Global Perspectives on Accounting Education
Volume 14, 2017, 44-66

THE IMPACT OF CLASS SCHEDULING ON ACADEMIC PERFORMANCE IN QUANTITATIVE AND QUALITATIVE BUSINESS DISCIPLINES

Adriana S. Cordis
Department of Accounting, Finance, and Economics
Winthrop University
Rock Hill, SC
USA

Barbara G. Pierce
School of Accounting
University of South Carolina
Columbia, SC
USA

ABSTRACT

Recent research indicates that class scheduling influences the academic performance of college students. Specifically, it finds that grades for early-morning classes are systematically lower than for classes held later in the day, all else being equal. Building on this research, we present evidence that class scheduling has a larger influence on student performance in quantitative business disciplines, such as accounting and finance, than in qualitative business disciplines, such as management. We also find that the scheduling effect is more pronounced for single-section classes than for multiple-section classes, which is consistent with a self-selection effect that acts to mitigate the adverse impact of early morning classes on grades.

Data availability: Data are available upon request

INTRODUCTION

Does the time of day at which a class meets have an impact on the academic performance of students? Several recent studies suggest that this is indeed the case (Dills and Hernandez-Julian, 2008; Carrell et al., 2011; Edwards, 2012; Pope, 2014). For example, Dills and Hernandez-Julian (2008) examine the relation between academic achievement of college students, as measured by letter grades earned, and the time of day at which classes are held. They report that after controlling for factors such as class size, semester taught, meetings per week, and fixed student and class characteristics, grades are lower for early-morning classes, a finding that they interpret as consistent with medical evidence regarding the effects of sleep on learning.[1]

This paper examines the impact of class scheduling on the academic performance of undergraduate business students with an eye towards assessing the implications for accounting education. Our analysis is motivated by the view that sleep disruption is likely to have a larger impact on the ability of students to learn problem-solving skills, like those required to conduct a cost-volume-profit analysis, than on their ability to learn descriptive material. To support this view, we draw on the findings of studies in the neurophysiology and behavioral neuroscience literature that suggest that sleep loss has a stronger detrimental impact on acquiring and retaining procedural knowledge (the ability to perform a given task or solve a particular problem) than on acquiring and retaining declarative knowledge (the ability to recall specific facts).[2]

We begin by noting that some business disciplines, such as accounting and finance, place considerably more emphasis on solving quantitative problems than others, such as management. In general, we expect the skills required to solve such problems to reflect procedural knowledge rather than declarative knowledge. We therefore posit that the class scheduling effect documented in prior research (lower grades in early-morning classes) is more pronounced for classes in quantitative business disciplines than for those in qualitative business disciplines. In other words, our core hypothesis is that the magnitude of the scheduling effect for early-morning quantitative classes is larger than that for early-morning qualitative classes.

The set of business disciplines considered includes accounting, finance, marketing, management, economics, quantitative methods, and general business administration. We classify accounting, finance, economics and quantitative methods as quantitative disciplines, and marketing, management, and general business administration as qualitative disciplines. Admittedly, our discipline-based classification scheme may be overly broad. A class in accounting ethics, for example, is unlikely to be more quantitative than a class in marketing research. Nonetheless, similarly broad classifications have been used in prior studies to demonstrate that distinguishing between qualitative and quantitative classes matters with respect to various learning outcomes.[3]

---

[1] Interestingly, the medical profession has become actively involved in the class scheduling debate. In a recent policy statement, the American Academy of Pediatrics notes that "delaying school start times is an effective counter measure to chronic sleep loss and has a wide range of potential benefits to students with regard to physical and mental health, safety, and academic achievement." The statement recommends against scheduling middle and high school classes any earlier than 8:30 a.m. (Au et al., 2014).

[2] Curcio et al. (2006) provides a good overview of this research.

[3] For instance, Kidwell and Kidwell (2008) classify accounting, economics, finance, statistics, and related fields as quantitative disciplines, and management, marketing, ethics, and related fields as qualitative disciplines. For other examples, see Loo (2002), Schlee (2005), Burke et al. (2009), and Schlee and Harich (2014). Our classifications are also supported by an analysis of differences in course prerequisites (mathematics and statistics classes) across disciplines, and an analysis of stated learning objectives across disciplines.

The data for the analysis are from the administrative records of Winthrop University, a public, coeducational, liberal arts university in South Carolina. They cover all undergraduate classes offered by the College of Business during the fall and spring semesters of calendar years 2012 and 2013. Our strategy for assessing the relation between class scheduling and student performance relies on multiple regression techniques. For example, to assess whether classes that meet from 8:00 a.m. to 9:15 a.m. (the earliest available time slot) produce lower grades than those that meet later in the day, we fit regression specifications in which one of the regressors is a dummy variable that equals one for 8:00 a.m. classes and zero otherwise. The estimated slope on the 8:00 a.m. dummy is the estimated marginal effect of moving the class from a later time slot to 8:00 a.m.

Our baseline regression specification, which controls for unobserved student heterogeneity, but does not distinguish between qualitative and quantitative classes, produces little support for the proposition that early morning classes are detrimental to academic performance. Although the estimated coefficient on the 8:00 a.m. dummy is negative, the associated t-statistic indicates that it is statistically indistinguishable from zero. In other words, we find no statistically significant evidence of the anticipated scheduling effect.

It is important to note, however, that the lack of statistically significant evidence is not necessarily at odds with our core hypothesis. Suppose that, as we hypothesize, the effect of early-morning classes on student performance is more pronounced for quantitative disciplines than for qualitative disciplines. Under these circumstances, restricting the coefficient on the 8:00 a.m. dummy to be the same for all classes could tend to mask the existence of the class scheduling effect. In other words, the presence of observations for qualitative classes could make it difficult to detect the class scheduling effect if it is relatively weak for these classes.

To assess whether the scheduling effect differs across qualitative and quantitative disciplines, we replace the 8:00 a.m. dummy with two separate variables: a dummy for qualitative 8:00 a.m. classes and a dummy for quantitative 8:00 a.m. classes. The resulting regression estimates are consistent with the hypothesis that class scheduling has a larger impact on grades for quantitative disciplines than for qualitative disciplines. The estimated slope on the quantitative dummy is both negative and statistically significant, while the estimated slope on the qualitative dummy is small in magnitude and statistically indistinguishable from zero.

The contrast between the estimated slopes on the qualitative and quantitative dummies suggests that the benefits of altering class scheduling policies differ across qualitative and quantitative disciplines. In quantitative disciplines, such as accounting and finance, our analysis suggests that moving a class from 8:00 a.m. to later in the day increases grades by one to two tenths of a grade point, depending on the specification. If this increase in grades accurately reflects student learning, then one can make a reasoned argument in favor of reducing the prevalence of early morning quantitative classes. Although any change in scheduling policy would naturally require an evaluation of the costs and benefits, our findings suggest that faculty and administrators should be open to revamping traditional scheduling practices.

We also assess whether the class scheduling effect is influenced by students choosing between two different sections of the same class based on start time. Prior research suggests that if students are allowed to choose between multiple sections of a class with different start times, they make this choice in a way that leads to better grade outcomes on average. To identify this self-selection effect, we fit regression specifications that include a dummy for single-section 8:00 a.m. classes and a dummy for multiple-section 8:00 a.m. classes. As anticipated, the regression estimates are clearly suggestive of a self-selection effect. The estimated slope on the single-section dummy is both negative and statistically significant, while the estimated slope on the multiple-section dummy is small in magnitude and statistically indistinguishable from zero. Further analysis suggests that the self-selection effect is stronger for quantitative classes, although our conclusions in this regard are tentative due to sample size considerations.

Finally, we conduct several additional tests to check the robustness of our findings. It is conceivable, for example, that 9:30 a.m. classes are “early-morning classes” in the sense that grades for these classes are systematically lower than those for classes that meet during later time slots. If this is true, then classifying 9:30 a.m. classes as “later in the day” when estimating the scheduling effect for 8:00 a.m. classes could bias our estimates downwards. The evidence is broadly consistent with our conjecture. Once we exclude 9:30 a.m. classes from the dataset used to fit the regressions, we find that it is no longer necessary to distinguish between quantitative and qualitative classes to obtain statistically-significant estimates of the scheduling effect.

Overall our analysis suggests that department chairs in quantitative disciplines may want to exercise some control over the type of classes that are offered during early morning time slots. In accounting, for example, offering a quantitatively-demanding class such as advanced cost accounting during an afternoon time slot might lead to better learning outcomes than offering it at 8:00 a.m. Because most institutions provide some flexibility in choosing the time slots for different classes within the same discipline, adopting a policy of scheduling the least quantitatively-demanding classes during the early morning might prove beneficial.

BACKGROUND AND HYPOTHESES DEVELOPMENT

Student academic achievement is undoubtedly influenced by a large number of factors. The grade earned by a student in a particular class could be a function of innate ability, class size, instructor effectiveness, the method of instructional delivery, the nature of the material covered, and a host of other things. Many of the factors that influence academic achievement, such as student ability, are outside the control of instructors and administrators once the students have been admitted. But others, such as the method of instructional delivery, are clearly within their purview to change.

In recent years, class scheduling has received increasing attention as a possible factor in academic achievement. Some studies, such as Edwards (2012), investigate the relation between scheduling and student performance using data from public schools. Others, such as Dills and Hernandez-Julian (2008) and Carrell et al. (2011), focus on the impact of scheduling in the higher-education setting. In each case the empirical evidence suggests that the time of day at which classes are held has an impact on student performance, all else being equal. For instance, Dills and Hernandez-Julian (2008) find that controlling for factors such as class size, semester taught, meetings per week, and fixed student and class characteristics, student grades are lower for early-morning classes than for classes held later in the day.

Researchers have identified several phenomena that may contribute to poor performance in early-morning classes. First, the intuitive proposition that sleepy students perform worse than well-rested students is supported by a number of studies (see, e.g., Austin et al., 1988; Medeiros et al., 2001). Second, students are often reluctant to adjust their sleep habits in order to accommodate class schedules. If classes are scheduled early in the day, then many of the students enrolled in these classes will end up sleeping less than would otherwise be the case (Shinkoda et al., 2000; Wolfson and Carskadon, 2003). Third, research on circadian rhythms suggests that alertness fluctuates over the course of the day (Carskadon et al., 1998, 1999).

Although there is certainly some variation in circadian rhythms across individuals, the evidence suggests that the lowest level of alertness in teens and young adults typically occurs between 3:00 a.m. and 7:00 a.m. In addition, a lack of sufficient sleep can both intensify the drop in alertness and extend the duration of the low-alertness period by as much as several hours. Thus early-morning classes have the most potential to be detrimental to student academic performance due to normal circadian rhythms in conjunction with the exacerbating influence of sleep loss from having to wake up early. We therefore consider the following hypothesis.

Hypothesis 1: *Business students achieve lower grades in early morning classes than in classes held later in the day, all else being equal.*

We anticipate in view of the existing evidence that the data will produce support for Hypothesis 1. But the overall effect of class scheduling is not the principal focus of our analysis. Our interest centers on a heretofore unexplored aspect of the relation between class scheduling and academic achievement. In particular, we investigate whether this relation differs by the type of discipline: qualitative versus quantitative. We argue that it may be important to distinguish between qualitative and quantitative disciplines for a couple of reasons.

One reason to distinguish between qualitative and quantitative disciplines is based on the findings from the business education literature showing that this distinction matters in a variety of contexts. Burke et al. (2009) investigate the effectiveness of PowerPoint-based lectures across business disciplines, and conclude that such lectures are less effective in quantitative disciplines than in qualitative disciplines. Schlee and Harich (2014) measure the ability of business students to think creatively along six different dimensions, and conclude that "students in the quantitative business disciplines of accounting, finance, economics and information systems outperformed other business majors in some categories of creative thinking." Loo (2002) studies the learning styles of business students, and concludes that there are significant differences in styles between quantitative and qualitative disciplines. Although these studies do not address class scheduling per se, they do serve to illustrate the importance of distinguishing between qualitative and quantitative disciplines when investigating other aspects of business education.

Another reason to distinguish between qualitative and quantitative disciplines is that studies in the neurophysiology and behavioral neuroscience literature suggest that the effect of sleep loss is likely to differ between qualitative and quantitative disciplines, because sleep loss has a stronger detrimental impact on acquiring procedural knowledge than on acquiring declarative knowledge.[4] For example, the Tower of Hanoi puzzle, which stresses basic logical reasoning, has been used to investigate acquisition of procedural knowledge. It consists of a board with three vertical rods and a set of discs of varying diameter that can be stacked on the rods. Initially, the discs are stacked on a single rod in order of size (largest disc on the bottom). The task is to transfer the discs to a different rod by moving one disc at a time, subject to certain rules: only the top disc of a given stack is eligible to be moved, it must be placed on the top of another stack when it is moved, and it can only be placed on top of a larger-diameter disc.

[4] Procedural memory is a type of nondeclarative memory. Walker (2008) notes that declarative memory refers to "memories of fact-based information (i.e., knowing what)," whereas nondeclarative memory "includes procedural memory (i.e., knowing how), such as the learning of actions, habits, and skills." Cohen and Squire (1980) provide an interesting example of the difference between procedural and declarative learning by studying the ability of amnesic patients to learn to read words reflected in a mirror. The patients learned this procedural task "at a rate equivalent to that of matched control subjects," and did so "despite amnesia for the words that had been read." Hence, Cohen and Squire (1980) conclude that operations governed by rules or procedures "have information-processing and memory characteristics different from those operations that depend on specific, declarative, data-based material."

A number of studies, such as Smith (1993) and Conway and Smith (1994), investigate the role of sleep in learning to solve the Tower of Hanoi puzzle and in other procedural learning tasks. The evidence from these studies indicates that the loss of episodes of rapid eye movement (REM) sleep impairs performance on procedural tasks, but not on declarative tasks. Based on these findings, Smith (1995, 2001) argues that REM sleep is tied to the processing of procedural memories, and that REM sleep is not involved in forming declarative memories. In a related study, Smith and Smith (2003) report that loss of REM sleep due to alcohol consumption impairs memory for procedural tasks, but not declarative tasks. These findings are consistent with evidence that sleep plays a fundamental role in consolidating procedural memories, i.e., in the process by which a newly-acquired, unstable memory is gradually strengthened and integrated into existing knowledge (Plihal and Born, 1997; Peigneux et al., 2004; Marshall and Born, 2007).

In our view, the evidence in the neuroscience literature provides a sound basis for arguing that the relation between class scheduling and academic performance may differ by the type of discipline. Quantitative disciplines generally place a strong emphasis on acquiring problem solving skills. For example, students in an advanced cost accounting class might be asked to use actual company data to conduct a cost-volume-profit analysis, and identify the profit-maximizing level of production under a certain set of constraints. Developing the skills necessary to conduct such an analysis seems much more closely aligned with the definition of procedural knowledge than with that of declarative knowledge. Qualitative disciplines, on the other hand, tend to stress the acquisition of descriptive knowledge. For example, marketing students might be asked to describe some common models of distribution and retailing, or discuss the advantages and disadvantages of different promotional tactics, such as advertising and direct marketing.

We recognize that some might question the strength of the association between procedural knowledge and the sort of quantitative material covered in business classes, and we acknowledge that this question is not unreasonable. It is clear, for instance, that the Tower of Hanoi task discussed earlier is quite simplistic compared with the problems that students are required to solve in quantitative business disciplines. Nonetheless, disciplines that place an emphasis on problem solving seem much more likely to draw on procedural knowledge than those in which problem solving plays a relatively minor role. We therefore believe that there are sound reasons to argue that the extent to which classes focus on quantitative material should be meaningful in the context of our investigation. These considerations lead us to the following hypothesis.

Hypothesis 2 *The effect of early morning classes on the grades of business students is more pronounced for quantitative classes than for qualitative classes, all else being equal.*

Hypothesis 2 is the primary focus of our empirical tests. However, we must also confront the issue of student selection given that in many cases more than one section of a class is offered in a given term. Because different sections are typically held at different times of the day, this presents students with some ability to tailor their class schedule to their individual needs and preferences. For example, sleep habits will undoubtedly vary from one student to another. A student who prefers to go to sleep early and wake up early might be well rested even for early morning classes. If this type of student also has a preference for early morning classes (i.e., preferentially selects into these classes), then the selection effect could work to offset the sleep deprivation effect, thereby making it more difficult to find a relation between class scheduling and student performance in the data. Thus we have the following hypothesis.

Hypothesis 3 *The effect of early morning classes on the grades of business students is more pronounced for single-section classes than for multiple-section classes, all else being equal.*

Note that Hypothesis 3 does not presuppose that every student who enrolls in a given section of a multiple-section class is expressing unambiguous scheduling preferences. There may be many instances in which students are essentially forced to take a particular section due to time conflicts with other required classes. The key point is that, on average, multiple-section classes provide students with more scheduling flexibility than single-sections classes. Whether this flexibility is sufficient to be clearly reflected in grades is an empirical question.

DATA AND RESEARCH DESIGN

Our strategy for testing the three hypotheses is similar to that employed by Dills and Hernandez-Julian (2008). It entails fitting a number of multiple regression specifications to the data in an attempt to both isolate the class scheduling effect and determine whether it differs across quantitative and qualitative disciplines.

The data employed in the empirical analysis are from Winthrop University, a public, coeducational, liberal arts university in South Carolina. Specifically, we use administrative records for all undergraduate three-credit-hour classes offered by the College of Business during the fall and spring semesters of calendar years 2012 and 2013 to construct the dataset. Classes in the College of Business meet twice a week on a Monday/Wednesday, Tuesday/Thursday, or Wednesday/Friday schedule. They are 75 minutes in length, and begin at 8:00 a.m., 9:30 a.m., 11:00 a.m., 12:30 p.m., 2:00 p.m., 3:30 p.m., and 5:00 p.m. Some evening classes are also offered. These classes typically meet once a week at 6:30 p.m. for 150 minutes.

Our measure of student academic performance is the letter grade earned by the student in the class. Winthrop uses a standard letter-grade scale for undergraduate classes. Thus the possible grades for an undergraduate student in a given class are A, B, C, D, and F.[5] Following prior studies, we assign each letter grade a numerical value to construct the dependent variable used for the econometric analysis. In other words, we recode the letter grades such that an A corresponds to 4 points, a B to 3 points, and so forth.

The dataset also contains a number of other variables for each grade observation, including the discipline and course catalog number for the class, the semester it was offered, the name of the instructor, the time at which the class met, the number of meetings per week, and the class size (number of enrolled students). We also know the gender and age of the students. In most cases there are multiple grade observations for a given student: one for each class that the student completed during the two years covered by our sample. The total number of grade observations contained in the sample is 10,039.

Strategy for Testing Hypothesis 1

We assess the relation between class scheduling and student academic performance using

[5] Instructors at Winthrop have the option to use plus and minus modifiers for letter grades. However, relatively few do so in our dataset. We follow the Dills and Hernandez-Julian (2008) approach by rounding off the plus or minus grades to the closest unmodified letter grade (e.g., we code an A− as an A). This has the advantage of maintaining consistency across instructors.

multiple regression techniques. To illustrate, consider the *n*th observation in the sample. This observation consists of the grade earned in some class by one of the students who appears in our dataset along with information about the class and student. Let $early_n$ equal 1 if the class start time for the *n*th observation is 8:00 a.m. and 0 otherwise.

Our general strategy for estimating the effect of an early class start time on grades is to fit a multiple regression model of the form

$$grade_n = \alpha + \beta early_n + \lambda' x_n + \varepsilon_n, \tag{1}$$

where λ is a vector of slope coefficients and x_n is a vector of controls: class size, student age, and dummy variables that identify classes that meet one day a week and students who are male. The coefficient on $early_n$ measures the expected difference between grades for 8:00 a.m. classes and grades for classes held later in the day, holding the values of the controls constant. That is, β is the marginal effect of an 8:00 a.m. start time on grades. Hypothesis 1 predicts that $\beta < 0$.

As it stands, however, equation (1) is unlikely to yield satisfactory tests of Hypothesis 1. We say this because it includes a relatively small set of controls. This is unavoidable because we have only limited information about student, instructor, and class characteristics. Accordingly, we augment equation (1) with a full set of student dummy variables, a full set of instructor dummy variables, and a full set of class dummy variables.[6] These dummy variables are included in the specification used to test Hypothesis 1 to account for unobserved heterogeneity at the student, instructor, and class levels. Suppose, for example, that some of the factors that influence academic achievement, such as gender, are fixed for a given student. Omitting these fixed factors from the model would typically result in biased estimates of the coefficients of interest.

By including a full set of student dummy variables, we avoid the bias that would otherwise arise from fixed student-specific factors. The dummy variable for a given student equals 1 for all grades earned by that student and 0 for all grades earned by other students. Thus standard results imply that the resulting estimate of β is identical to the estimate obtained by demeaning the dependent and explanatory variables using student-specific means, and then fitting the model using the demeaned observations. Demeaning causes all determinants of grades that are fixed for a given student to drop out of the model. Similarly, including a full set of instructor dummy variables and class dummy variables causes all determinants of grades that are fixed for a given instructor and a given class to drop out of the model.

We specify the augmented version of equation (1) as follows. Let *S*, *C*, and *I* denote the number of students, classes, and instructors that appear in the dataset, respectively. The model is

$$grade_n = \alpha + \beta early_n + \sum_{s=1}^{S} \delta_s\, stud_{ns} + \sum_{c=1}^{C} \gamma_c class_{nc} + \sum_{i=1}^{I} \kappa_i inst_{ni} + \lambda' x_n + \varepsilon_n, \tag{2}$$

where $stud_{ns}$, $class_{nc}$, and $inst_{ni}$ denote student, class, and instructor dummy variables, e.g., $stud_{ns}$ takes a value of 1 if the *n*th grade observation is for student *s* and 0 otherwise.

There are several reasons to believe that the model in equation (2) should be more robust than that in equation (1). First, by including student dummy variables, we not only account for the impact of innate ability on grades, we also account for a number of other fixed factors that could impact grades, such as gender and socioeconomic status. Second, by including class dummy variables, we account for factors that are fixed for a given class, but vary across classes. Third, by

[6] In the econometrics literature, our approach is known as least-squares dummy variable estimation, and the student, instructor, and class dummy variables are known as student, instructor, and class fixed effects. See, for example, chapter 10 of Wooldridge (2010).

including instructor dummy variables, we account for the impact of differences in instructor quality and idiosyncratic grading practices. This is important because a significant portion of the variation in grades across classes may be due to differences in grading across instructors (Baird, 1984).

But we should also point out that robustness to unobserved heterogeneity at the student level comes at a cost. Earlier we noted that the estimate of β obtained by adding student dummies to equation (1) is identical to the estimate obtained by demeaning the dependent and explanatory variables using student-specific means, and then fitting the model using the demeaned observations. Thus one cost is that a student must have more than one grade observation in the dataset to be included in the sample used to estimate the model. The larger issue, however, is that the estimate of β is determined by the within-individual differences in grades (i.e., the variation in each student's grades around his or her average grade across all classes taken). In effect, the information about differences in average grades between students is discarded, making it more difficult to obtain precise estimates of the effect of an early class start time on grades.[7]

In general, therefore, we anticipate that regressions that include student dummies to produce conservative estimates of the class scheduling effect, because the estimates do not utilize the information conveyed by differences in average grades between students. In other words, we expect these regressions to understate the true impact of class scheduling on grades. This is the price that we have to pay in order to guard against the bias that could be introduced by failing to account for unobserved heterogeneity at the student level.

Strategy for Testing Hypothesis 2

To test Hypothesis 2, we need to identify the effect of an early class start time for qualitative classes separately from that for quantitative classes. We do so by fitting a regression of the form

$$grade_n = \alpha + \beta_{qt} early\_quant_n + \beta_{ql} early\_qual_n + \sum_{s=1}^{S} \delta_s\, stud_{ns} + \sum_{c=1}^{C} \gamma_c class_{nc} + \sum_{i=1}^{I} \kappa_i inst_{ni} + \lambda' x_n + \varepsilon_n, \tag{3}$$

where $early\_quant_n$ equals 1 if the nth observation is for a quantitative class with an 8:00 a.m. start time and 0 otherwise, and $early\_qual_n$ equals 1 if the nth observation is for a qualitative class with an 8:00 a.m. start time and 0 otherwise. The coefficients β_{qt} and β_{ql} are the marginal effects of an 8:00 a.m. start time on grades in quantitative and qualitative classes, respectively. Hypothesis 1 predicts that $\beta_{\mathrm{qt}} < 0$ and $\beta_{\mathrm{ql}} < 0$. Hypothesis 2 predicts that $\beta_{\mathrm{qt}} < \beta_{\mathrm{ql}}$.

Strategy for Testing Hypothesis 3

To test Hypothesis 3, we need to identify the effect of an early class start time for single-section classes separately from that for classes with multiple sections. We do so by fitting a regression of the form

---

[7] Similar arguments apply with respect to the instructor dummies and class dummies. However, the number of unique classes and unique instructors is small compared to the number of unique students. Thus most of the information loss is due to the inclusion of student dummies.

$$grade_n = \alpha + \beta_s early\_sing_n + \beta_m early\_mult_n + \sum_{s=1}^{S} \delta_s \, stud_{ns} + \sum_{c=1}^{C} \gamma_c class_{nc} + \sum_{i=1}^{I} \kappa_i inst_{ni} + \lambda' x_n + \varepsilon_n, \tag{4}$$

where $early\_sing_n$ equals 1 if the nth observation is for a single-section class with an 8:00 a.m. start time and 0 otherwise, and $early\_mult_n$ equals 1 if the nth observation is for a multiple-section class with an 8:00 a.m. start time and 0 otherwise. The coefficients β_s and β_m are the marginal effects of an 8:00 a.m. start time on grades in single- and multiple-section classes, respectively. Hypothesis 1 predicts that $\beta_s < 0$ and $\beta_m < 0$. Hypothesis 3 predicts that $\beta_s < \beta_m$.

One limitation of the model in equation (3) is that it does not address the issue of whether the self-selection mechanism that unpins Hypothesis 3 differs across quantitative and qualitative classes. If the type of class (quantitative versus qualitative) influences the propensity of students to engage in self-selection, then equation (3) is too restrictive. To investigate whether this is the case, we also fit a regression of the form

$$grade_n = \alpha + \beta_{qts} early\_quant\_sing_n + \beta_{qtm} early\_quant\_mult_n + \beta_{qls} early\_qual\_sing_n + \beta_{qlm} early\_qual\_mult_n + \sum_{s=1}^{S} \delta_s \, stud_{ns} + \sum_{c=1}^{C} \gamma_c class_{nc} + \sum_{i=1}^{I} \kappa_i inst_{ni} + \lambda' x_n + \varepsilon_n, \tag{5}$$

where $early\_quant\_sing_n$ equals 1 if the nth observation is for a single-section quantitative class with an 8:00 a.m. start time and 0 otherwise, $early\_quant\_mult_n$ equals 1 if the nth observation is for a multiple-section quantitative class with an 8:00 a.m. start time and 0 otherwise, $early\_qual\_sing_n$ equals 1 if the nth observation is for a single-section qualitative class with an 8:00 a.m. start time and 0 otherwise, and $early\_qual\_mult_n$ equals 1 if the nth observation is for a multiple-section qualitative class with an 8:00 a.m. start time and 0 otherwise.

Alternative Methods of Estimation and Inference

The models discussed thus far allow us to compare the grades earned in classes held during the earliest time slot used by the university to the grades earned in all other classes. In specifying these models, we have implicitly interpreted the phrase "later in the day" from Hypothesis 1 to mean "a 9:30 a.m. or later start time." It is conceivable that this interpretation could lead us to underestimate the effect of early morning classes on grades. Perhaps classes that start at 9:30 a.m. are also "early" in the sense that grades in these classes are systematically lower than grades in classes held during later time slots. If so, then adopting a narrower interpretation of "later in the day," such as "an 11:00 a.m. or later start time," would be more appropriate. We therefore assess the sensitivity of our findings to how "later in the day" is interpreted by fitting equations (3), (4), and (5) with 9:30 a.m. classes excluded from the dataset.

EMPIRICAL RESULTS

Table 1 provides descriptive statistics for the dataset. The top panel reports the mean, standard deviation, minimum, and maximum of the variables used in the analysis. The bottom panel reports the same statistics by discipline for the grade, time, and class size variables. We show the number of grade observations for each discipline next to its abbreviated name in parentheses.

The student characteristics look about as expected. The average age is a little over 23, with a range of 18 to 62, and about 46% of the students are male. Most of the older students are enrolled in evening classes that meet once a week. The average grade earned across all disciplines is 2.82 with a standard deviation of 1.04. This may seem somewhat high at first glance. However, most business classes are closed to freshman and sophomores, so the bulk of the grade observations are for upperclassmen. The average class size is about 33 students, with a range of 1 to 53. The low end of this range reflects the way in which a small number of honors classes are taught. Each honors class is combined with a regular section of the same class for lecture purposes, i.e., the two sections meet at the same time of day in the same classroom and are taught by the same instructor. This makes it feasible to have honors classes without setting a minimum on enrollments.

As anticipated, breaking the statistics down by discipline reveals that the quantitative ones — accounting, economics, finance, and quantitative methods — have lower average grades than the qualitative ones. The average for the former ranges from 2.44 for economics to 2.63 for accounting, while that for the latter ranges from 3.02 for management to 3.22 for general business administration. Recall that we control for these differences in our multiple regression specifications by including a full set of class dummy variables. There is also some variation in the average time at which classes are held across disciplines. The range is from about 11:15 a.m. for general business administration to about 2:00 p.m. for quantitative methods.

Table 2 illustrates the unconditional relation between grades and average class start time. The first panel is for classes that meet twice per week. The second is for a small number of evening classes that meet once per week. Overall there is little indication of a clear pattern for either type of class. If anything, the figures in the first panel suggest that low grades are more likely to occur in classes that meet later in the day. The average start time for an A grade is about 12:30 p.m. while that for a D grade is about 1:00 pm.

It would be premature, however, to form conclusions about the relation between grades and class scheduling based solely on Table 2. A simple breakdown of average class time by grade fails to control for any of the other factors that might influence academic achievement. To draw reliable inferences, we need to examine the results from the multiple regression analysis.

Evidence from the Baseline Regression Specifications

To highlight the impact of controlling for fixed student-specific factors on grades, we report the results obtained both with and without the use of student dummy variables. Table 3 is for the case in which we omit the student dummies. The *, **, and *** superscripts indicate statistical significance at the 10%, 5%, and 1% levels, respectively, taking into account the directional nature of our hypotheses. For example, let $\hat{\beta}$ denote our estimate of β in equation (2). If $\hat{\beta}$ has one star, then we reject the restriction $\beta = 0$ in favor of the one-sided alternative $\beta < 0$ at the 10% significance level.[8] That is, we conclude that the data provide sufficient evidence that the sign of

[8] If Hypothesis 1 holds, then the regression residuals for classes offered during a given time slot are likely to be correlated. We therefore cluster the standard errors used to compute the *t*-statistics by class start time. That is, we compute the standard errors in a way that captures the impact of any correlation between the residuals for classes offered during a given time slot. This is consistent with the approach used by other studies that employ similar data and methods (see, for example, Dills and Hernandez-Julian, 2008). We refer the interested reader to Thompson (2011) for a detailed discussion of clustering standard errors using standard statistical software packages.

β is negative, as predicted by Hypothesis 1.

We begin with column (1) of Table 3, which shows the estimated slope on the 8:00 a.m. dummy ($\hat{\beta}$) for the model in equation (2). It is –0.082, has a standard error of 0.037, and is statistically significant at the 5% level using a one-tailed t-test. Thus the estimated marginal effect of starting a class at 8:00 a.m. rather than later in the day is about eight hundredths of a grade point, and the sign of the estimate is consistent with Hypothesis 1. In other words, the regression evidence indicates that early morning classes have an adverse impact on the academic performance of business students.

Next we consider the results in column (2), which reports the estimated slopes on the 8:00 a.m. dummies for quantitative and qualitative classes ($\hat{\beta}_{qt}$ and $\hat{\beta}_{ql}$) for the model in equation (3). Both estimates are negative, as predicted by Hypothesis 1, but $\hat{\beta}_{qt}$ is larger than $\hat{\beta}_{ql}$ in magnitude: –0.188 versus –0.036. In addition, the former is statistically significant at the 1% level, while the latter is statistically insignificant, and we reject the restriction $\beta_{qt} = \beta_{ql}$ in favor of $\beta_{qt} < \beta_{ql}$ at the 5% level (p-value of 0.01). Thus the regression evidence is consistent with the predictions of Hypothesis 2. It indicates that the deleterious effect of early morning classes on grades is more pronounced for quantitative classes than for qualitative classes.

The results in column (3), which are for the model in equation (4), follow a similar pattern. The estimated slopes on the 8:00 a.m. dummies for single- and multiple-section classes ($\hat{\beta}_s$ and $\hat{\beta}_m$) are negative, but the former is larger than the latter in magnitude: –0.186 versus –0.074. Furthermore, $\hat{\beta}_s$ is statistically significant at the 5% level, and $\hat{\beta}_m$ is statistically significant at the 10% level. It turns out, however, that we cannot reject the restriction $\beta_s = \beta_m$ at the 10% level (p-value of 0.11). Although the estimates are clearly suggestive of the self-selection effect that unpins Hypothesis 3, the evidence is insufficient to conclude that the deleterious effect of early morning classes on grades is more pronounced for single-section classes.

The results in column (4), which are for the model in equation (5), provide insights on whether the self-selection effect differs across class types. But we should point out that data limitations are an important consideration in this regard. Of the 833 grade observations for 8:00 a.m. classes that are contained in our dataset, 583 are for multiple-section qualitative classes, 190 are for multiple-section quantitative classes, 58 are for single-section qualitative classes, and 2 are for single-section quantitative classes. There are only 2 observations for the final category because the only class that falls into this category for our sample period is an honors section of an economics class. We obviously need to be cautious about using an estimate based on only 2 data points to draw anything other than tentative inferences, especially given that the academic performance of honors students is likely to be considerably better than that of the average student.

Indeed, we find that the estimated slope on the 8:00 a.m. dummy for single-section quantitative classes ($\hat{\beta}_{qts}$) is positive, albeit not statistically significant, which indicates that the average grade in the 8:00 a.m. honors section is higher than the average grade for all other classes held later in the day. This finding is consistent with a scenario in which the impact of the early morning class time is masked by the higher innate ability of honors students. Because this sort of conflating of effects can be overcome by fitting specifications that include student-specific dummy variables, we will focus on the remaining three categories of classes for the time being.

We begin by noting that the values of the 8:00 a.m. dummies for multiple-section quantitative, single-section qualitative, and multiple-section qualitative classes ($\hat{\beta}_{qtm}$, $\hat{\beta}_{qls}$ and $\hat{\beta}_{qlm}$) are negative, as predicted by Hypothesis 1, with both $\hat{\beta}_{qtm}$ and $\hat{\beta}_{qls}$ displaying statistical significance at the 1% level. The estimates also provide clear support for Hypothesis 2, with the restriction

$\beta_{qtm} = \beta_{qlm}$ rejected in favor of $\beta_{qtm} < \beta_{qlm}$ at the 1% level (p-value of 0.009), and stronger support for Hypothesis 3 than when we do not distinguish between class types. The spread between $\hat{\beta}_{qls}$ and $\hat{\beta}_{qlm}$ is wider than the spread between $\hat{\beta}_s$ and $\hat{\beta}_m$, which points to a larger self-selection effect, and the restriction $\beta_{qls} = \beta_{qlm}$ is rejected in favor of $\beta_{qls} < \beta_{qlm}$ at the 5% level (p-value of 0.03). Thus the evidence indicates that the deleterious effect of early morning qualitative classes on grades is more pronounced for single-section classes than for multiple-section classes.

Table 4 illustrates how the regression results change when we include a full set of student dummy variables in the model. In comparison to Table 3, the value of $\hat{\beta}$ for the model in equation (2) rises from –0.082 to –0.017, and its standard error rises from 0.037 to 0.049. We therefore find that discarding information about differences in average grades between students (i.e., including student-specific dummy variables) weakens the evidence of the class scheduling effect to the point that the $\hat{\beta}$ is no longer statistically significant.

As mentioned earlier, finding weaker evidence of a class scheduling effect after controlling for unobserved student heterogeneity is not unanticipated. Note that the adjusted R-squared for the model increases from 18.2% in Table 3 to 71.1% in Table 4, which indicates that fixed student-specific characteristics explain a large portion of the variation in grades between students. Although it is challenging to obtain precise estimates of the class scheduling effect based only on the variation in grades around student-specific means, the results in Table 3 suggest the effect of early morning classes on grades is more pronounced for quantitative disciplines than for qualitative disciplines. If this is true, then restricting the coefficient on the 8:00 a.m. dummy variable to be the same for all classes would tend to obscure the existence of the class scheduling effect.

The results in column (2) of Table 4 are consistent with a larger scheduling effect for quantitative disciplines. The values of $\hat{\beta}_{qt}$ and $\hat{\beta}_{ql}$ are –0.069 and 0.001, with standard errors of 0.046 and 0.057. Thus $\hat{\beta}_{qt}$ has the sign predicted by Hypothesis 1, and it is statistically significant at the 10% level. In addition, the restriction $\beta_{qt} = \beta_{ql}$ is rejected in favor of $\beta_{qt} < \beta_{ql}$ at the 10% level (p-value of 0.089), which is consistent with the predictions of Hypothesis 2. This finding reinforces the results reported in Table 3. In both cases, the evidence indicates that the deleterious effect of early morning classes on grades is more pronounced for quantitative classes than for qualitative classes.

Similarly, the results in column (3) are supportive of Hypothesis 3. The values of $\hat{\beta}_s$ and $\hat{\beta}_m$ are –0.210 and –0.006, with standards error of 0.117 and 0.049. Thus both estimates have the sign predicted by Hypothesis 1, and $\hat{\beta}_s$ is statistically significant at the 10% level based on a one-tailed t-test. In addition, we reject the restriction $\beta_s = \beta_m$ in favor of $\beta_s < \beta_m$ at the 10% level (p-value of 0.052), which is consistent with Hypothesis 3. Once again, therefore, the findings reinforce the results reported in Table 3. The regression evidence indicates that the deleterious effect of early morning classes on grades is more pronounced for single-section classes.

Turning to the final set of estimates in column (4), we see some interesting changes from the results in Table 3. First, $\hat{\beta}_{qts}$ falls from 0.251 to –0.622. Although it remains statistically insignificant, it now has the sign predicted by Hypothesis 1. The negative sign indicates that students in the 8:00 a.m. honors section earned grades for the class that were well below their overall average grades. Putting aside the issue of statistical significance, the drop in the value of $\hat{\beta}_{qts}$ from Table 3 illustrates that controlling for fixed student-specific factors, such as innate ability, can produce notable changes in the regression estimates.

Second, $\hat{\beta}_{qtm}$, $\hat{\beta}_{qls}$ and $\hat{\beta}_{qlm}$ have larger standard errors than in Table 3. This is not surprising given that the student-specific dummy variable explain a large portion of the variation in grades between students. The increase in the standard errors, along with changes in the coefficient estimates, leads to a reduction in statistical significance relative to Table 3. Although $\hat{\beta}_{qtm}$ and $\hat{\beta}_{qls}$ have the anticipated sign, only $\hat{\beta}_{qls}$ is estimated with sufficient precision to be statistically significant at the 10% level.

Evidence from Alternative Regression Specifications

We now take up the question of whether our regression estimates are sensitive to how we define the phase "later in the day" in Hypothesis 1. As noted in Section 3, it is conceivable that 9:30 a.m. classes should also be classified as "early" in the sense that grades for these classes are systematically lower than those for classes that meet during later time slots. If this is true, then we should not be treating 9:30 a.m. classes as "later in the day" when estimating the scheduling effect for 8:00 a.m. classes, as this would bias our estimates of the scheduling effect downwards. To assess whether this is a concern, we fit equations (3), (4) and (5) a second time with 9:30 a.m. classes excluded from the dataset. Table 5 presents the results.

Interestingly, the evidence that students perform worse in 8:00 a.m. classes is more definitive than that in Table 4. The value of $\hat{\beta}$ for the model in equation (2) is –0.095, and it is statistically significant at the 5% level using a one-tailed t-test. Hence we no longer have to distinguish between quantitative and qualitative classes to find evidence consistent with Hypothesis 1. In addition, $\hat{\beta}_{qt}$ and $\hat{\beta}_{ql}$ for the model in equation (3) are negative, statistically significant (at the 5% and 10% level, respectively), and the former is larger than the latter in magnitude, as predicted by Hypothesis 2. Similarly, $\hat{\beta}_{s}$ and $\hat{\beta}_{m}$ are negative, statistically significant (at the 5% and 10% level, respectively), and the former is larger than the latter in magnitude, as predicted by Hypothesis 3. Overall, however, these estimates offer weaker support for Hypotheses 2 and 3 than those in Table 4 from the standpoint of producing statistically-significant t-statistics for the restrictions $\beta_{qt} = \beta_{ql}$ and $\beta_{s} = \beta_{m}$.

The most dramatic change from Table 4 is that $\hat{\beta}_{qts}$ falls from –0.622 to –1.193. Hence the estimated scheduling effect for students in the 8:00 a.m. honors section is more than a full grade point. Although the standard error is quite large, the estimate is statistically significant at the 5% level. Of course we should always maintain a healthy skepticism about estimates based on very small samples. Nonetheless, it is interesting that the estimated scheduling effect is so large for the only single-section quantitative class that meets at 8:00 a.m. As for $\hat{\beta}_{qtm}$, $\hat{\beta}_{qls}$ and $\hat{\beta}_{qlm}$, these estimates are much smaller than $\hat{\beta}_{qts}$ in magnitude. But all are negative, and each is further away from zero than the corresponding estimate in Table 4. In addition, we find that both $\hat{\beta}_{qtm}$ and $\hat{\beta}_{qls}$ are statistically significant at the 10% level. All of these findings suggest that it may be appropriate to characterize 9:30 a.m. classes as "early" in the sense of Hypothesis 1.

To investigate further, we look at the direct estimates of the scheduling effect for 9:30 a.m. classes. This is accomplished by fitting equations (2), (3), (4) and (5) using 9:30 a.m. classes as the early category, with 8:00 a.m. classes excluded from the dataset. Note that excluding 8:00 a.m. classes allows us to cleanly identify the scheduling effect of interest (i.e., the effect of holding a class at 9:30 a.m. as opposed to later in the day). Table 6 presents the results.

We find that $\hat{\beta}$ for the model in equation (2) is –0.046, and it is statistically significant at the 1% level using a one-tailed t-test. This is consistent with Hypothesis 1. We also find that the values of $\hat{\beta}_{qt}$ and $\hat{\beta}_{ql}$ for the model in equation (3) are negative, and the former is larger than the latter in magnitude, as predicted by Hypothesis 2. As in Table 4, however, $\hat{\beta}_{ql}$ is not statistically significant. The spread between $\hat{\beta}_{qt}$ and $\hat{\beta}_{ql}$ is also about half as large as in Table 4, and we fail to reject the restriction $\beta_{qt} = \beta_{ql}$ at the 10% level. In other words, excluding 8:00 a.m. classes from the analysis weakens the evidence that the effect of early morning classes on grades is more pronounced for quantitative classes than for qualitative classes.

The values of $\hat{\beta}_s$ and $\hat{\beta}_m$ for the model in equation (4) have similar implications with respect to the self-selection effect. Although both are negative, the former is *smaller* than the latter in magnitude, which goes against the prediction of Hypothesis 3. We therefore find that excluding 8:00 a.m. classes from the analysis weakens the evidence that the effect of early-morning classes on grades is more pronounced for single-section classes than for multiple-section classes.

In view of the results in Table 6, we conclude that Hypotheses 2 and 3 find less support in the data when we exclude 8:00 a.m. classes from the analysis. But this finding is not wholly unexpected. In an earlier study of class scheduling effects, Dills and Hernandez-Julian (2008) report that grades increase in a gradual fashion as the class start time is moved from early morning to later in the day. Accordingly, we should not be surprised to find that our analysis using 8:00 a.m. classes produces more definitive evidence of scheduling effects than that 9:30 a.m. classes. This outcome simply suggests that the scheduling effects are smaller, and hence harder to detect, for 9:30 a.m. classes than for 8:00 a.m. classes. If this is true, then finding support for Hypotheses 2 and 3 becomes a more difficult task as well.

CONCLUSIONS

The proposition that class scheduling has an impact on academic achievement finds empirical support in several recent studies. Building on this research, we find that class scheduling has a stronger impact on student performance for quantitative classes than for qualitative classes. The evidence also indicates that the scheduling effect is more pronounced for single-section classes than for multiple-section classes. The evidence that scheduling is less important for multiple-section classes points to self-selection as one mechanism that students use to mitigate the adverse impact of early-morning classes on grades.

How should faculty and administrators respond to mounting empirical evidence that class scheduling policies have some influence on student grades? To answer this question, we need to consider both the costs and benefits of any proposed policy changes. Our analysis suggests that the benefits differ across disciplines. In quantitative disciplines, such as accounting and finance, we estimate that moving the start time of a class from 8:00 a.m. to an 11:30 a.m. or later time slot increases grades by about one to two tenths of a grade point, all else being equal. If this estimate accurately reflects student learning, then one can make a reasoned argument in favor of reducing the prevalence of early morning quantitative classes.

The counterargument is that revamping traditional scheduling practices might entail substantial costs. For example, it might be impossible to eliminate 8:00 a.m. classes without adding additional classroom space to accommodate the increased number of classes that would need to be taught later in the day. Implementing such a policy might also require additional faculty lines. The cost of reducing or eliminating early morning classes is likely to be institution specific, or even department specific, so determining the appropriate policy must necessarily occur at the local level. If the costs of rescheduling early morning classes are not prohibitive, then doing so may be worthwhile.

In addition, department chairs in quantitative disciplines may want to pay more attention to the nature of the classes that are offered during early morning time slots. It might be better, for example, to offer an advanced cost accounting class during the late morning or early afternoon than at 8:00 a.m. Class scheduling practices are typically flexible enough to permit some swapping of time slots for different classes within the same discipline. If this is the case, then scheduling the least quantitatively demanding classes during the early morning might prove beneficial.

References

Au, R., M. Carskadon, R. Millman, A. Wolfson, P. Braverman, W. Adelman, C. Breuner et al. 2014. School start times for adolescents. *Pediatrics* 134: 642–649.

Austin, T., R. Fennell, and C. Yeager. 1988. Class scheduling and academic achievement in a non-traditional graduate program. *Innovative Higher Education* 12: 79–90.

Baird, L. 1984. Predicting predictability: The influence of student and institutional characteristics on the prediction of grades. *Research in Higher Education* 21: 261–279.

Burke, L., K. James, and M. Ahmadi. 2009. Effectiveness of PowerPoint-based lectures across different business disciplines: An investigation and implications. *Journal of Education for Business* 84: 246–251.

Cameron, A., and P. Trivedi. 2005. Microeconometrics: Methods and Applications. Cambridge University Press, New York.

Carskadon, M., S. Labyak, C. Acebo, and R. Seifer. 1999. Intrinsic circadian period of adolescent humans measured in conditions of forced desynchrony. *Neuroscience Letters* 260: 129–132.

Carskadon, M., A. Wolfson, C. Acebo, O. Tzischinsky, and R. Seifer. 1998. Adolescent sleep patterns, circadian timing, and sleepiness at a transition to early school days. *Sleep* 21: 871–81.

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

Cohen, N., and L. Squire. 1980. Preserved learning and retention of pattern-analyzing skill in amnesia: dissociation of knowing how and knowing that. *Science* 210: 207–210.

Conway, J., and C. Smith. 1994. REM sleep and learning in humans: a sensitivity to specific types of learning tasks. Proceedings of the 12th Congress of the European Sleep Research Society.

Curcio, G., M. Ferrara, and L. De Gennaro. 2006. Sleep loss, learning capacity and academic performance. *Sleep Medicine Reviews* 10: 323–337.

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

Edwards, F. 2012. Early to rise? The effect of daily start times on academic performance. *Economics of Education Review* 31: 970–983.

Kidwell, L., and R. Kidwell. 2008. Do the Numbers Add Up to Different Views? Perceptions of Ethical Faculty Behavior Among Faculty in Quantitative Versus Qualitative Disciplines. *Journal of Business Ethics* 78:141–151.

Loo, R. 2002. The distribution of learning styles and types for hard and soft business majors. *Educational Psychology* 22: 349–360.

Marshall, L., and J. Born. 2007. The contribution of sleep to hippocampus-dependent memory consolidation. *Trends in Cognitive Sciences* 11: 442–450.

Medeiros, A., D. Mendes, P. Lima, and J. Araujo. 2001. The relationships between sleep-wake cycle and academic performance in medical students. *Biological Rhythm Research* 32: 263–70.

Peigneux, P., S. Laureys, S. Fuchs, F. Collette, F. Perrin, J. Reggers, C. Phillips, C. Degueldre, G.

Del Fiore, J. Aerts, A. Luxen, and P. Maquet. 2004. Are spatial memories strengthened in the human hippocampus during slow wave sleep? *Neuron* 44: 535–545.

Petersen, M. 2009. Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies* 22: 435–480.

Plihal W., and J. Born. 1997. Effect of early and late nocturnal sleep on declarative and procedural memory. *Journal of Cognitive Neuroscience* 9: 534–547.

Pope, N. 2014. How time of day affects productivity: Evidence from school schedules. Working paper. The University of Chicago.

Schlee, R. 2005. Social styles of students and professors: Do students' social styles influence their preferences for professors? *Journal of Marketing Education* 27: 130–142.

Schlee, R., and K. Harich. 2014. Teaching creativity to business students: How well are we doing? *Journal of Education for Business* 89: 133–141.

Shinkoda, H., K. Matsumoto, Y. Park, and H. Nagashima. 2000. Sleep-wake habits of school children according to grade. *Psychiatry and Clinical Neurosciences* 54: 287–289.

Smith, C. 1993. REM sleep and learning: some recent findings. In: The Functions of Dreaming, edited by Moffitt A, Kramer M, Hoffmann R, editors. Albany, NY: State University of New York Press, 341–361.

Smith, C. 1995. Sleep states and memory processes. *Behavioural Brain Research* 69: 137–145

Smith C. 2001. Sleep states and memory processes in humans: procedural versus declarative memory systems. *Sleep Medicine Reviews* 5: 491–506.

Smith, C., and D. Smith. 2003. Ingestion of ethanol just prior to sleep onset impairs memory for procedural but not declarative tasks. *Sleep* 26: 185–191.

Thompson, S. 2011. Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics* 99: 1–10.

Walker, M. 2008. Sleep-dependent memory processing. *Harvard Review of Psychiatry* 16: 287–298.

Wooldridge, J. 2010. Econometric Analysis of Cross Section and Panel Data. The MIT Press. Cambridge, MA.

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

Table 1: Descriptive Statistics for Full Sample and by Discipline

Variable	Mean	SD	Min	Max
		Full Sample (N=10,039)		
Grade	2.82	1.04	0.00	4.00
Time	12.83	3.00	8.00	18.50
Class Size	33.31	8.44	1.00	53.00
Age	23.25	4.36	18.00	62.00
Male	0.46	0.50	0.00	1.00
Freshman	0.04	0.21	0.00	1.00
Sophomore	0.12	0.33	0.00	1.00
Junior	0.27	0.44	0.00	1.00
Senior	0.56	0.50	0.00	1.00
		By Discipline		
ACCT (N=1,424)				
Grade	2.63	1.14	0.00	4.00
Time	13.59	2.75	8.00	18.50
Class Size	30.95	7.02	6.00	42.00
BADM (N=1,128)				
Grade	3.22	0.83	0.00	4.00
Time	11.26	3.66	8.00	18.50
Class Size	37.50	4.70	1.00	43.00
ECON (N=1,762)				
Grade	2.44	1.16	0.00	4.00
Time	11.96	2.51	8.00	17.00
Class Size	36.36	9.10	1.00	46.00
FINC (N=642)				
Grade	2.59	1.04	0.00	4.00
Time	13.71	2.13	9.50	18.50
Class Size	25.57	8.18	6.00	36.00
MGMT (N=3,012)				
Grade	3.02	0.86	0.00	4.00
Time	13.45	3.03	8.00	18.50
Class Size	31.40	7.51	5.00	45.00
MKTG (N=1,135)				
Grade	3.08	0.81	0.00	4.00
Time	11.73	2.61	8.00	17.00
Class Size	32.02	9.34	5.00	53.00
QMTH (N=936)				
Grade	2.52	1.22	0.00	4.00
Time	13.96	2.57	9.50	18.50
Class Size	39.10	6.14	1.00	53.00

Notes: The top panel of the table reports the mean, standard deviation (SD), minimum (Min), and maximum (Max) of the key variables used in the empirical analysis. The bottom panel reports the same statistics by discipline for the grade, time, and class size variables. The seven disciplines considered are accounting (ACCT), general business administration (BADM), economics (ECON), finance (FINC), management (MGMT), marketing (MKTG), and quantitative methods (QMTH). The number of grade observations for each discipline is reported next to its abbreviated name in parentheses.

Table 2: Average Class Start Time by Grade Earned

Grade	N	Mean	SD
	Two Days per Week Courses		
A	2, 727	12.55	2.94
B	3, 661	12.47	2.88
C	2, 183	12.70	2.78
D	625	12.96	2.59
F	375	12.69	2.55
	One Day per Week Courses		
A	154	18.12	0.65
B	192	17.71	0.75
C	104	17.88	0.74
D	9	18.17	0.66
F	9	18.00	0.75

Notes: The table illustrates the unconditional relation between grades and average class start time. The first panel is for classes that meet two days per week. The second is for the small number of evening classes that meet one day per week, primarily at 6:30 p.m. for 150 minutes.

Table 3. Estimates of Scheduling Effects for 8:00 a.m. Classes with No Student Dummies

Dummy variable (coefficient)	Model 1 $\hat{\beta}$/se	Model 2 $\hat{\beta}$/se	Model 3 $\hat{\beta}$/se	Model 4 $\hat{\beta}$/se
8:00 a.m. (β)	-0.082**			
	(0.037)			
8:00 a.m. Quant (β_{qt})		-0.188***		
		(0.030)		
8:00 a.m. Qual (β_{ql})		-0.036		
		(0.053)		
8:00 am Single (β_s)			-0.186**	
			(0.069)	
8:00 am Multiple (β_m)			-0.074*	
			(0.040)	
8:00 am Single Quant (β_{qts})				0.251
				(0.202)
8:00 am Multiple Quant (β_{qtm})				-0.193***
				(0.033)
8:00 am Single Qual (β_{qls})				-0.207***
				(0.065)
8:00 am Multiple Qual (β_{qlm})				-0.020
				(0.058)
R-Squared	0.182	0.182	0.182	0.183
Sample Size	10039	10039	10039	10039
H2 t-statistic		-2.904**		-3.100***
H3 t-statistic			-1.319	-2.264**

Notes: Models 1, 2, 3, and 4 are the regression specifications obtained by omitting the student-specific dummy variables (i.e., student fixed effects) from equations (2), (3), (4), and (5), respectively. We report standard errors (clustered by class start time) below the estimates in parentheses. The *, **, and *** superscripts indicate that we reject the restriction that the coefficient is zero in favor of the one-sided alternative that it is less than zero at the 10%, 5%, and 1% levels, respectively. All specifications include the following controls: log of class size, age of the student, gender of the student, and a dummy variable for classes that meet only once a week. The t-statistics reported for Models 2, 3, and 4 are for tests of the following restrictions: H_0: $\beta_{qt} = \beta_{ql}$ versus H_A: $\beta_{qt} < \beta_{ql}$, H_0: $\beta_s = \beta_m$ versus H_A: $\beta_s < \beta_m$, H_0: $\beta_{qtm} = \beta_{qlm}$ versus H_A: $\beta_{qtm} < \beta_{qlm}$, and H_0: $\beta_{qls} = \beta_{qlm}$ versus H_A: $\beta_{qls} < \beta_{qlm}$. Consider, for example, Model 2. The reported t-statistic is the ratio of $\hat{\beta}_{qt} - \hat{\beta}_{ql}$ to the standard error of this quantity. A negative statistic of sufficient magnitude yields a rejection of the restriction in favor of the one-sided alternative. Rejections at the 10%, 5%, and 1% levels are again indicated using *, **, and *** superscripts.

Table 4. Estimates of Scheduling Effects for 8:00 a.m. Classes

Dummy variable (coefficient)	Model 1 $\hat{\beta}$/se	Model 2 $\hat{\beta}$/se	Model 3 $\hat{\beta}$/se	Model 4 $\hat{\beta}$/se
8:00 a.m. (β)	-0.017			
	(0.049)			
8:00 a.m. Quant (β_{qt})		-0.069*		
		(0.046)		
8:00 a.m. Qual (β_{ql})		0.001		
		(0.057)		
8:00 am Single (β_s)			-0.210*	
			(0.117)	
8:00 am Multiple (β_m)			-0.006	
			(0.049)	
8:00 am Single Quant (β_{qts})				-0.622
				(0.521)
8:00 am Multiple Quant (β_{qtm})				-0.063
				(0.047)
8:00 am Single Qual (β_{qls})				-0.192*
				(0.123)
8:00 am Multiple Qual (β_{qlm})				0.016
				(0.059)
R-Squared	0.711	0.711	0.711	0.711
Sample Size	10039	10039	10039	10039
H2 t-statistic		-1.499*		-1.694*
H3 t-statistic			-1.874*	-1.801*

Notes: Models 1, 2, 3, and 4 are the regression specifications in equations (2), (3), (4), and (5), respectively. We report standard errors (clustered by class start time) below the estimates in parentheses. The *, **, and *** superscripts indicate that we reject the restriction that the coefficient is zero in favor of the one-sided alternative that it is less than zero at the 10%, 5%, and 1% levels, respectively. All specifications include the following controls: log of class size, age of the student, gender of the student, and a dummy variable for classes that meet only once a week. The t-statistics reported for Models 2, 3, and 4 are for tests of the following restrictions: H_0: $\beta_{qt} = \beta_{ql}$ versus H_A: $\beta_{qt} < \beta_{ql}$, H_0: $\beta_s = \beta_m$ versus H_A: $\beta_s < \beta_m$, H_0: $\beta_{qtm} = \beta_{qlm}$ versus H_A: $\beta_{qtm} < \beta_{qlm}$, and H_0: $\beta_{qls} = \beta_{qlm}$ versus H_A: $\beta_{qls} < \beta_{qlm}$. Consider, for example, Model 2. The reported t-statistic is the ratio of $\hat{\beta}_{qt} - \hat{\beta}_{ql}$ to the standard error of this quantity. A negative statistic of sufficient magnitude yields a rejection of the restriction in favor of the one-sided alternative. Rejections at the 10%, 5%, and 1% levels are again indicated using *, **, and *** superscripts.

Table 5. Alternative Estimates of Scheduling Effects for 8:00 a.m. Classes

Dummy variable (coefficient)	Model 1 $\hat{\beta}$/se	Model 2 $\hat{\beta}$/se	Model 3 $\hat{\beta}$/se	Model 4 $\hat{\beta}$/se
8:00 a.m. (β)	-0.095**			
	(0.040)			
8:00 a.m. Quant (β_{qt})		-0.124**		
		(0.062)		
8:00 a.m. Qual (β_{ql})		-0.083*		
		(0.045)		
8:00 am Single (β_s)			-0.273**	
			(0.104)	
8:00 am Multiple (β_m)			-0.081*	
			(0.047)	
8:00 am Single Quant (β_{qts})				-1.193**
				(0.490)
8:00 am Multiple Quant (β_{qtm})				-0.114*
				(0.064)
8:00 am Single Qual (β_{qls})				-0.238*
				(0.123)
8:00 am Multiple Qual (β_{qlm})				-0.066
				(0.054)
R-Squared	0.725	0.725	0.725	0.725
Sample Size	8051	8051	8051	8051
H2 t-statistic		-0.702		-0.842
H3 t-statistic			-1.601*	-1.295

Notes: We construct alternative estimates of the scheduling effects by excluding 9:30 a.m. classes from the dataset. Models 1, 2, 3, and 4 are the regression specifications in equations (2), (3), (4), and (5), respectively. We report standard errors (clustered by class start time) below the estimates in parentheses. The *, **, and *** superscripts indicate that we reject the restriction that the coefficient is zero in favor of the one-sided alternative that it is less than zero at the 10%, 5%, and 1% levels, respectively. All specifications include the following controls: log of class size, age of the student, gender of the student, and a dummy variable for classes that meet only once a week. The t-statistics reported for Models 2, 3, and 4 are for tests of the following restrictions: H_0: $\beta_{qt} = \beta_{ql}$ versus H_A: $\beta_{qt} < \beta_{ql}$, H_0: $\beta_s = \beta_m$ versus H_A: $\beta_s < \beta_m$, H_0: $\beta_{qtm} = \beta_{qlm}$ versus H_A: $\beta_{qtm} < \beta_{qlm}$, and H_0: $\beta_{qls} = \beta_{qlm}$ versus H_A: $\beta_{qls} < \beta_{qlm}$. Consider, for example, Model 2. The reported t-statistic is the ratio of $\hat{\beta}_{qt} - \hat{\beta}_{ql}$ to the standard error of this quantity. A negative statistic of sufficient magnitude yields a rejection of the restriction in favor of the one-sided alternative. Rejections at the 10%, 5%, and 1% levels are again indicated using *, **, and *** superscripts.

Table 6. Estimates of Scheduling Effects for 9:30 a.m. Classes

Dummy variable (coefficient)	Model 1 $\hat{\beta}$/se	Model 2 $\hat{\beta}$/se	Model 3 $\hat{\beta}$/se	Model 4 $\hat{\beta}$/se
9:30 a.m. (β)	-0.046***			
	(0.007)			
9:30 a.m. Quant (β_{qt})		-0.066**		
		(0.029)		
9:30 a.m. Qual (β_{ql})		-0.029		
		(0.031)		
9:30 a.m. Single (β_s)			-0.011	
			(0.034)	
9:30 a.m. Multiple (β_m)			-0.061***	
			(0.007)	
9:30 a.m. Single Quant (β_{qts})				0.064
				(0.066)
9:30 a.m. Multiple Quant (β_{qtm})				-0.101**
				(0.039)
9:30 a.m. Single Qual (β_{qls})				-0.032
				(0.026)
9:30 a.m. Multiple Qual (β_{qlm})				-0.026
				(0.042)
R-Squared	0.716	0.716	0.716	0.716
Sample Size	9223	9223	9223	9223
H2 t-statistic		-0.637		-0.931
H3 t-statistic			1.256	-0.134

Notes: We construct estimates of the scheduling effects for 9:30 a.m. classes by excluding 8:00 a.m. classes from the dataset. Models 1, 2, 3, and 4 are the regression specifications in equations (2), (3), (4), and (5), respectively. We report standard errors (clustered by class start time) below the estimates in parentheses. The *, **, and *** superscripts indicate that we reject the restriction that the coefficient is zero in favor of the one-sided alternative that it is less than zero at the 10%, 5%, and 1% levels, respectively. All specifications include the following controls: log of class size, age of the student, gender of the student, and a dummy variable for classes that meet only once a week. The t-statistics reported for Models 2, 3, and 4 are for tests of following restrictions: H_0: $\beta_{qt} = \beta_{ql}$ versus H_A: $\beta_{qt} < \beta_{ql}$, H_0: $\beta_s = \beta_m$ versus H_A: $\beta_s < \beta_m$, H_0: $\beta_{qtm} = \beta_{qlm}$ versus H_A: $\beta_{qtm} < \beta_{qlm}$, and H_0: $\beta_{qls} = \beta_{qlm}$ versus H_A: $\beta_{qls} < \beta_{qlm}$. Consider, for example, Model 2. The reported t-statistic is the ratio of $\hat{\beta}_{qt} - \hat{\beta}_{ql}$ to the standard error of this quantity. A negative statistic of sufficient magnitude yields a rejection of the restriction in favor of the one-sided alternative. Rejections at the 10%, 5%, and 1% levels are again indicated using *, **, and *** superscripts.