

NBER WORKING PAPER SERIES

THE IMPACT OF TIME BETWEEN COGNITIVE TASKS ON PERFORMANCE:
EVIDENCE FROM ADVANCED PLACEMENT EXAMS

Ian Fillmore
Devin G. Pope

Working Paper 18436
<http://www.nber.org/papers/w18436>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2012

We are grateful to Derek Neal and seminar participants at the University of Chicago for helpful suggestions. Ian Fillmore would like to thank the Institute of Education Sciences (IES) for funding through the Pre-doctoral Interdisciplinary Research Training Program at the University of Chicago. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Ian Fillmore and Devin G. Pope. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Impact of Time Between Cognitive Tasks on Performance: Evidence from Advanced Placement Exams

Ian Fillmore and Devin G. Pope
NBER Working Paper No. 18436
October 2012
JEL No. D03,I20

ABSTRACT

In many education and work environments, economic agents must perform several mental tasks in a short period of time. As with physical fatigue, it is likely that cognitive fatigue can occur and affect performance if a series of mental tasks are scheduled close together. In this paper, we identify the impact of time between cognitive tasks on performance in a particular context: the taking of Advanced Placement (AP) exams by high-school students. We exploit the fact that AP exam dates change from year to year, so that students who take two subject exams in one year may have a different number of days between the exams than students who take the same two exams in a different year. We find strong evidence that a shorter amount of time between exams is associated with lower scores, particularly on the second exam. Our estimates suggest that students who take exams with 10 days of separation are 8% more likely to pass both exams than students who take the same two exams with only 1 day of separation.

Ian Fillmore
University of Chicago
5545 S Kenwood Ave
2nd Floor Rear
Chicago, IL 60637
ianfillmore@uchicago.edu

Devin G. Pope
Booth School of Business
University of Chicago
5807 South Woodlawn Avenue
Chicago, IL 60637
and NBER
devin.pope@chicagobooth.edu

In life, we all have many projects and tasks that demand our physical and mental energies. These competing demands often require us to make careful tradeoffs as to where we devote our time and strength. This is especially true when we have to schedule two tasks close together. For physically demanding tasks, it is clear that the amount of time between events can significantly affect performance. For example, running two consecutive miles is much harder than running two miles with a rest period in the middle. In fact, some physical events, such as ultra-marathons, leave athletes needing several weeks of recuperation before they can return to peak performance (Chambers et al., 1998). For cognitive tasks, however, the impact of time between events on performance is less clear. For example, imagine a lawyer who is required to argue two cases in the same week or a student who must take two exams a few days apart. Will the temporal closeness of the deadlines negatively affect overall performance? Are one or two days between tasks enough separation to allow for a return to peak mental acuity?

These questions are difficult to study with observational data because of the endogeneity inherent in these types of situations.¹ People who are assigned or volunteer to complete tasks that are scheduled close together may be very different from people who don't face competing deadlines. Selection bias in both the types of tasks and the people who complete them can result in misleading conclusions about the importance of the temporal proximity of tasks on performance.

In this paper, to identify the causal relationship between the temporal proximity of cognitive tasks and performance, we exploit a novel natural experiment made possible by the timing of Advanced Placement (AP) exams. In May of each year, hundreds of thousands of high-school students take AP exams administered by the College Board. For most students, these exams are the culmination of a year of study in an AP course intended to be comparable to college-level work.

¹ These questions are difficult to answer in laboratory settings as well since the experiment would necessarily have to run multiple days and require large incentives to motivate survey participants.

There are currently 33 exams, each covering a different subject area such as Calculus, Chemistry and European History.

We analyze administrative data for a 10% sample of all AP exam takers between 1996 and 2001 who took two (and only two) exams in the same year. We take advantage of the fact that exam dates change from year to year. Therefore, our sample consists of thousands of students who were tested on the same two AP subjects but with varied time gaps between their two exams. Because of this natural variation, we are able to control for the type of students taking the exams and identify the impact on exam performance of plausibly exogenous differences in the time between exams.

Our results indicate that performance significantly improves with more days between exams. Increasing the number of days between exams from 1 to 10 improves the combined point total on the two exams, which ranges from 2 to 10, by approximately 0.15 points (0.07 standard deviations). Increasing the number of days between exams from 1 to 10 improves the probability of passing both exams by 8%. Rather remarkably, this relationship is almost entirely linear, which suggests that increasing the time between exams from 1 to 3 days has a similar impact on performance as increasing the time between exams from 8 to 10 days.

Ancillary analyses show that the benefit of more spacing of exams is driven nearly entirely by the increase in performance on the second exam. We also find significant differences across demographic groups. Females and Asians benefit the most from increasing the time between exams.

Our findings contribute to several strands of literature in both economics and psychology. In psychology, researchers have long recognized the possibility of cognitive fatigue (Ebbinghaus, 1896-1897). A large amount of work has focused on the impact of task length (e.g. total exam time) on average performance. For example, Ackerman and Kanfer (2009) provide a nice review of the psychology literature of cognitive fatigue. They argue that the evidence is inconclusive regarding the impact of exam length on performance and produce empirical results that actually find that

performance can increase with exam length. There is also a large body of literature in psychology that addresses the impact of cognitive load on a variety of outcomes (see Paas, Renkl, and Sweller (2004) for a review). Cognitive load theory is based on the idea that working memory is limited and that performance, reasoning, and learning degrades as the working memory becomes loaded. There is also a large body of literature in cognitive psychology looking at memory and how distributed study can improve recall (see Cepeda, et al. (2006) for a review of this literature). This literature suggests that proper temporal spacing of study can lead to enhanced learning and memory. In economics, recent work has explored the phenomenon of multitasking. Coviello, Ichino, and Persico (2010) show that Italian judges who were randomly assigned to work on several trials in parallel spent more time than if they did the trials one after the other.²

Our paper builds on these literatures by providing causal empirical evidence of the impact of time between tasks on performance in a natural setting. Our findings suggest that time between exams is an important component in the AP-exam-taking setting. These findings are directly useful for testing agencies, parents, and students who must choose classes and exam schedules. Our findings also indirectly inform other non-test-taking environments, such as task assignment and shift length in the work force. Given the possibility of cognitive fatigue in a large number of important economic situations (job search, training programs, etc.) our findings can hopefully lead to increased research on how to optimally account for cognitive fatigue in various economic situations.

The paper is organized in the following way: In Section 1, we provide background information about the Advanced Placement Exam program and discuss the data that we use in our study. In Section 2, we lay out our empirical strategy. We report our results in Section 3, and we conclude with a discussion of our findings and their broader implications in Section 4.

² There is also work in behavioral economics that explores the impact that time-inconsistent preferences can have on performance when there are varying amounts of task separation (Ariely and Wertenbroch (2002) and see DellaVigna (2009) for a review of this literature).

I. Advanced Placement Exams and Data

In May of each year, Advanced Placement (AP) exams are administered to high-school students by the College Board (the same company that administers the SAT college admissions exam). For most students, these exams are the culmination of a year's worth of study in an AP course intended to be comparable to college-level work. In 2011, more than 1.9 million students took at least one AP exam, resulting in nearly 3.5 million total exams taken.³ Exams are currently offered on 33 different subjects and include both multiple-choice and free-response sections. They are graded by college professors and other individuals with expertise in a subject field. Each exam is given a whole-number score from 1 (lowest) to 5 (highest), with the cutoffs for each number determined freshly every year for each subject exam. Students are typically highly motivated to perform well on these exams for at least two reasons. First, high scores on AP exams are thought to impress college admissions committees. More importantly, many colleges and universities offer course credit for passing marks on AP exams.

We obtained administrative data for a 10% random sample of all AP exam takers who took exactly two AP exams in a given year from 1996 through 2001.⁴ This 10% sample results in 238,138 AP exams taken by 119,069 students. Table 1 lists the AP exams taken by the students in our dataset, ordered by subject popularity. United States History, English Language, English Composition, and Calculus have been the most popular exams. Very few students take exams such as Physics C: Electricity and Magnetism, French Language and Culture, or Latin.

No matter what the subject, all of the exams are offered in morning and afternoon sessions from Monday through Friday during a two-week (10 day) period every May. Students are required to

³ This information was obtained from the College Board's website (professionals.collegeboard.com) on Dec. 5, 2011. <http://professionals.collegeboard.com/profdownload/AP-Program-Summary-Report.pdf>

⁴ We thank the College Board for making these data available to us.

take the exam on a specified morning or afternoon.⁵ Figure 1 provides a typical exam schedule (for the year 2012). The College Board clearly schedules exams so as to minimize the likelihood that a student will have to take two exams on the same day. Popular exams tend to be spread out over the two-week period. However, actual exam dates do not remain the same from year to year, a fact important to the empirical strategy that we outline in the next section.

Table 2 provides basic summary statistics for the students in our sample. More than 80% of students are high-school seniors. The average AP exam score for these seniors is lower than the average score for juniors and sophomores, suggesting that there is positive selection on juniors and sophomores who take two AP exams in a given year. 55% of the students in our sample are female and 66% are white. Black and Hispanic students are underrepresented in AP exam taking and receive lower scores on average. We also provide summary statistics for the number of days between the two exams that each student in our sample took. Most students had 2-6 days between their exams. However, 14% of students took exams on consecutive days and 3.6% of students took two exams on the same day. Approximately 10% of students had more than seven days between exams.

II. Empirical Strategy

Estimating the impact of time between two tasks on performance using observational data can be difficult. In most situations, the time between tasks is likely to be endogenous. Selection bias in both the types of tasks and the types of people who complete them can result in misleading conclusions about how the temporal proximity of tasks affects performance. For example, an employee who has to make several important presentations in a short period of time may be very

⁵ Students must take the exam during its scheduled session. If a student is taking two exams that are scheduled at the same time, she may take one of the exams during a make-up period several weeks later. Therefore, we drop any student who is taking exams that are scheduled at the same time.

different from an employee who only rarely makes presentations. Accurate performance metrics also pose a challenge to credibly identifying how time between tasks affects performance in field settings.

We argue that the AP exam program provides an ideal context in which to test the impact of time between two tasks on performance. This context provides both plausibly exogenous differences in time between tasks *and* the opportunity to assess standardized measures of performance.

Our empirical strategy involves comparing the between-students performance on the same two AP subject exams with varying numbers of days of separation due to exam schedules changing every year. For example, in 1998, the Calculus AP exam was offered on Friday of Week 1 and the United States History exam was offered on Monday of Week 2. In 1999, the Calculus exam was offered on Thursday of Week 1 and the United States History exam was offered on Friday of Week 1. Thus, students who were tested on these subjects in 1998 had three days between exams but in 1999 had only one day between exams. Our empirical strategy allows us to ask whether students who had three days between exams scored significantly better than students who had only one day between exams.

The baseline model that we estimate in this paper is

$$(1) Y_{ijt} = \beta D_{ijt} + \gamma X_i + \theta_j + \alpha_t + \varepsilon_{ijt}$$

where Y_{ijt} is an outcome variable (e.g. exam score) for student i , taking exam pair j , in year t . D_{ijt} is the number of days separating the two exams taken by student i in year t , X_i is a set of student-level controls, θ_j are exam-pair fixed effects, and α_t are year fixed effects.⁶

⁶ We define exam-pair fixed effects to be two exams that were given in a particular order. For example, a student who took United States History followed by Calculus will have an exam-pair fixed effect associated with United States History and Calculus. In some cases, two exams will be reversed in order (a student may take Calculus followed by United States History in a subsequent year). We define this to be a new exam-pair fixed effect. Defining exam-pair fixed effects in this way creates some additional exam pairs, but is useful when testing the impact of days between exams on the first exam only or on the second exam only (see the Results section).

Our key assumption is that the difference in time between exams is exogenous after controlling for students taking the same two exam pairs. We argue that this variation, which results from the College Board changing the exam dates from year to year, creates a credible natural experiment. It is not entirely clear why the College Board tinkers with the exam dates so much from year to year. However, discussions with the College Board suggest that the changes are unlikely to be related to student characteristics. For example, a representative of the Board in a private correspondence indicated that a major reason that exam dates change from year to year is the introduction of new exams and the elimination of exams that are no longer offered. These additions and eliminations can lead to a reshuffling of the exam schedule—which, fortuitously, leads to a primary source of the identifying variation used in our analysis.

Even if, as it appears, the changes made by the College Board are unrelated to student characteristics, one potential concern is that students react to these changes and select in or out of taking the exams. For example, our findings could be biased if the better (or worse) students in any given year decide to not take two exams because they are scheduled very close together. Because students must sign up and pay to take an exam long before the test date, it is unlikely that students and parents base their decisions on the test schedule. We can provide two pieces of empirical evidence that suggest that this type of selection is not driving the results that we find. First, we demonstrate in the Results section that time between exams has a large and significant effect on the second exam a student takes but not on the first. Although simple self-selection (better or worse students choose not to take two exams when in close proximity) could explain why students score better or worse on both exams, selection effects alone are unable to explain a systematic difference in performance on the first exam relative to the second exam.

Second, we can directly test whether more students sign up to take exams that have greater spacing. To do this, we aggregate the data to the exam-pair*year level and regress the number of

students taking a given exam pair on the difference in days between the two exams while controlling for exam-pair fixed effects. We find no evidence that days between exams impacts the number of students taking the two exams.⁷

It is also worth considering whether the manner in which AP exams are graded could bias the results. Specifically, one might worry that these exams are graded on a “curve”⁸ and that if everyone does very poorly one year, the strictness of the grading is simply changed to ensure that a relatively stable percentage of people pass the exam. We argue that grade curving would work against finding any meaningful results because if two exams are close in time, causing everyone to do worse on those two exams, the curve would eliminate the ability to find any differences. However, we do not think this is a major concern. The vast majority of students who take an AP exam only take one exam. Thus the “curve” for any one exam will largely be set by the many individuals who take just one exam. Our identification strategy allows us to essentially compare students who took two exams close together one year and further apart a different year to students who took just one exam both years.

III. Results

Table 3 produces the first set of results based on the model specified in Equation (1). The first three columns of Table 3 use the combined total of the first and second exam scores as the dependent variable. Because scores on any individual exam range from 1 to 5, the dependent variable for these three columns ranges from 2 to 10. Column 1 provides the most basic results from the regression of the combined total exam score on days between exams while controlling for exam-

⁷ Regressing the log number of exam takers on days between exams while including year and exam-group dummies gives a coefficient of -0.0018 and a robust standard error of 0.0062.

⁸ The College Board claims not to base its scores off of a strict curve. Rather, their stated goal is to match exam performance with how a college Freshman would have done in the typical corresponding college course. Thus, an AP score of 5 corresponds to an A, a 4 to a B, and so forth. Indeed, score distributions differ considerably across different subjects in ways that seem inconsistent with the College Board adhering to a simple “curve.”

pair fixed effects and year fixed effects. We find a positive and statistically significant relationship between exam scores and days between exams. The coefficient suggests that having 1 more day between exams leads to a higher combined exam score of 0.016. Thus, increasing the number of days between exams from 1 to 10 increases the combined point total on the two exams by approximately 0.144 points (0.07 standard deviations).

The specification in Column 1 requires a linear relationship between days between exams and total exam scores. It is possible, however, that while a student can benefit greatly from having at least a couple of days between exams, more days beyond that are not needed. We test for this kind of nonlinear relationship in Column 2 by including dummy variables for the number of days between exams. The omitted category for this regression is that the exams were taken on the same day (very rare) or 1 day apart. The results suggest that students who take exams that are 2-5 days apart score 0.03 - 0.04 points higher than students with 0-1 days between exams. Exams taken 6-7 days apart yield a 0.08 increase relative to the omitted category and exams taken 8-11 days apart yield a 0.12 increase relative to the omitted category. Thus, these results provide support for an approximately linear relationship between days between exams and exam score outcomes. We further explore the nonparametric relationship that might exist between days between exams and combined exam scores by estimating a model with a dummy variable for every possible number of days between exams (0-11). In Figure 2, we plot the predicted margins and standard error bars for the combined exam score by the number of days between exams. We also plot the linear relationship estimated in Column 1 of Table 3. Once again, a linear relationship appears to concisely describe the function relating days between exams and exam scores.

In Column 3 of Table 3, we once again estimate the linear impact of days between exams on the combined exam score, but also control for the demographic variables (gender, class, and race) in our

dataset. The estimated effect of days between exams on performance shrinks a bit when these controls are added, but it still remains large and statistically significant.

Although the combined exam score is a good indication of performance on the two exams, perhaps a more important outcome to students themselves is whether they pass the exams. Typically, a student must achieve a score of 3 on an exam in order to pass and receive college credit. In Columns 4-6 of Table 3, we use as our dependent variable the number of exams passed, which can be equal to 0, 1, or 2. We see a pattern of results that is very similar to what was found in Columns 1-3. Overall we find that one more day between exams leads to a 0.006 - 0.008 increase in the number of exams passed.

In Table 4, we estimate the same specifications as those found in Table 3, but do so using an ordered probit model, which is more natural than OLS given the non-continuous nature of the dependent variables.⁹ Qualitatively, we find very similar results to those found in Table 3 using OLS. The coefficients themselves from the ordered probit model are harder to interpret. As is often suggested (e.g. Wooldridge (2002, p. 506)), one way to interpret the results is by focusing on the predicted values that the coefficients generate. For example, using the specification in Column 4 of Table 4, we can calculate the probabilities associated with passing none, one, or both exams by the number of days between exams. We report these probabilities in Table 5. The probability of passing both exams when there is one day separating the two exams is 49.0%. With 10 days separating the two exams, the probability increases to 52.9%. Thus, the probability of passing both exams increases by approximately 8% (3.9 percentage points) when given 10 days between exams relative to just 1 day.

In Table 6, we explore whether the effect that we find varies by demographic group. To do this, we include interaction effects of days between exams and gender (Column 1), class (Column 2),

⁹ In Appendix Table 1, we also present results from a Bivariate Probit Model. Once again, we find results that are consistent with those found using OLS.

and race (Column 3). Column 1 indicates that females benefit significantly more than males when the exams are farther apart. In fact, the effect size for females is roughly three times larger than that of males (males = .008; females = .008 + .014). Column 2 finds no statistically significant differences across class. The coefficient on the interaction between sophomores and days between exams appears to be large but it is imprecise due to the small number of sophomore-year exam takers in our dataset. The final column explores the interaction of our main effect with race. Relative to the omitted category (White students), Black and Hispanic students benefit less from a greater number of days between exams. Asian students benefit even more than White students from a greater number of days between exams.

The reason behind these heterogeneous effects is unclear and our data do not allow us to distinguish among various explanations. One possible scenario is that certain groups of students simply don't recover their mental acuity as fast or they get stressed or "burned out" quicker than other groups. Another, perhaps more plausible, mechanism is that the students with larger effect sizes (e.g. females, Asians) study more for AP exams than their counterparts and thus have a higher value of having extra days between exams so as to have more time to "cram" for the second exam.

A natural question regarding our findings regards the extent to which having extra days between exams increases scores on the first and the second exam. An obvious hypothesis would be that students who take two exams close in time perform particularly poorly on the second exam due to mental and physical fatigue. However, it is also possible that students score poorly on the *first* exam taken because they are worried about second exam, spend time cramming for it before they take the first exam, and/or were using up their working memory by retaining information about the second subject while taking the first exam.

We directly test this question in Table 7 by running our baseline specification (Column 1 of Table 3) using the score received on the first exam as the dependent variable in Column 1 and the

score received on the second exam as the dependent variable in Column 2. The results show that having more time between exams has no significant effect on the first exam but rather has a very large and significant effect on the second exam. In fact, these results suggest that almost the entire effect that we found in Tables 3 is driven by poor performance on the second exam when the exams are close together. In Column 3 of Table 7, we use as the dependent variable the ratio of the first exam score to the second exam score. This specification is particularly useful in that it controls for any differences in overall student ability that might exist. A change in the ratio of exam scores cannot be explained by the overall quality of the student. Not surprisingly, given the results in the first 2 columns, we find that the Exam 1 to Exam 2 ratio decreases significantly as the number of days between exams increases. These results lend credence to the mechanism that we suspect helps to explain the results: that the study and attention required to perform two cognitive tasks that close in time result in mental and physical fatigue. The fatigued state is most noticeable towards the end of the cognitive tasks.

IV. Conclusions

Understanding how students and employees handle multiple tasks that occur in close temporal proximity is an important, albeit difficult issue to address empirically. In this paper, we use data generated by the AP exam system to analyze how between-exam delays affect student performance. We find that students who took two exams did significantly better on the second exam the more days they had between the two exams. We identify an approximately linear relationship between 1 and 10 days between exams.

Our results have many direct implications. Testing agencies, students, teachers, and parents can all benefit from understanding the impact that exam proximity can have on performance. Importantly, these results may also have broader-reaching implications such as the impact of

temporally-proximate, mentally-demanding tasks in the workforce. Although the generalizability of our findings to situations outside of test taking is not entirely clear (test taking is a very specific situation that often requires memorization, “cramming,” performance under pressure, etc.), this paper provides clean evidence that suggests that cognitive fatigue may be an important factor in other related domains. Future research can hopefully build on these results by exploring other sources of quasi-random variation in task assignment on performance in a variety of situations.

References:

- Ackerman, P. and R. Kanfer. 2009. "Test Length and Cognitive Fatigue: An Empirical Examination of Effects on Performance and Test-Taker Reactions." *Journal of Experimental Psychology: Applied*, 15(2), 163-181.
- Ariely, D. and Wertenbroch, K. 2002. "Procrastination, Deadlines, and Performance: Self-Control by Precommitment." *Psychological Science*, 13(3): 219-224.
- Cepeda, N.J., Pashler, H., Vul, E., Wixted, J.T., and Rohrer, D. 2006. "Distributed practice in verbal recall tasks: A review and quantitative synthesis." *Psychological Bulletin*, 132, 354-380.
- C. Chambers, T.D. Noakes, E.V. Lambert, & M.I. Lambert (1998): Time course of recovery of vertical jump height and heart rate versus running speed after a 90-km foot race. *Journal of Sports Sciences*, 16:7, 645-651.
- Coviello, D., Ichino, A., and Persico, N. 2010. "Don't spread yourself too thin: the impact of task juggling on workers' speed of job completion." NBER Working Paper #16502.
- DellaVigna, S. 2009. "Psychology and Economics: Evidence from the field." *Journal of Economic Literature*, 47, 315-372.
- Ebbinghaus, H. 1896-1897. (O. Wilhem, trans.) "On a new method for testing mental abilities and its use with school children." *Zeitschrift fur Psychologie und Physiologie der Sinnesorgane*, 13, 401-459.
- Paas, F., Renkl, A., & J. Sweller. 2004. "Cognitive Load Theory: Instructional Implications of the Interaction between Information Structures and Cognitive Architecture." *Instructional Science*, 32, 1-8.
- Woolridge, J. M. 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press.

Figure 1. Exam Calendar Example. This figure provides an example of an AP exam calendar (2012).
 Source: Collegeboard (<http://www.collegeboard.com/student/testing/ap/cal/cal2.html>).

2012 Exam Calendar - Week 1		
	Morning - 8 a.m.	Afternoon - 12 p.m.
Monday, May 7	Chemistry Environmental Science	Psychology
Tuesday, May 8	Computer Science A Spanish Language	Art History
Wednesday, May 9	Calculus AB Calculus BC	Chinese Language and Culture
Thursday, May 10	English Literature and Composition	Japanese Language and Culture Latin: Vergil
Friday, May 11	German Language and Culture U.S. History	European History
Studio Art: Last day for your school to submit digital portfolios and to gather 2-D Design and Drawing students for the physical portfolio assembly. Students should have forwarded their completed digital portfolios to their teachers well before this date.		

2012 Exam Calendar - Week 2			
	Morning - 8 a.m.	Afternoon - 12 p.m.	Afternoon - 2 p.m.
Monday, May 14	Biology Music Theory	Physics B Physics C: Mechanics	Physics C: Electricity and Magnetism†
Tuesday, May 15	U.S. Government and Politics	Comparative Government and Politics French Language and Culture	
Wednesday, May 16	English Language and Composition	Statistics	
Thursday, May 17	Macroeconomics World History	Microeconomics Italian Language and Culture	
Friday, May 18	Human Geography Spanish Literature		

Figure 2. Predicted Values by Days Between Exams. Each dot in this figure is the predicted combined exam score for the two AP exams when the number of days between the two exams varied from 0 to 11. These predicted values were obtained from estimating a regression similar to Column 1 of Table 3, but with a separate dummy variable for each possible day between exam, and then obtaining predicted values treating all factor variables as balanced (default margins command in Stata). Standard error bars are also included. The linear line is the predicted values obtained from the regression in Column 1 of Table 3.

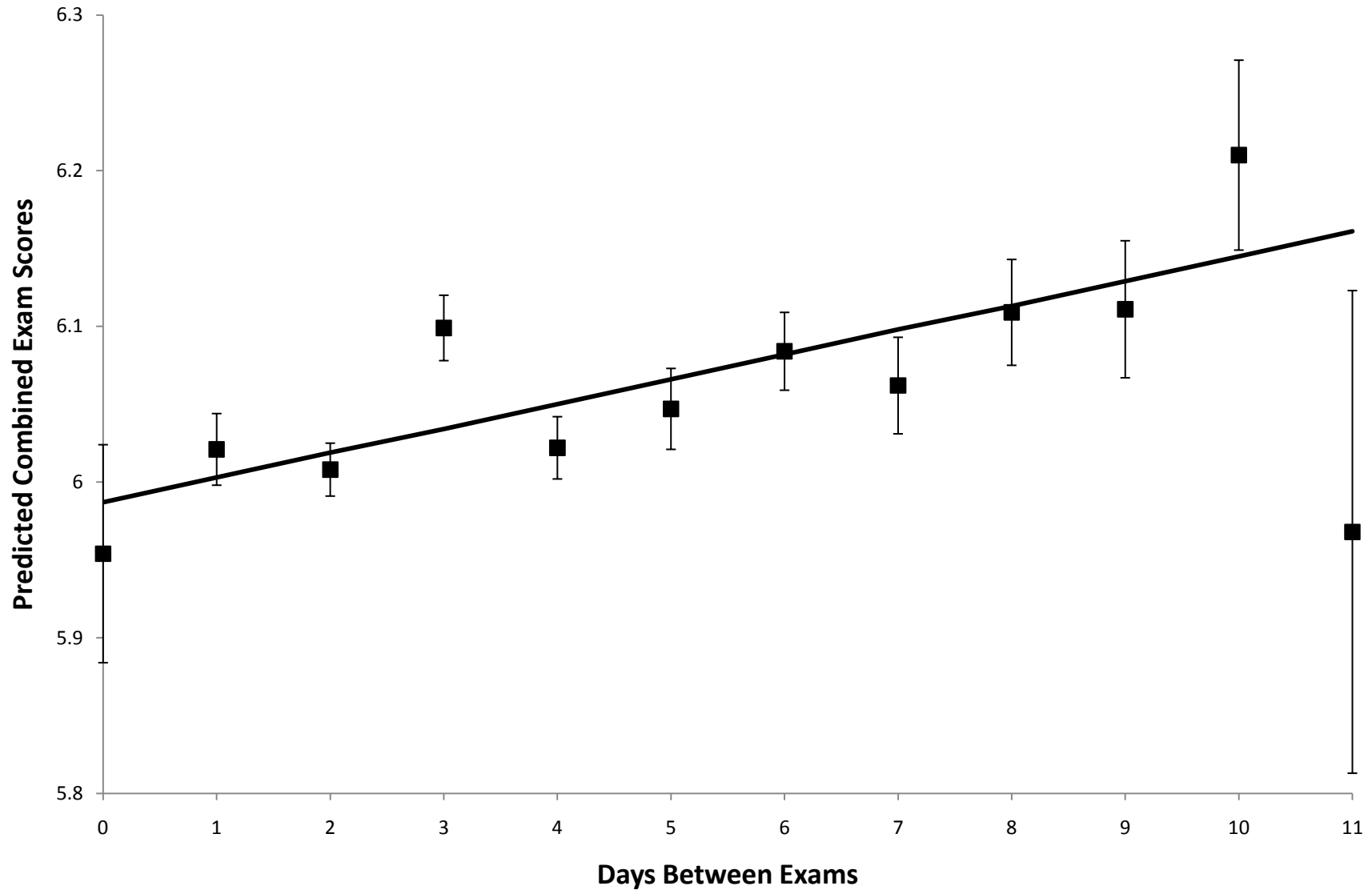


Table 1. AP Exams: Ordered By Popularity

Exam	# of Students	% of Students	Avg Score
US History	45,684	38.37%	2.98
English Literature	32,941	27.67%	3.04
English Language	28,761	24.15%	3.02
Calculus AB	25,914	21.76%	3.01
Biology	19,685	16.53%	3.19
Chemistry	13,207	11.09%	2.88
Gov't & Politics: US	10,443	8.77%	2.72
Spanish Language	9,176	7.71%	3.41
European History	8,340	7.00%	3.21
Physics B	6,474	5.44%	2.83
Psychology	5,479	4.60%	3.24
Calculus BC	4,913	4.13%	3.52
Statistics	3,990	3.35%	2.78
Economics: Macro	2,831	2.38%	2.72
French Language	2,657	2.23%	2.77
Computer Science A	2,226	1.87%	2.79
Economics: Micro	2,176	1.83%	2.71
Art History	1,849	1.55%	3.14
Environmental Science	1,575	1.32%	2.78
Spanish Literature	1,557	1.31%	3.20
Physics C: Mechanics	1,417	1.19%	2.80
Gov't & Politics: Comparative	1,246	1.05%	2.63
Computer Science B	1,173	0.99%	3.34
Art Studio: General	979	0.82%	3.18
Latin: Vergil	877	0.74%	3.01
Music Theory	760	0.64%	3.38
German Language	589	0.49%	3.12
Art Studio: Drawing	454	0.38%	3.37
Latin: Literature	445	0.37%	2.68
French Literature	286	0.24%	3.22
Physics C: Electricity & Magnetism	34	0.03%	3.35
Total	238,138	200.00%	3.02

Notes: All averages and frequencies are for our sub-sample of students taking exactly two exams, which is based on a 10% random sample of all AP exam takers. Not all exams were offered in every year.

Table 2. Summary Statistics

	Number	Percent	Avg Score
Class			
Senior	95,637	80.3%	2.97
Junior	22,514	18.9%	3.23
Sophomore	918	0.8%	3.51
Gender			
Female	65,539	55.0%	2.95
Male	53,530	45.0%	3.11
Race			
White	79,055	66.4%	3.09
Asian	15,739	13.2%	3.04
Hispanic	9,046	7.6%	2.75
Black	4,652	3.9%	2.24
Other	10,577	8.9%	3.12
Days Between Exams			
0	4,239	3.6%	2.88
1	16,660	14.0%	3.03
2	21,800	18.3%	2.97
3	16,557	13.9%	3.05
4	16,192	13.6%	3.06
5	12,385	10.4%	3.12
6	11,975	10.1%	3.04
7	7,816	6.6%	2.98
8	6,199	5.2%	2.94
9	3,703	3.1%	2.99
10	1,308	1.1%	3.16
11	235	0.2%	3.16

Notes: All statistics are for our sub-sample of students taking exactly two exams, which is based on a 10% random sample of all AP exam takers. Days between exams indicates the number of full days between the two exams taken (e.g. 0 indicates that the two exams were taken on the same day).

Table 3. The Effect of Days Between Exams on Exam Outcomes - OLS

	First Exam Score + Second Exam Score			Number of Exams Passed		
Days Between Exams	0.016 (0.004)**		0.012 (0.004)**	0.008 (0.002)**		0.006 (0.002)**
2-3 Days Between		0.040 (0.024)			0.020 (0.009)*	
4-5 Days Between		0.032 (0.030)			0.019 (0.011)	
6-7 Days Between		0.077 (0.033)*			0.043 (0.013)**	
8-11 Days Between		0.120 (0.038)**			0.058 (0.015)**	
Male			0.236 (0.012)**			0.075 (0.005)**
Sophomore			0.764 (0.069)**			0.214 (0.023)**
Junior			0.462 (0.016)**			0.153 (0.006)**
Hispanic			-0.864 (0.024)**			-0.378 (0.009)**
Black			-1.584 (0.029)**			-0.601 (0.012)**
Asian			-0.204 (0.018)**			-0.090 (0.007)**
Other Race			0.017 (0.021)			-0.009 (0.008)
Exam Group Fixed Effects	X	X	X	X	X	X
Year Fixed Effects	X	X	X	X	X	X
R-Squared	0.054	0.054	0.095	0.037	0.037	0.078
Observations	119,069	119,069	119,069	119,069	119,069	119,069

Notes: This table presents coefficients and robust standard errors from the OLS regression of total exam score on both exams (first exam score + second exam score) in Columns 1 - 3 and the number of exams passed (a score of 3 or higher) in Columns 4 - 6 on the days between exams and in some specifications demographic characteristics. Each regression includes fixed effects for the two exams being taken by each student and year fixed effects.

* $p < .05$; ** $p < .01$;

Table 4. The Effect of Days Between Exams on Exam Outcomes - Ordered Probit

	First Exam Score + Second Exam Score			Number of Exams Passed		
Days Between Exams	0.008		0.006	0.011		0.00900
	(0.002)**		(0.002)*	(0.003)**		(0.003)**
2-3 Days Between		0.018			0.032	
		(0.012)			(0.014)*	
4-5 Days Between		0.013			0.030	
		(0.015)			(0.017)	
6-7 Days Between		0.035			0.064	
		(0.017)*			(0.019)**	
8-11 Days Between		0.056			0.087	
		(0.019)**			(0.022)**	
Male			0.122			0.116
			(0.006)**			(0.007)**
Sophomore			0.402			0.373
			(0.037)**			(0.044)**
Junior			0.239			0.239
			(0.008)**			(0.010)**
Hispanic			-0.462			-0.549
			(0.013)**			(0.014)**
Black			-0.853			-0.867
			(0.016)**			(0.018)**
Asian			-0.108			-0.136
			(0.01)**			(0.011)**
Other Race			0.006			-0.010
			(0.011)			(0.012)
Exam Group Fixed Effects	X	X	X	X	X	X
Year Fixed Effects	X	X	X	X	X	X
Pseudo R-Squared	0.013	0.013	0.024	0.019	0.019	0.038
Observations	119,004	119,004	119,004	119,004	119,004	119,004

Notes: This table presents coefficients and robust standard errors from the ordered probit regression of total exam score on both exams (first exam score + second exam score) in Columns 1 - 3 and the number of exams passed (a score of 3 or higher) in Columns 4 - 6 on the days between exams and in some specifications demographic characteristics. Each regression includes fixed effects for the two exams being taken by each student and year fixed effects.

* $p < .05$; ** $p < .01$;

Table 5. Predicted Probability of Passing Exams Using Ordered Probit Estimates

	Number of Days Between Exams												
	0	1	2	3	4	5	6	7	8	9	10	11	
Predicted Number of Exams Passed													
None	22.1%	21.8%	21.5%	21.2%	20.8%	20.5%	20.2%	19.9%	19.6%	19.3%	19.0%	18.7%	
One	29.3%	29.2%	29.1%	29.0%	28.8%	28.7%	28.6%	28.5%	28.3%	28.2%	28.1%	27.9%	
Two	48.6%	49.0%	49.4%	49.9%	50.3%	50.8%	51.2%	51.6%	52.1%	52.5%	52.9%	53.4%	

Notes: This table provides the predicted probabilities for the exact number of exams passed (0, 1, 2) by days between exams using the ordered probit regression specification found in Column 4 of Table 4.

Table 6. OLS Estimates of the Effect of Days Between Exams on Total Score: By Subgroup

	Gender	Grade	Race
Days Between Exams	0.008 (0.005)	0.011 -0.005	0.017 (0.005)**
Female * Days Between Exams	0.014 (0.005)**		
Soph * Days Between Exams		0.017 (0.032)	
Junior * Days Between Exams		0.002 (0.008)	
Asian * Days Between Exams			0.023 (0.008)*
Black * Days Between Exams			-0.014 (0.012)
Hispanic * Days Between Exams			-0.038 (0.009)**
Other * Days Between Exams			-0.009 (0.009)
Exam Pair Fixed Effects	X	X	X
Year Dummies	X	X	X
Observations	119069	119069	119069
R-squared	0.0581	0.0614	0.0858
F-stat (for interactions)	7.94*	0.16	7.96**

Notes: This table presents coefficients and robust standard errors from the OLS regression of total exam score on both exams (first exam score + second exam score) on the days between exams and days between exams interacted with gender, class, and race demographics. The omitted categories are male, senior, and white. Each regression includes fixed effects for the two exams being taken by each student and year fixed effects.

* p < .05; ** p < .01;

Table 7. The Effect of Days Between Exams on the Scores of the First and Second Exam Taken

	Exam 1 Score	Exam 2 Score	Ratio of Exam 1 and Exam 2 Scores
Days Between Exams	0.003 (0.002)	0.013 (0.003)**	-0.007 (0.001)**
Exam Group Fixed Effects	X	X	X
Year Fixed Effects	X	X	X
R-Squared	0.060	0.064	0.071
Observations	119,069	119,069	119,069

Notes: This table presents coefficients and robust standard errors from the OLS regression of the score received on the first exam taken (Column 1), the second exam taken (Column 2), and the score ratio of exam 1 and exam 2 on the days between exams. Each regression includes fixed effects for the two exams being taken by each student and year fixed effects.

* p < .05; ** p < .01;

**Appendix Table 1. The Effect of Days Between Exams on Exam Outcomes -
Bivariate Probit**

		Fail Second Exam	Pass Second Exam
1 Day Between Exams:	Fail First Exam	21.67%	10.78%
	Pass First Exam	18.69%	48.86%
		Fail Second Exam	Pass Second Exam
10 Day Between Exams:	Fail First Exam	19.14%	12.47%
	Pass First Exam	15.33%	53.06%

Notes: This table presents predicted probabilities from a bivariate probit model. The numbers in the boxes are the predicted probabilities for passing some combination of the first and second AP exams taken when there is one day between exams (first box) or 10 days between exams (second box).