00889
	(0.00977)	(0.00929)	(0.00891)	(0.00839)	(0.0284)	(0.0153)	(0.0148)
Teacher Controls	Y	Y	Y	Y	Y	Y	Y
Individual FEs	N	Y	N	Y	N	N	N
N	20655	20655	28825	28825	4503	8963	8979
R2	0.256	0.697	0.257	0.678	0.273	0.253	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 11: Fixed Effect and Peer Effect Robustness Specifications
Outcome: Normalized Grade

	(1)	(2)	(3)	(4)
Period 2	0.139***	0.147***	0.138***	0.0901***
	(0.0253)	(0.0248)	(0.0258)	(0.0324)
Period 3	0.154***	0.171***	0.158***	0.139***
	(0.0284)	(0.0281)	(0.0286)	(0.0399)
Period 4	0.130***	0.135***	0.140***	0.141***
	(0.0332)	(0.0330)	(0.0343)	(0.0526)
Period 5	0.124***	0.192***	0.110**	0.102
	(0.0404)	(0.0431)	(0.0450)	(0.0703)
Period 6	0.195***	0.272***	0.186***	0.216***
	(0.0414)	(0.0451)	(0.0475)	(0.0823)
Period 7	0.110**	0.200***	0.142***	0.143
	(0.0465)	(0.0499)	(0.0523)	(0.101)
Credits/Day	-0.00129	-0.00325	-0.00431	-0.00517
	(0.00325)	(0.00322)	(0.00325)	(0.00325)
Back-to-Back Classes	-0.0402**	-0.0421**	-0.0530***	-0.0565***
	(0.0196)	(0.0195)	(0.0197)	(0.0201)
Cumulative Classes	-0.0388*	-0.0433*	-0.0360	-0.0362
	(0.0223)	(0.0222)	(0.0223)	(0.0226)
Cumulative Squared	0.00429	0.00552	0.00608	0.00654
	(0.00445)	(0.00442)	(0.00444)	(0.00453)
Cumulative Taught	-0.00523	0.00206	-0.00151	-0.0155
	(0.0234)	(0.0234)	(0.0245)	(0.0365)
Taught Back-to-Back	0.00505	-0.0171	0.00150	0.0233
	(0.0201)	(0.0199)	(0.0211)	(0.0275)
Teach Cumul Squared	0.0227***	0.0150*	0.00878	0.00781
	(0.00800)	(0.00802)	(0.00836)	(0.00956)
Peer Controls	N	Y	Y	Y
Course x Sem x MT FEs	N	N	Y	Y
Teacher FEs	N	N	N	Y
Individual FEs	N	N	N	N
N	22445	22445	22445	22445
R2	0.156	0.172	0.196	0.255

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 and controls for student characteristics. Each column shows estimates when including additional controls and fixed effects. Standard errors are clustered by student.

Table 12: Aggregated Coefficients

	Cumulative Classes						
	0	1	2	3	4	5	6
Period							
1	0						
2	0.0901						
$2^{\dagger}$		0.0039					
3	0.1390	0.1028					
$3^{\dagger}$		0.0463	0.0101				
4	0.1410	0.1049	0.0687				
$4^{\dagger}$		0.0484	0.0122	-0.0240			
5	0.1020	0.0659	0.0297	-0.0065			
$5^{\dagger}$		0.0094	-0.0268	-0.0630	-0.0992		
6	0.2160	0.1798	0.1436	0.1074	0.0712		
$6^{\dagger}$		0.1233	0.0871	0.0509	0.0147	-0.0215	
7	0.1430	0.1070	0.0708	0.0346	-0.0016	-0.0378	
$7^{\dagger}$		0.0505	0.0143	-0.0219	-0.0581	-0.0943	-0.1305

Estimates are taken from column 1 of Table 6. Period 1 is set to 0. Gives expected value based on time of day, number of cumulative classes and whether or not class is a back-to-back one. † represents a class being back-to-back.

Table 13: Simulation Results: Homogenous Schedules

	All			Bottom Quartile		
	Mean	SD	N	Mean	SD	N
$G\hat{P}A_{if}$	0.0	0.452	4,536	-0.586	0.228	1125
$Sim\hat{G}PA_{if}^{homo}$	0.0	0.434	4,536	-0.564	0.221	1125
Difference	0.0			0.022		

Table 14: Simulation Results: Heterogenous Schedules

	All			Bottom		Middle		Top	
	Mean	SD	N	Mean	SD	Mean	SD	Mean	SD
$G\hat{P}A_{if}$	0.0	0.395	4,536	-0.391	0.248	-0.034	0.122	0.447	0.207
$Sim\hat{G}PA_{if}^{homo}$	0.012	0.371	4,536	-0.353	0.225	-0.024	0.112	0.434	0.198
$Difference$	0.012			0.038		0.010		-0.013	