Blueprint Labs

Discussion Paper #2016.02

The Impact of Computer Usage on Academic Performance: Evidence from a Randomized Trial at the United States Military Academy

Susan Payne Carter
Kyle Greenberg
Michael Walker

May 2016

MIT Department of Economics
77 Massachusetts Avenue, Bldg. E53-390
Cambridge, MA 02139

National Bureau of Economic Research
1050 Massachusetts Avenue, 3rd Floor
Cambridge, MA 02138

The Impact of Computer Usage on Academic Performance: Evidence from a Randomized Trial at the United States Military Academy
Susan Payne Carter, Kyle Greenberg, and Michael Walker
Blueprint Labs Discussion Paper #2016.02
May 2016

ABSTRACT

We present findings from a study that prohibited computer devices in randomly selected classrooms of an introductory economics course at the United States Military Academy. Average final exam scores among students assigned to classrooms that allowed computers were 18 percent of a standard deviation lower than exam scores of students in classrooms that prohibited computers. Through the use of two separate treatment arms, we uncover evidence that this negative effect occurs in classrooms where laptops and tablets are permitted without restriction and in classrooms where students are only permitted to use tablets that must remain flat on the desk surface.

* The views expressed herein are those of the authors and do not reflect the position of the United States Military Academy, the Department of the Army, or the Department of Defense. We are grateful for the invaluable contributions to this study from Perry Bolding, Bill Skimmyhorn, Rekha Balu, Cassandra Hart, and the USMA economics program, especially those instructors participating in our study.

THE IMPACT OF COMPUTER USAGE ON ACADEMIC PERFORMANCE: EVIDENCE FROM A RANDOMIZED TRIAL AT THE UNITED STATES MILITARY ACADEMY*

By

Susan Payne Carter
United States Military Academy, West Point

Kyle Greenberg
United States Military Academy, West Point

Michael S. Walker
United States Military Academy, West Point

May 2016

Abstract

We present findings from a study that prohibited computer devices in randomly selected classrooms of an introductory economics course at the United States Military Academy. Average final exam scores among students assigned to classrooms that allowed computers were 18 percent of a standard deviation lower than exam scores of students in classrooms that prohibited computers. Through the use of two separate treatment arms, we uncover evidence that this negative effect occurs in classrooms where laptops and tablets are permitted without restriction and in classrooms where students are only permitted to use tablets that must remain flat on the desk surface.

* The views expressed herein are those of the authors and do not reflect the position of the United States Military Academy, the Department of the Army, or the Department of Defense. We are grateful for the invaluable contributions to this study from Perry Bolding, Bill Skimmyhorn, and the USMA economics program, especially those instructors participating in our study.

I. INTRODUCTION

Internet-enabled classroom technology is nearly universal at all levels of education in the United States. Between 1994 and 2005, the percentage of U.S. public school classrooms with Internet access increased from 3 percent to 94 percent, while the ratio of students to computers with Internet access in these classrooms decreased from 12.1 to 3.8 (Wells & Lewis, 2006). Further improvement of classroom Internet access continues to serve as a major policy initiative for the U.S. government. In 2013, President Obama introduced the ConnectED initiative, which included a goal of providing "next generation" broadband Internet access to 99 percent of U.S. students by 2018 through classrooms and libraries.[1] More recently, the U.S. Department of Education emphasized its policy commitment to Internet-enabled pedagogical reform in the 2016 National Education Technology Plan.[2]

At the college level, campus Internet access has become a competitive margin as schools battle to attract the best students. Students have become accustomed to near-constant Internet access at home and in the classroom. As a result, reduced bandwidth and/or Internet "dead zones" may negatively impact student perceptions of the quality of a university's education. College rating services, noting these student preferences, rank institutions according to their wireless connectivity, and undergraduate institutions market the ease of student access to the Internet as a recruiting tool.[3] Beyond satisfying student preferences, increased connectivity also provides opportunities for students and teachers to collaborate outside of the classroom, convenient options

[1] See https://www.whitehouse.gov/issues/education/k-12/connected for a full explanation of the ConnectED initiative and its components.

[2] See *2016 National Education Technology Plan*, page 6.

[3] UNIGO ranked the "Top 10 Wired Schools on the Cutting Edge of Technology" in 2013, relying upon WiFi coverage, student access to computers, and required computer science courses (among other factors) as evidence of a school's commitment to technology.

for student research via university library-enabled online search engines, and continuous access to web-based curricula, to name a few of the potential benefits touted by technology proponents.

In addition to other Internet-enabled classroom innovations, the development of electronic textbooks has accompanied the proliferation of web-based curriculum and wireless access at undergraduate institutions. “Enhanced” textbooks offer students the capability to watch embedded videos, follow hyperlinks to pertinent articles on the Internet, and carry their entire curriculum with them at all times.[4] These e-textbooks also provide publishers with an ability to avoid competition with their own secondary market, reduce marginal publication costs, and easily update content. “E-texts” undoubtedly offer new and desirable features, which would be impossible to achieve with the standard text.

The platforms required for use of the e-texts (e.g., laptop and tablet computers) also provide students with access to a host of potential distractions if allowed in the classroom. As institutions, including the one in the present study,[5] continue to push for ever faster and continuous access to wireless Internet to support the proliferation of web-enabled educational resources, it is unclear whether the benefits of Internet-enabled computer usage in the classroom outweigh its potential costs to student learning. In fact, anecdotal evidence suggests that professors and teachers are increasingly banning laptop computers, smart phones, and tablets from their classrooms.[6]

In an effort to inform the debate surrounding student Internet access in the classroom, we evaluate the effects of an experiment that randomly allowed student access to laptop and tablet

---

[4] There are other advantages for the professor as well. For example, certain “e-text” programs enable professors to capture the rate at which students progress through reading assignments and, thus, to confirm whether students have completed these assignments prior to class.

[5] See, for example, “By the Numbers,” *West Point Magazine*, Summer 2015, p. 46.

[6] See, for example, Gross (2014), “This year, I resolve to ban laptops from my classroom,” *Washington Post*, available from https://www.washingtonpost.com.

computers during an introductory economics course at the United States Military Academy at West Point, NY. We divided classrooms into a control group or one of two treatment groups. Classrooms in the first treatment group permitted students to use laptops and tablets without restriction. In the second treatment group, hereafter referred to as the “modified-tablet” treatment group, students were only permitted to use tablets, but the tablet had to remain flat on the desk surface. Meanwhile, students assigned to classrooms in the control group were not permitted to use laptops or tablets in any fashion during class.

The results of our study suggest that permitting computing devices in the classroom reduces final exam scores by 18 percent of a standard deviation. By way of comparison, this effect is as large as the average difference in exam scores for two students whose cumulative GPAs at the start of the semester differ by one-third of a standard deviation. These results are nearly identical for classrooms that permit laptops and tablets without restriction as they are for classrooms that only permit modified-tablet usage. This result is particularly surprising considering that nearly 80 percent of students in the first treatment group used a laptop or tablet at some point during the semester while only 40 percent of students in the second treatment group ever used a tablet. We also find modest evidence that computer usage is most detrimental to male students and to students who entered the course with a high grade point average (GPA).

This study adds to the existing literature concerning the effects of classroom technology usage on student performance. Our research moves beyond the measurement of student attitudes toward computer usage in the classroom (e.g., Barak, et al., 2006) and observational studies of correlation between technology and cohort performance (e.g., Wurst, et al., 2008). Instead, we attempt to isolate the causal effect of Internet-enabled computer usage on individual student performance during a semester-long undergraduate course. Our randomized controlled trial is most similar to previous laboratory-style studies, many of which demonstrate the potentially negative effects of computer usage on student outcomes (e.g., Hembrooke and Gay, 2003; Sana, et al., 2013; Mueller and Oppenheimer, 2014). In contrast to the laboratory-style research, however, our study measures the cumulative effects of Internet-enabled classroom technology over the course of a semester, as opposed to its impact on immediate or short-term (less than one week) recall of knowledge. Furthermore, our research design intentionally seeks to limit the influence of artificial behaviors caused by experimental conditions or treatment design. This outcome might occur when the experimental design requires students to perform tasks or behave in a way that is abnormal or out of character, such as forcing students to multi-task, as in Sana, et al. (2013), or requiring students to use computers, as in Mueller and Oppenheimer, (2014).[7]

While laboratory experiments certainly allow the researcher to limit the potential channel through which computers can affect learning, students may behave differently when the outcome of interest is performance on an inconsequential or random topic than when faced with an assessment that may impact their GPA. Thus, investigation of the effects of technology in the context of an actual course is an important extension of laboratory research. Our study also, therefore, adds to existing research that has attempted to measure the effect of computer usage in an actual classroom environment (e.g., Grace-Martin and Gay, 2001; Fried, 2008; and Kraushaar and Novak, 2010). This research tends to show a negative correlation between Internet-enabled computer usage and student performance on course-specific events.

Our RCT design allows us to improve upon existing results. First, we are able to control for selection into computer usage and avoid the problems associated with student self-reporting of

[7] In Sana, et al.(2013), the authors experimental design required students in a treatment group to complete a pre-determined list of twelve web-enabled tasks during a classroom lecture. These tasks primarily required the student "multi-tasker" to answer questions irrelevant to the lecture material.

computer activity. Second, our comprehensive dataset allows us to control for a wide range of relevant observable characteristics, which has been an insurmountable issue for many of the aforementioned researchers. Finally, we examine the effect on final exam scores where students are incentivized to do well both for their GPA and for their class rank which affects their future job choice. Although many aspects of West Point differ from typical 4-year undergraduate institutions, there are many reasons to believe that permitting computers in traditional lecture-style classrooms could have even more harmful effects than those found in this study. Students at West Point are highly incentivized to earn high marks, professors are expected to interact with their students during every lesson, and class sizes are small enough that it is difficult for students to be completely distracted by their computer without the professor noticing.

The paper proceeds as follows. Section II provides background on West Point for the purposes of generalization and Section III discusses our experimental design. Sections IV and V discuss our empirical framework, data sample, and evidence of successful random assignment. Section VI presents the results of our regression analysis, Section VII discusses results from additional robustness checks, and Section VIII concludes.

II. BACKGROUND ON WEST POINT

The United States Military Academy at West Point, NY, is a 4-year undergraduate institution with an enrollment of approximately 4,400 students. In addition to a mandatory sequence of engineering courses, students complete a liberal arts education with required courses in math, history, English, philosophy, and most importantly for this paper, introductory economics. This principles-level economics course, which combines micro and macroeconomics in a single semester, is typically taken during a student's sophomore year.

West Point's student composition is unique, due primarily to its mission of generating military officers and the unique requirements of its admissions process. Admission to West Point is accompanied by the equivalent of a "full-ride" scholarship, but when a student graduates, he/she is commissioned as an officer in the U.S. Army and incurs an 8 year service obligation with a 5-year active duty requirement. In preparation for this service obligation, West Point requires all students to be physically active through competitive sports (intramurals, club, or varsity) and to complete required military education courses, in addition to a rigorous academic course load. These requirements likely lead to a student body that is more athletic and physically fit, on average, than at typical universities. Furthermore, to gain admission to West Point, applicants must receive a nomination from one of their home state's Congressional members on top of the typical elements of a college admissions file (e.g., standardized test scores, letters of recommendation, etc.).[8] Due to this admissions requirement and limits placed on the number of students a Congressperson can have at West Point at any given time, students are more geographically diverse than students at a typical undergraduate institution.

To alleviate concerns regarding the generalizability of our findings, we report summary statistics comparing students to other schools in Table 1. West Point is currently ranked 22nd on U.S. News and World Report's list of National Liberal Arts Colleges.[9] In Panel A, we show gender, race, and home location breakdowns for West Point relative to five other schools ranked in the top 25 of the same poll. West Point is about twice the size of other similar schools but has a similar student to faculty ratio. West Point has a much lower female to male ratio with female

[8] Applicants may also receive a nomination from the U.S. Vice President and/or the Secretary of the Army. In addition to these political nominations, applicants may receive a nomination from categories related to the student's own prior military service or a parent's military service.

[9] See http://colleges.usnews.rankingsandreviews.com/best-colleges, accessed 29 April 2016, for the full set of rankings.

students accounting for only 17 percent of the undergrad population. It also has a much lower percentage of non-resident aliens and a slightly higher percentage of people from out of state, both direct impacts of West Point's unique admissions process. On the other hand, ACT and SAT scores at West Point are comparable to scores at other high-ranked liberal arts colleges, as is the share of minority students. In Panel B, we compare West Point to all 4-year public schools, 4-year public schools with a student body between 1,000 and 10,000, all 4-year schools (including private non-profit and private for-profit), and all 4-year schools with a population between 1,000 and 10,000. West Point's study body consists of fewer women, has fewer minorities, and has slightly higher ACT and SAT scores than the average 4-year institution. Overall, while there are clear differences between the U.S. Military Academy and other civilian institutions, West Point does have many similarities with liberal arts colleges and smaller 4 year public schools.

III. EXPERIMENTAL DESIGN

To test the impact of allowing Internet-enabled laptops and tablets in classrooms, we randomized classrooms into either a control group or one of two treatment groups. Control group classrooms were "technology-free," indicating that students were not allowed to use laptops or tablets at their desk. In our first treatment group, students were permitted to use laptops and/or tablets during class for the purposes of note-taking and classroom participation (e.g., using the "e-text" version of the course textbook). However, professors had discretion to stop a student from using a computing device if the student was blatantly distracted from the class discussion. This treatment was intended to replicate the status quo collegiate classroom environment: students using Internet-enabled technology at will during lecture and discussion. Classrooms in our second treatment group, or "tablet-only" group, allowed students to use their tablet computers, but professors in this group required tablets to remain flat on the desk (i.e., with the screen facing up and parallel to the desk surface). This modified-tablet usage enabled students to access their tablets to reference their e-text or other class materials, while allowing professors to observe and correct student access to distracting applications. Therefore, the second treatment more closely replicated the "intended" use of Internet-enabled technology in the classroom.

West Point provides an ideal environment for conducting a classroom experiment for a number of reasons. As part of West Point's "core" curriculum, the principles of economics course has a high enrollment (approximately 450 students per semester). Class size, however, remains relatively small due to an institutional commitment to maintaining a low faculty to student ratio, which is generally near 1:15 in the principles course and is capped at 1:18 per class by Academy policy. Despite the large enrollment and small class size, student assessment in the course is highly standardized. All classes use an identical syllabus with the same introductory economics textbook and accompanying online software package. Students complete all homework, midterms, and final exams (consisting of multiple choice, short answer, and essay questions) via an online testing platform. With 30 different sections of the course, taught by approximately ten different professors, most professors teach between two and four sections of the economics course each semester. This course structure allowed us to randomize treatment and control groups among classrooms taught by the same professor. As part of this process, we limited our study to professors who taught at least two sections of the course in a single semester and ensured that each professor taught at least one section in the control group and at least one section in either treatment group.[10]

Second, within a class hour, students are randomized into their particular class. West Point centrally generates student academic schedules, which are rigidly structured due to the substantial

[10] It is important to note that West Point professors do not have teaching assistants.

number of required courses. Students cannot request a specific professor and, importantly, students are unaware prior to the first day of class whether computers will be allowed in their classroom or not. After the first day of class, there is virtually no switching between sections.

Third, West Point's direct link between student performance and post-graduation employment provides motivation for students to do well in the economics course. The higher a student ranks in her graduating class, the greater her chances of receiving her first choice of military occupation and duty location upon graduating. For those students incapable of seeing the long term consequences of poor academic performance, West Point's disciplinary system provides additional, immediate reinforcement. If their professor elects to report the incident, a student who misbehaves in class (whether by arriving late, falling asleep, skipping class, or engaging in distracting behavior) will be disciplined by the officer in charge of her military training.[11] Fourth and finally, all students at West Point are on equal footing in terms of access to the educational resources that may differentially impact our experiment. West Point required all students in our study to purchase laptop computers and tablets, and each academic building at West Point was equipped with wireless Internet access at the time of our experiment. Furthermore, each student is required to complete an introductory computer science course during their freshman year, which falls before the economics course in West Point's core curriculum sequence.

IV. EMPIRICAL FRAMEWORK

To compare outcomes between students assigned to classrooms that permitted laptop or tablet

[11] This "discipline" takes many forms, depending on the severity of the infraction and the student's personal disciplinary background. For example, the officer in charge may elect to employ everything from counseling techniques to monotonous physical tasks (e.g., "walking hours") in correcting unacceptable behavior. Unsurprisingly, these disciplinary measures often take place during the student's valuable weekend hours.

usage and students assigned to classrooms that prohibited computer usage, we estimate the following model of undergraduate academic achievement:

$$(1) \qquad Y_{ijht} = \kappa_{jt} + \lambda_{ht} + \gamma' X_i + \pi Z_{jht} + \eta_{ijht}.$$

Y_{ijht}, is the final exam score of student *i* who had professor *j* during class-hour *h* and semester *t*. Z_{jht} is an indicator for an individual being in a classroom which allows laptops or tablets. X_i is a vector of individual controls, the term κ_{jt} includes fixed effects for each combination of professor and semester, λ_{ht} includes fixed effects for each combination of class-hour and semester, and η_{ijht} is the error term. By including semester and professor controls, we compare students within the same semester while also controlling for unobserved mean differences in academic performance across professors and across class-hours.[12] As laptop and tablets are randomly prohibited in certain classrooms, estimates of π capture the causal effect of allowing computers in the classroom on final exam scores.[13]

We also report OLS and 2SLS estimates of the effects of computer usage on academic performance. In this setting, the causal equation of interest is:

$$(2) \qquad Y_{ijht} = \alpha_{jt} + \beta_{ht} + \delta' X_i + \rho D_{ijht} + \varepsilon_{ijht}.$$

where D_{ijht} is an indicator variable that equals 1 if student *i* uses a laptop or a tablet and equals 0 otherwise. Because students individually choose whether to use a laptop or tablet in the classroom,

[12] Each student only has one observation in the data for the analysis that follows. The within semester comparison is critical for at least two reasons. First, the students participating in the experiment spanned two separate class years at West Point, which may have been subject to different admissions policies and/or admissions personnel. Second, professors in charge of the introductory course and its primary textbook changed between the semesters. Both textbooks were published by the same company and used an identical online assessment platform, but the curricular sequence of the course changed slightly in the second semester to accommodate the layout of the new textbook.

[13] Since the treatment in this experiment varies at the classroom level, it would normally be appropriate to cluster standard errors on classrooms. However, we mainly report robust standard errors in the results that follow because they are more conservative than clustered standard errors. We explore alternative standard error estimates below.

OLS estimates of ρ in equation (2) may be biased by unobservable factors that are correlated with both computer usage and test scores. Under the assumption that assignment to a classroom that allows laptops or tablets only influences academic performance through a student's propensity to use her own computing device, we can use assignment to a classroom that permits computers (Z_{jht} from equation (1)) as an instrumental variable for actual computer usage (D_{ijht}).

A key concern for interpreting 2SLS estimates of ρ as causal is that the exclusion restriction is violated if computer usage by a student's peers provides a strong enough distraction to influence her own performance. The possibilities of such spillovers in West Point classrooms are likely minimized by the small class sizes, class layout, and unique levels of professor-teacher interaction in the classroom. For example, West Point professors typically arrange desks in a "U-shape" within the classroom, reducing the number of students with obstructed views of the teacher and front of the classroom. Additionally, West Point encourages its professors to engage with all students in the classroom over the course of a class hour. Nevertheless, we cannot rule out the possibility of spillovers and therefore urge caution when interpreting 2SLS estimates as causal.[14]

V. DATA, STUDENT CHARACTERISTICS, AND COVARIATE BALANCE

Our sample consists of students enrolled in West Point's Principles of Economics during the spring semester of the 2014-2015 academic year or the fall semester of the 2015-2016 academic year. We limit the sample to students who took the class as sophomores and further exclude

[14] Empirical evidence of a "distraction effect" is mixed. Aguilar-Roca et al. (2012) randomly assign students to classrooms with "laptop-free" seating zones in a large-enrollment biology course. They observe no impact of the seating arrangements on student performance, suggesting that computer usage of other students does not impact academic performance. On the other hand, Fried (2008) finds that 64% of students who reported in-class distractions due to laptop use cited other students' laptop usage as a distractor. Additionally, Sana, et al (2013) find that students able to view peer "multi-tasking" on a laptop scored 17 percentage points lower on an immediate comprehension test than students not in of the multi-tasking behavior. The authors found that this effect was larger (17 percentage points versus 11) than the negative effect of own laptop usage in a separate experiment.

students enrolled in classrooms of professors who chose not to participate in the experiment, resulting in a final sample of 726 students.[15]

Columns 1 through 3 of Table 2 report descriptive statistics for students assigned to the control group, where laptops and tablets are not allowed, treatment group 1, where laptop and tablet computers are allowed without restriction, and treatment group 2, where tablets are permitted if students keep them face up on the desk at all times. As expected, the racial and ethnic composition of students in the sample is similar to that of the West Point student body, with women comprising roughly 1 in 5 students in each group, African Americans and Hispanics comprising roughly 1 in 4 students, and Division I athletes comprising 1 in 3 students. Average composite ACT scores are between 28 and 29, and average baseline (pre-treatment) GPAs are between 2.8 and 2.9 for all three groups.[16]

Subsequent columns of Table 2 investigate the quality of the randomization of classrooms to treatment arms by comparing differences in demographic characteristics, baseline GPAs, and ACT scores between treatment arms and the control group. The numbers reported in column 4 are regression-adjusted differences between students assigned to a classroom in either treatment group and students assigned to a classroom in the control group. The regressions used to construct these estimates only include fixed effects for each combination of professor and semester and fixed effects for each combination of class hour and semester. The differences in column 4 are generally small and statistically insignificant, suggesting that the assignment of classrooms to either treatment group was as good as random. The P-value from a test of the joint hypothesis that all

[15] Nearly 95 percent of students enrolled in Principles of Economics are sophomores. Limiting the sample to sophomores ensures that no student appears in our data twice. Two professors informed the authors of their intention to not participate prior to the randomization of classrooms to treatment arms.

[16] For students who did not take the ACT, we converted SAT scores to ACT scores using the ACT-SAT concordance table found here: http://www.act.org/solutions/college-career-readiness/compare-act-sat/.

differences in baseline characteristics are equal to zero, reported at the bottom of the column, is 0.61, further supporting the argument that classrooms assigned to either treatment group were not meaningfully different from classrooms assigned to the control group.

Columns 5 and 6 of Table 2 report results from the same covariate balance check as column 4, but this time separately comparing differences in baseline characteristics between students in treatment group 1 and the control group and students in treatment group 2 and the control group, respectively. On the whole, there are relatively few significant differences in observable characteristics between groups. Students assigned to classrooms that permitted unrestricted use of laptops and tablets are 7.5 percentage points more likely to be Division I athletes than students assigned to classrooms where computers were prohibited. Although this is likely a chance finding, we control for baseline characteristics in our analysis below to ensure that our estimates are not confounded by this or any other differences.

We derive outcomes in this experiment from a final exam that was mandatory for all students in the course. This exam consisted of a combination of multiple choice, short answer (mostly fill-in-the-blank questions and problems requiring graphical solutions), and essay questions that were mapped directly to learning objectives in the course textbook and syllabus.[17] Students had 210 minutes to complete the exam in an online testing platform, which required the students to use a computer to answer questions.[18] The testing software automatically graded all

[17] The final exam accounts for 25 percent of the total course points (250 of 1000). Students are informed on the first day of class that failure to pass the final exam could constitute grounds for failure of the entire course, regardless of performance on pervious events. Each type of question is weighted differently. For example, multiple choice questions are typically assigned 2 points, and short answer questions are worth 4-6 points each. Each essay question is worth 10 points. Points from multiple choice, short answer, and essay questions account for roughly 65, 20, and 15 percent, respectively, of the exam's total possible points.

[18] To be clear, this testing format required students in all three classroom types (treatment 1, treatment 2, and control) to use a computer on the final exam, regardless of whether they were allowed to use a computer in regular class meetings.

multiple choice and short answer questions, but professors manually scored all essay responses.[19] Notably, nearly all students in our sample sat for the final exam. Only 15 of the 726 students who began the semester did not have final exam scores, implying an attrition rate of roughly two percent.[20]

One potential concern with using final exam scores as an outcome is the possibility that a student's exam score might not only reflect her understanding of the material, but also the relative leniency or severity of her professor's grading. By including professor fixed effects in our regression model, we account for any idiosyncratic grading procedures that a professor applies to all of his students. However, if professors develop a bias against (or in favor of) students who use computers in the classroom, or if a professor's degree of grading leniency is influenced by a student's performance on other parts of the exam, then professor grading procedures could be correlated with assignment to one of our treatment arms. Neither of these concerns is relevant to multiple choice and short answer questions, which automatically receive grades from the online testing platform, but they are germane to essay questions. When a professor begins grading a new exam, he is immediately prompted by the online testing platform to input a grade for the first essay question. While deciding the essay question score, the professor can observe the graded student's name and current performance on all multiple choice and short answer questions. This concurrent knowledge of a student's "running average" may influence the professor's grading decisions on the essay questions.

[19] For short answer graphing questions, the testing software automatically awards a zero if a student answers any element of a multi-part graphing question incorrectly. Therefore, the course director issues grading guidance for these multi-part questions to professors prior to the exam. This step aids in standardizing the process of awarding "partial credit" across the course. For essay questions, the course director enters an example of a full credit answer in the professor's answer key. However, it does not specify point allocations for each element of the essay answer, and professor discretion plays a major role in determining student essay grades.

[20] Attrition is not significantly correlated with assignment to either treatment group.

To investigate the possibility that grades reflect grader bias rather than academic achievement, Appendix Table 1 compares the percentage of variation in test scores explained by professor fixed effects (the partial R-squared when adding professor fixed effects) for multiple choice, short answer, and essay questions. Column 1 of each panel reports estimates of equation (1) where Z_{jht} is an indicator variable that equals 1 if the classroom identified by professor j, class hour h, and semester t is assigned to either treatment arm. Column 2 reports estimates of an analogous equation that excludes professor fixed effects. A comparison of the R-squared reported in columns 1 and 2 of panel A indicates that professor fixed effects explain roughly 3 percent of the variation in multiple choice test scores (0.479-0.451=0.028). Similarly, professor fixed effects explain only 4 percent of the variation in short answer test scores. On the other hand, professor fixed effects explain 32 percent of the variation in essay question test scores. It is also noteworthy that the standard error of the coefficient for Z_{jht} triples when professor fixed effects are excluded from essay score estimates. Furthermore, baseline GPAs and ACT scores exhibit substantially less correlation with essay scores than they do with multiple choice and short answer scores.

Taken together, the evidence in Appendix Table 1 indicates that essay scores do not provide an accurate measurement of student achievement. Therefore, while we report estimates for all three types of questions in our analysis, our preferred outcome is the composite of a student's multiple choice and short answer scores.[21] For this particular outcome, the average score among students in our sample was roughly 72 percent, with a standard deviation of 9.2 percentage points.

[21] For students who took the introductory economics course in the fall semester of the 2015-2016 academic year, final exam scores exclude six multiple choice and short answer questions that pertained to lesson objectives covered during the personal finance block of the course. All students were required to use laptop computers during the personal finance classes. The six personal finance questions constituted 5 percent of the total final exam grade and were not part of the final exam for the 2014-2015 academic year. Below we investigate whether students in classrooms that permitted computers scored higher on personal finance questions than students in the control group.

Throughout our remaining analysis, we standardize test scores to have a mean of zero and a standard deviation of one for all students who took the exam in the same semester.

VI. RESULTS

A. Effects of permitting laptops or tablets on academic performance

We begin our analysis by comparing exam scores of students in classrooms assigned to either treatment arm to the scores of students assigned to classrooms where laptops and tablets were prohibited. Panel A of Table 3 reports estimates of equation (1) where the outcome is the composite of a student's multiple choice and short answer scores. The point estimate of -0.21, reported in column 1, indicates that exam scores among students in classrooms that permitted laptops and tablets (treatment groups 1 and 2) were 0.21 standard deviations (hereafter σ) below the exam scores of students in classrooms that prohibited computers (control group). In columns 2, 3 and 4 we add demographic, baseline GPA, and ACT scores, respectively and ACT / baseline GPA, respectively.[22] The estimated coefficient falls but remains statistically significant at -0.18σ.

To provide context for the magnitude of this estimate, we can compare the effect of permitting computer usage on exam scores to the estimated effect of baseline GPAs on the same outcome. As seen in column 3 of panel A, the effect of being assigned to a classroom that permits computers is roughly 17 percent as large as the association between a one point reduction in baseline GPAs and final exam scores $\left(\frac{-0.19}{1.13} = 0.17\right)$. To put this another way, a student in a

[22] The full set of controls for the regression estimates reported in column 4 include indicators for gender, white, black, Hispanic, prior military service, and Division I athlete as well as linear terms for age, composite ACT score, and baseline GPA.

classroom that prohibits computers is on equal footing with her peer who is in a classroom that allows computers but has a GPA that is one-third of a standard deviation higher than her GPA.[23]

Subsequent panels of Table 3 report estimates for multiple choice scores, short answer scores, and essay scores. Permitting laptops or computers appears to reduce multiple choice and short answer scores, but has no effect on essay scores, as seen in Panel D. Our finding of a zero effect for essay questions, which are conceptual in nature, stands in contrast to previous research by Mueller and Oppenheimer (2014), who demonstrate that laptop note-taking negatively affects performance on both factual and conceptual questions. One potential explanation for this effect could be the predominant use of graphical and analytical explanations in economics courses, which might dissuade the verbatim note-taking practices that harmed students in Mueller and Oppenheimer's study. However, considering the substantial impact professors have on essay scores, as discussed above, the results in panel D should be interpreted with considerable caution.

B. Distinguishing between treatment arms.

Interestingly, the reduction in exam performance associated with permitting computer usage appears to occur in both classrooms that permit unrestricted computer usage and classrooms that permit only modified-tablet usage. Table 4 reports estimates that are similar to those reported in Table 3, except that they only compare students in classrooms that permitted laptops and tablets without restriction (treatment group 1) to students in classrooms that prohibited computers. The precisely estimated -0.18σ, reported in column 4 of panel A, suggests that allowing computers in the classroom reduces average grades by roughly one-fifth of a standard deviation. It is worth noting that including demographic, baseline GPA, and ACT controls attenuates the estimates in

[23] The standard deviation of baseline GPAs is 0.53 among students in our sample.

panel A of Table 4 from -0.28σ to -0.18σ. This is due to random differences in the composition of students between the first treatment arm and the control group. Importantly, however, these estimates are statistically indistinguishable. The other results in Table 4 indicate that unrestricted computer usage reduces multiple choice and short answer scores but has no effect on essay scores, consistent with our results from Table 3.

Table 5 reports estimates of equation (1) after restricting the sample to students in either modified-tablet classrooms (treatment group 2) or in classrooms that prohibited computers. When the full set of controls are included, permitting modified-tablet usage reduces exam scores by 0.17σ, which is nearly identical to the estimated effect of permitting unrestricted laptop or tablet usage (compare column 4, panel A, of Table 4 to column 4, panel A of Table 5). Thus, it appears that even requiring students to use computing devices in a manner that is conducive to professor monitoring still negatively impacts student performance.

C. *Effects by subgroup*

Table 6 explores whether treatment effects vary by subgroups by conditioning the sample based on gender, race, baseline GPA, ACT scores, and predicted scores. Although differential treatment effects by subgroup are generally not statistically distinguishable, Table 6 does reveal some interesting differences. In particular, permitting computers reduces male academic performance by 0.21σ but appears to have little effect on the academic performance of women (columns 1 and 2 of Panel A). Columns 3 and 4 of Panel A reveal that the negative impact exists for both nonwhite and white students. There is also modest evidence that permitting computers is most harmful to students with relatively strong baseline academic performance. This can be seen in panel B of Table 6, which reports estimates for students who fall within the lower, middle, and upper-third of the distribution of baseline GPAs. Permitting computers appears to lower exam scores of students in the upper-tercile of baseline GPAs by 0.25σ, but only lowers exam scores by 0.10σ for students in the lowest tercile of baseline GPAs. Panel C reveals a similar pattern: permitting computers has a small, statistically insignificant effect on students with relatively low ACT scores, but it reduces exam performance by 0.24σ for students in upper-tercile of the ACT distribution.

To further investigate whether computer and tablet usage is most harmful for students who would otherwise perform well in the absence of treatment, we use predicted exam scores to bin individuals into the bottom, middle, and top of the class. Following the method suggested by Abadie, Chingos, and West (2013), we first compute predicted exam scores for those in the *control* group, using the leave-out fitted values. This method minimize the possibility that outcomes will be mechanically correlated with predicted performance:

(3) $$Y_k = \beta'_{(-i)} X_k + \varepsilon_k; \;\; k \neq i,$$

Y_k is individual test score and X_k includes individual covariates (gender, race, age, prior military service, Division I athlete, baseline GPA, and composite ACT score). We leave out each person individually (*i*) when predicting their exam score. We then use covariate information on students in the control group to construct predicted exam scores for students in the treatment groups:[24]

(4) $$\hat{Y}_i = \hat{\beta}'_{(-i)} X_i.$$

Panel D of Table 6 reports estimates of equation (1) for students within the lower, middle, and upper-third of the distribution of $\hat{Y}_i$. Consistent with the results in panels B and C, those with the highest predicted exam score are most negatively affected by the treatment. It could be that

[24] Note that equation (6) also constructs predicted exam scores for students who are not in the control group. Because only students in the control group are used in the estimation of $\hat{\beta}'_{(-i)}$, leave-out fitted values and leave-in fitted values are identical for students in laptop or modified-tablet classrooms.

students with relatively low predicted achievement find the course curriculum challenging, regardless of available enablers or distractions. Alternatively, professors might make more of an effort to engage with students who are not performing well in the class. Still, the point estimates in all three columns of panels B, C, and D are statistically indistinguishable, so these could be chance findings.

D. 2SLS estimates

As discussed above, if we assume that allowing computers in class only influences a student's academic performance through her own propensity to use a computer, then we can produce 2SLS estimates of computer usage on academic performance. In this setting, assignment to either treatment group is an instrument for actual computer usage, which we asked professors to record on three separate occasions during each semester. Just over 60 percent of students assigned to classrooms that permitted laptops or tablets used a computing device during at least one of these three classes during the semester, which we henceforth define as any computer usage. This can be seen column 1, panel A, of Table 7, which reports first stage estimates from a regression that is similar to equation (1), but where the outcome is an indicator for any computer usage. Scaling the reduced form effect of being assigned to a classroom in either treatment group by this first stage suggests that computer usage reduces exam performance by approximately 0.28σ. OLS estimates comparing exam scores of students with any computer usage to students with no computer usage, reported in column 3, are noticeably smaller than the corresponding 2SLS estimates, potentially suggesting positive selection into computer usage (students who would have performed better on exams are more likely to use computers).

Columns 4 through 6 of Table 7 report estimates that are analogous to the estimates reported in columns 1 through 3, except now we define our endogenous variable as average computer use over the semester. For example, a student observed using a computer during only one of the three days where professors recorded computer usage has an average computer use value of one-third. Since not all students ever recorded as using a computing device used one each day of class, 2SLS estimates constructed from the average usage variable are larger in magnitude than those constructed using the any use variable (-0.42σ relative to -0.28σ). While average computer use might provide a more accurate measure of the prevalence of computing devices on a typical class day, we believe first stage estimates using this endogenous variable will be biased downwards (and therefore 2SLS estimates will be biased upwards) because not all professors reported attendance along with computer usage. Thus, students who are absent from class are recorded as not having used a computing device.

Panels B and C of Table 7 report first stage, 2SLS, and OLS estimates of unrestricted laptop or tablet usage (panel B) and modified-tablet usage (panel C) separately.[25] The first stage estimates reported in columns 1 and 4 suggest that requiring students to keep their tablets face-up on the desk substantially reduces the number of computing devices in the classroom. Whereas nearly 80 percent of students in classrooms that permitted unrestricted computer usage ever used a laptop or tablet (panel B), only 41 percent of students in modified-tablet classrooms used a tablet on at least one of the three days where professors recorded usage (panel C). Since both treatments had similar impacts on average classroom performance (Tables 4 and 5), it is surprising that 2SLS estimates

[25] As with Tables 4 and 5, panel B restricts the sample to students in the control group and treatment group 1 while panel C restricts the sample to students in the control group and treatment group 2. Although we did not require professors to distinguish between laptop and tablet usage classrooms that permitted unrestricted computer use, most professors who taught classrooms in treatment group 1 indicated that laptops were far more common than tablets. We again emphasize that laptop and tablet usage at West Point are not impacted by differences in student resources or differential access to the Internet. West Point "issues" a laptop and tablet computer to all students and each classroom in the study was equipped with wireless Internet at the time of the experiment.

for modified-tablet usage are twice as negative as 2SLS estimates for unrestricted computer usage in Table 6. However, given the relatively large standard errors for the comparison of modified-tablet to prohibited-use classrooms, owing mainly to the relatively small first stage, we urge considerable caution in making this comparison.[26]

VII. ROBUSTNESS CHECKS

A. Clustering

Although it would normally be appropriate to cluster standard errors at the classroom level, inference based on robust standard errors is actually more conservative than inference based on clustered standard errors. To see this more clearly, Appendix Table 3 compares robust, conventional, and clustered standard error estimates for the specification described by equation (1). Robust and conventional standard errors are nearly identical, but, surprisingly, clustered standard errors are substantially smaller than robust standard errors. With 50 classrooms in the experiment, it is unlikely that clustered standard errors are biased downwards as a result of having too few clusters. On the other hand, estimates of the interclass correlation coefficient are indistinguishable from 0 for our preferred outcome of multiple choice and short answer questions, suggesting that little correlation exists in test scores within classrooms after including professor and class hour fixed effects.

To further substantiate the precision of our estimates, we construct estimates of equation (1) using the two-step grouped-data estimation procedure for models with microcovariates

[26] For completeness, Appendix Table 2 reports 2SLS estimates of laptop or modified usage by subgroup, where assignment to a classroom in either treatment group instruments for any laptop or tablet usage. The results are similar to those reported in Table 6.

described by Angrist and Pischke (2009; pp. 313-314). In the first step of this procedure, we construct covariate-adjusted classroom effects by estimating:

(5) $$Y_{ijht} = \mu_{jht} + \gamma' X_i + \eta_{ijht}$$

Here, μ_{jht} is an indicator, or fixed effect, for each classroom in our experiment. In the second step, we regress the estimated classroom fixed effects from equation (5), $\widehat{\mu_{jht}}$, on classroom-level variables, where each observation (i.e. each classroom) is weighted by the number of students in the classroom:

(6) $$\widehat{\mu_{jht}} = \kappa_{jt} + \lambda_{ht} + \pi Z_{jht} + \epsilon_{jht}$$[27]

Column 4 of Appendix Table 3 reports standard errors of π using this two-step method, with P-values based on inference from a t-distribution with 27 degrees of freedom.[28] Even this conservative method of inference indicates that the effect of permitting computers on our preferred outcome (multiple choice plus short answer scores) is significant at the 1 percent level. The effects on multiple choice and short answer scores, estimated separately, are still significant at the 1 and 5 percent levels, respectively.

B. *Additional placebo checks*

The combination of random assignment of classrooms to treatment arms and the inability of students to select their professor or class hour makes it unlikely that our results suffer from omitted variable bias. Still, students assigned to either treatment arm could potentially have had a stronger baseline knowledge of economics than students assigned to the control group. To check

[27] As a reminder, Z_{jht} is equal to 1 for students in either treatment group, the term κ_{jt} includes fixed effects for each combination of professor and semester, λ_{ht} includes fixed effects for each combination of class-hour and semester.

[28] This follows the suggestion of Donald and Lang (2007). With 50 classrooms, 16 combinations of professor and semester, and 8 combinations of class hour and semester, the residual degrees of freedom is 50-15-7-1=27.

for this possibility, we constructed a pre-exam, modeled after the Test of Understanding in College Economics (TUCE), and asked professors to proctor it at the beginning of the semester. Unfortunately, we only received permission to implement this exam during the spring semester of the 2014-2015 academic year but not during the fall semester of the 2015-2016 academic year.[29] In column 3 of Appendix Table 4 we run our same reduced-form regression as Table 3 but restrict the sample to those individuals in the spring who took the TUCE exam. We again find that access decreases exam scores (coefficient = -0.15). In column 4, we then run the same sample with the TUCE (pre-exam) score as the outcome and find a coefficient of +0.13. While the estimates reported in columns 3 and 4 of Appendix Table 4 are not statistically significant, they indicate that among the subsample of students who sat for both the final exam and the pre-exam, those assigned to classrooms that permitted laptops or tablets performed worse on the final exam, but better on the pre-exam, than those assigned to classrooms that prohibited computers.

Students who took the course in the fall semester of the 2015-2016 academic year did not take a pre-exam, but their final exam did cover material from a four lesson personal finance block where all students, including those in the control group, were required to bring computers to class as part of in-classroom instruction.[30] Considering that no classrooms were prohibited from using computers during these lessons, assignment to classrooms that permitted laptops or tablets throughout the semester should not be associated with a decrease in exam scores derived from questions based on the personal finance lessons. In column 5, we restrict the sample to students in the fall semester and find similar results to our main findings (column 1) and those in the spring

[29] One professor also chose not to administer the pre-exam during the spring semester of the 2014-2015 academic year.

[30] Students who took the course in the spring semester of the 2014-2015 academic year also received four classes of personal finance instruction, but material covered during these lessons was not tested on their final exam.

semester (column 2). In column 6, the dependent variable is instead the score on questions covered in the personal finance lessons. The estimate reported in column 6 reveals that on sections of the test covered in classes where all students were exposed to equal treatment of computer access, there is no difference in exam performance.

VIII. CONCLUSION

The results from our randomized experiment suggest that computer devices have a substantial negative effect on academic performance. Our estimates imply that permitting computers or laptops in a classroom lowers overall exam grades by around one-fifth of a standard deviation. Scaling this reduced-form estimate by the percentage of students in either treatment group who actually used computers implies that using a laptop or tablet reduces final exam scores by 0.28 standard deviations, although we admit that these 2SLS estimates can only be interpreted as causal under strong assumptions. Comparing each treatment arm to the control group separately, we estimate that unrestricted laptop or tablet access reduces test scores by 0.18σ, while modified-tablet access reduces test scores by 0.17σ. Given a standard deviation of 9.2, this amounts to about a 1.7 point reduction (or a 2.6 point reduction for 2SLS estimates) on a 100 point scale.

There are at least a few channels through which computer usage could affect students. First, students who are using their tablet or computer may be surfing the Internet, checking email, messaging with friends, or even completing homework for that class or another class. All of these activities could draw a student's attention away from the class, resulting in a lower understanding of the material. Second, Mueller and Oppenheimer (2014) find that students required to use computers are not as effective at taking notes as students required to use pen and paper, which could also lower test scores. Third, professors might change their behavior – either teaching differently to the whole class or interacting differently to students who are on their computer or tablet relative to how they would have otherwise. Regardless of the mechanism, our results indicate that students perform worse when personal computing technology is available. It is quite possible that these harmful effects could be magnified in settings outside of West Point. In a learning environment with lower incentives for performance, fewer disciplinary restrictions on distracting behavior, and larger class sizes, the effects of Internet-enabled technology on achievement may be larger due to professors' decreased ability to monitor and correct irrelevant usage.

The estimated effects of our two treatments are nearly identical, suggesting that even allowing students to use computer devices in a manner that is conducive to professor monitoring (e.g. tablets flat on the desk) can have harmful effects on classroom performance. The tablet computers used in this experiment (iPads) use a mobile device operating system, which allows for cloud access to web applications typically used on smart phones. Despite the professor's ability to monitor usage, students may have greater propensity to access distracting web applications or message with their friends via the tablet computer than with a laptop computer. An alternative explanation could be a lack of student familiarity in using a tablet computer for academic purposes. While students may have regularly used laptop or desktop computers in secondary school classrooms, tablet computers are a relatively new technology and may not be as fully integrated into high school education, and thus their ability to effectively take notes on a tablet may be limited.

Comparing our results to value-added estimates of the impact of teacher quality at the high-school level suggests that removing laptops and tablets from a classroom is equivalent to improving the quality of the teacher by more than a standard deviation.[31] Our estimate of -0.18σ

[31] Aaronson, Barrow, and Sander (2007) find that a one standard deviation improvement in teacher quality increases test scores by 0.15 standard deviations. Several other value-added studies find results of similar magnitude (between

is also more negative than the estimated effect of increasing class size by one standard deviation, as reported in Bandiera, Larcinese, and Rasul (2010). The results reported in this study also appear to be larger in magnitude than estimates of peer effects at the same institution: exploiting random assignment of cadets to social groups at West Point, Lyle (2009) finds that a standard deviation increase in peer-group 75—25 differential in math SAT scores increases the peer-group's average freshman math grade by 13 percent of a standard deviation while peer-group SAT means have no effect on freshman math grades.[32] Using an IV approach to study the effects of study time on grades, Stinebrickner and Stinebrickner (2008) find that an additional hour of studying increases GPA by around 0.36 points, or roughly half of a standard deviation.

Removing computers from the classroom could also have larger effects on student performance than merit-based financial aid reward programs. Angrist, Oreopoulos, and Williams (2014) measure the effect of a such a program and find that, while second year college students increased the number of courses in which they scored above the reward threshold, the program did not significantly increase overall GPA. They also summarize evidence of previous randomized control trials studying the impact of paying students on GPA and course grades. Most studies find very little impact, and when there are statistically significant results, the effects are mostly concentrated among women. In one example with a similar result to ours, Angrist, Lang, and Oreopoulos (2009) find that financial incentives and support services increase GPA by 0.3 standard deviations for women.

0.10 and 0.15) for students in lower grade levels. See Chetty, Friedman, and Rockoff (2014), Rivkin, Hanusheck, and Kain (2005), Rockoff (2005), and others.

[32] Lyle (2007) also finds no evidence that peer-group means impact academic performance, but Carrell, Fullerton, and West (2009), who explore the possibility of peer effects among students at the U.S. Air Force Academy, find that a 100-point increase in peer-group average SAT verbal scores increases individual GPAs by 0.4 grade points.

We want to be clear that we cannot relate our results to a class where the laptop or tablet is used deliberately in classroom instruction, as these exercises may boost a student's ability to retain the material. Rather, our results relate only to classes where students have the option to use computer devices to take notes. We further cannot test whether the laptop or tablet leads to worse note taking, whether the increased availability of distractions for computer users (email, facebook, twitter, news, other classes, etc.) leads to lower grades, or whether professors teach differently when students are on their computers. Given the magnitude of our results, and the increasing emphasis of using technology in the classroom, additional research aimed at distinguishing between these channels is clearly warranted.

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago Public High Schools. *Journal of Labor Economics*, 95-135.

Abadie, A., Chingos, M. M., & West, M. R. (2013). Endogenous stratification in randomized experiments (No. w19742). *National Bureau of Economic Research Working Paper*.

Aguilar-Roca, N. M., Williams, A. E., & O'Dowd, D. K. (2012). The impact of laptop-free zones on student performance and attitudes in large lectures. *Computers and Education, 59,* 1300-1308.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton, NJ: Princeton University Press.

Angrist, J., Lang, D., & Oreopoulos, P. (2009). Incentives and services for college achievement: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, 136-163.

Angrist, J., Oreopoulos, P., & Williams, T. (2014). When Opportunity Knocks, Who Answers? New Evidence on College Achievement Awards. *Journal of Human Resources*, *49*(3), 572-610.

Bandiera, O., Larcinese, V., & Rasul, I. (2010). Heterogeneous class size effects: New evidence from a panel of university students. *The Economic Journal, 120*(549), 1365-1398.

Barak, M., Lipson, A., & Lerman, S. (2006). Wireless Laptops as Means for Promoting Active Learning in Large Lecture Halls. *Journal of Research on Technology in Education*, 245-263.

Carrell, S. E., Fullerton, R. L., & West, J. E. (2009). Does your cohort matter? Measuring peer effects in college achievement. *Journal of Labor Economics*, 439-464.

Chetty, R., Friedman, J. N., & Rockoff, J. E. (2014). Measuring the Impacts of Teachers I: Evaluating Bias in Teacher Value-Added Estimates. *American Economic Review, 109*(9), 2593-2632.

Donald, S. G., & Lang, K. (2007, May). Inference with Difference-in-Differences and Other Panel Data. *The Review of Economics and Statistics, 89*(2), 221-233.

Fried, C. B. (2008). In-class laptop use and its effects on student learning. *Computers and Education, 50*(3), 906-914.

Grace-Martin, M., & Gay, G. (2001). Web Browsing, Mobile Computing, and Academic Performance. *Journal of Educational Technology & Society, 4*(3), 95-107.

Gross, T. (2014, December 30). This year, I resolve to ban laptops from my classroom. *The Washington Post*.

Hembrooke, H., & Gay, G. (2003). The Laptop and the Lecture: The Effects of Multitasking in Learning Environments. *Journal of Computing in Higher Education, 15*(1), 46-64.

Kraushaar, J. M., & Novak, D. C. (2010). Examining the Effects of Student Multitasking With Laptops During the Lecture. *Journal of Information Systems Education, 21*(2), 241-251.

Lyle, D. S. (2007). Estimating and interpreting peer and role model effects from randomly assigned social groups at West Point. *The Review of Economics and Statistics*, 289-299.

Lyle, D. S. (2009). The effects of peer group heterogeneity on the production of human capital at West Point. *American Economic Journal: Applied Economics*, 69-84.

Mueller, P. A., & Oppenheimer, D. M. (2014, April 23). The Pen is Mightier Than the Keyboard: Advantages of Longhand Over Laptop Note Taking. *Psychological Science*, 1-10. doi:10.1177/0956797614524581

National Liberal Arts Colleges Rankings. (2016). Retrieved May 1, 2016, from US News & World Report Education Rankings & Advice: http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/national-liberal-arts-colleges?int=a73d09

Rivkin, S. G., Hanusheck, E. A., & Kain, J. F. (2005). Teachers, Schools, and Academic Achievement. *Econometrica*, 417-458.

Rockoff, J. E. (2005). Teachers, Schools and Academic Achievement. *Econometrica, 73*, 417-458.

Sana, F., Weston, T., & Cepeda, N. J. (2012). Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education*, *62*, 24-31.

Stinebrickner, R., & Stinebrickner, T. R. (2008). The causal effect of studying on academic performance. *The BE Journal of Economic Analysis & Policy, 8*(1).

U.S. Department of Education, Office of Educational Technology. (2016). Future Ready Learning: Reimagining the Role of Technology in Education. Washington, DC.

Walstad, W. B., Watts, M., & Rebeck, K. (2007). *Test of Understanding in College Economics, Fourth Edition.* New York: National Council on Economic Education.

Wells, J., & Lewis, L. (2006). *Internet Access in U.S. Public Schools and Classrooms: 1994-2005 (NCES 2007-020).* Washington, D.C.: National Center for Education Statistics. Retrieved May 1, 2016, from http://nces.ed.gov/pubs2007/2007020.pdf

Wurst, C., Smarkola, C., & Gaffney, M. A. (2008). Ubiquitous laptop usage in higher education: Effects on student achievement, student satisfaction, and constructivist measures in honors and traditional classrooms. *Computers & Education, 51*, 1766-1783.

Table 1: Comparisons to Other Schools

	Panel A: 2014-2015 Common Data Sets						Panel B: 2013-2014 IPEDS				
								Public 4-Year Schools		All 4 Year Schools	
	United States Military Academy, NY (Ranked 22)	Williams College, MA (Ranked 1)	Pomona College, CA (Ranked 4)	Davidson College, NC (Ranked 9)	Washington and Lee, VA (Ranked 14)	Colorado College, CO (Ranked 25)	United States Military Academy	All	Pop between 1,000 & 10,000	All	Pop between 1,000 & 10,000
	Full-Time Degree Seeking Undergrads										
Undergraduate Population	4,414	2,014	1,625	1,765	1,876	2,036	4,591	9,215	4,645	2,841	3,213
Student to Faculty Ratio	7:1	7:1	8:1	10:1	8:1	10:1					
% Female	17%	51%	51%	51%	50%	53%	17%	56%	56%	56%	58%
Non-resident Aliens	1%	7%	9%	6%	4%	6%	1%	3%	2%	3%	3%
Hispanic	11%	12%	14%	7%	4%	9%	10%	12%	10%	12%	12%
Black / AA, non-Hispanic	9%	7%	7%	6%	2%	2%	8%	14%	17%	16%	15%
White, non-Hispanic	67%	56%	43%	69%	83%	66%	69%	59%	59%	54%	56%
American Indian	1%	0%	0%	1%	0%	0%	1%	2%	2%	1%	1%
Asian, Non-Hispanic	6%	11%	13%	6%	3%	5%	6%	4%	3%	4%	3%
Pacific Islander	0%	0%	0%	0%	0%	0%	1%	1%	1%	0%	0%
Two or more races	3%	7%	7%	4%	2%	8%	4%	3%	3%	2%	2%
Race Unknown	2%	0%	7%	2%	2%	3%	1%	3%	4%	7%	7%
% from Out of State	93%	88%	69%	77%	86%	82%					
	Freshman Profile										
ACT Composite											
25th Perc	26	31	31	28	30	28	27	20	19	20	21
75th Perc	31	34	34	32	33	32	30	25	24	25	26
SAT Critical Reading											
25th Perc	570	680	690	610	660	620	580	459	442	468	471
75th Perc	690	790	770	720	730	730	695	565	544	577	578
SAT Math											
25th Perc	590	670	690	620	660	630	600	474	453	477	480
75th Perc	700	770	770	720	730	730	690	581	556	585	587

Notes: This table compares The United States Military Academy, West Point to other 4 year undergraduate institutions. Panel A reports statistics from the 2014-2015 Common Datasets from West Point and other Schools in the top 25 of National Liberal Arts Schools. Data in panel B comes from the Integrated Postsecondary Education Data System for the 2013-2014 academic year.

Table 2. Summary Statistics and Covariate Balance

Baseline Characteristic	Mean Characteristics			Regression of LHS Var on Indicator for Intention-To-Treat		
	Control	Treatment 1 (Computers/Tablets)	Treatment 2 (Tablets, face up)	Both Treatments vs. Control	Treatment 1 vs. Control	Treatment 2 vs. Control
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.167	0.202	0.192	0.033	0.059	0.004
				(0.030)	(0.036)	(0.038)
White	0.637	0.673	0.663	0.023	0.019	0.019
				(0.039)	(0.045)	(0.051)
Black	0.115	0.097	0.111	-0.024	-0.021	-0.028
				(0.026)	(0.029)	(0.036)
Hispanic	0.126	0.129	0.087	0.003	0.020	-0.028
				(0.028)	(0.033)	(0.034)
Age	20.12	20.15	20.15	0.03	0.05	0.06
	[1.06]	[1.00]	[0.96]	(0.08)	(0.09)	(0.10)
Prior Military Service	0.193	0.185	0.159	-0.017	-0.002	-0.005
				(0.031)	(0.036)	(0.039)
Division I Athlete	0.289	0.395	0.346	0.052	0.075*	0.039
				(0.036)	(0.044)	(0.046)
GPA at Baseline	2.87	2.82	2.89	-0.01	-0.05	0.03
	[0.52]	[0.54]	[0.51]	(0.04)	(0.05)	(0.05)
Composite ACT	28.78	28.30	28.30	-0.34	-0.37	-0.54
	[3.21]	[3.46]	[3.27]	(0.26)	(0.31)	(0.33)
P Val (Joint χ^2 Test)				0.610	0.532	0.361
Observations	270	248	208	726	518	478

Notes: This table reports descriptive statistics of students in classrooms participating in the experiment. Columns (1), (2), and (3) report mean characteristics of the control group (classrooms where computers and tablets are prohibited), treatment group 1 (computers and tablets are permitted without restriction), and treatment group 2 (tablets are permitted if they are face up). Standard deviations are reported in brackets. Columns (4), (5), and (6) report coefficient estimates from a regression of the baseline charactersitics on an indicator variable that equals one if a student is assigned to a classroom in the indicated treatment group. The regressions used to construct estimates in columns (4), (5), and (6) include instructor fixed effects, class hour fixed effects, semester fixed effects, and (instructor) x (semester) and (class hour) x (semester) interactions. Robust standard errors are reported in parentheses. The reported P-values are from a joint test of the null hypothesis that all coefficients are equal to zero. ***,**, and * denote significance at the 1%, 5%, and 10% level, respecitvely.

Table 3. Laptop and Modified-Tablet Classrooms vs. Non-Computer Classrooms

	(1)	(2)	(3)	(4)
A. Dependent Variable: Final Exam Multiple Choice and Short Answer Score				
Laptop/Tablet Class	-0.21**	-0.20***	-0.19***	-0.18***
	(0.08)	(0.07)	(0.06)	(0.06)
GPA at start of course			1.13***	1.00***
			(0.06)	(0.06)
Composite ACT				0.06***
				(0.01)
Demographic Controls		X	X	X
Observations	711	711	711	711
B. Dependent Variable: Final Exam Multiple Choice Score				
Laptop/Tablet Class	-0.17**	-0.16**	-0.15**	-0.14**
	(0.08)	(0.07)	(0.06)	(0.06)
GPA at start of course			1.01***	0.89***
			(0.06)	(0.07)
Composite ACT				0.05***
				(0.01)
Demographics		X	X	X
Observations	711	711	711	711
C. Dependent Variable: Final Exam Short Answer Score				
Laptop/Tablet Class	-0.23***	-0.23***	-0.22***	-0.21***
	(0.08)	(0.07)	(0.06)	(0.06)
GPA at start of course			1.05***	0.93***
			(0.06)	(0.07)
Composite ACT				0.05***
				(0.01)
Demographics		X	X	X
Observations	711	711	711	711
D. Dependent Variable: Final Exam Essay Questions Score				
Laptop/Tablet Class	0.00	0.00	0.01	0.02
	(0.07)	(0.06)	(0.06)	(0.06)
GPA at start of course			0.76***	0.70***
			(0.06)	(0.07)
Composite ACT				0.03***
				(0.01)
Demographics		X	X	X
Observations	711	711	711	711

Notes: This table reports estimates from a regression of exam scores on an indicator for being assigned to a classroom that permits either laptops or tablets. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. All estimates include instructor fixed effects, class hour fixed effects, semester fixed effects, and (instructor) x (semester) and (class hour) x (semester) interactions. Demographic controls include indicators for female, white, black, hispanic, prior military service, athlete, and a linear term for age at the start of the course. Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respecitvely.

Table 4. Unrestricted Laptop/Tablet Classrooms vs. Non-Computer Classrooms

	(1)	(2)	(3)	(4)
A. Dependent Variable: Final Exam Multiple Choice and Short Answer Score				
Computer Class	-0.28***	-0.23***	-0.19***	-0.18***
	(0.10)	(0.09)	(0.07)	(0.07)
GPA at start of course			1.09***	0.92***
			(0.07)	(0.07)
Composite ACT				0.07***
				(0.01)
Demographic Controls		X	X	X
Observations	507	507	507	507
B. Dependent Variable: Final Exam Multiple Choice Score				
Computer Class	-0.24**	-0.19**	-0.16**	-0.15**
	(0.10)	(0.09)	(0.07)	(0.07)
GPA at start of course			0.96***	0.80***
			(0.07)	(0.08)
Composite ACT				0.07***
				(0.01)
Demographics		X	X	X
Observations	507	507	507	507
C. Dependent Variable: Final Exam Short Answer Score				
Computer Class	-0.26***	-0.23***	-0.19***	-0.18***
	(0.09)	(0.09)	(0.07)	(0.07)
GPA at start of course			1.04***	0.90***
			(0.07)	(0.08)
Composite ACT				0.06***
				(0.01)
Demographics		X	X	X
Observations	507	507	507	507
D. Dependent Variable: Final Exam Essay Questions Score				
Computer Class	-0.06	-0.03	0.00	0.00
	(0.08)	(0.08)	(0.07)	(0.07)
GPA at start of course			0.80***	0.72***
			(0.08)	(0.08)
Composite ACT				0.04***
				(0.01)
Demographics		X	X	X
Observations	507	507	507	507

Notes: This table reports estimates from a regression of exam scores on an indicator for being assigned to a classroom that permits laptop and unrestricted tablet usage. The sample used to construct this table excludes students in modified-tablet classrooms. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. All estimates include instructor fixed effects, class hour fixed effects, semester fixed effects, and (instructor) x (semester) and (class hour) x (semester) interactions. Demographic controls include indicators for female, white, black, hispanic, prior military service, athlete, and a linear term for age at the start of the course. Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respecitvely.

Table 5. Modified-Tablet Classrooms vs. Non-Computer Classrooms

	(1)	(2)	(3)	(4)
A. Dependent Variable: Final Exam Multiple Choice and Short Answer Score				
Modified-Tablet Class	-0.17*	-0.17*	-0.19***	-0.17**
	(0.10)	(0.09)	(0.07)	(0.07)
GPA at start of course			1.12***	1.01***
			(0.07)	(0.08)
Composite ACT				0.05***
				(0.01)
Demographic Controls		X	X	X
Observations	466	466	466	466
B. Dependent Variable: Final Exam Multiple Choice Score				
Modified-Tablet Class	-0.14	-0.14	-0.16**	-0.13*
	(0.10)	(0.09)	(0.08)	(0.07)
GPA at start of course			1.03***	0.93***
			(0.07)	(0.08)
Composite ACT				0.04***
				(0.02)
Demographics		X	X	X
Observations	466	466	466	466
C. Dependent Variable: Final Exam Short Answer Score				
Modified-Tablet Class	-0.23**	-0.24***	-0.26***	-0.23***
	(0.10)	(0.09)	(0.08)	(0.08)
GPA at start of course			1.00***	0.85***
			(0.08)	(0.09)
Composite ACT				0.06***
				(0.02)
Demographics		X	X	X
Observations	466	466	466	466
D. Dependent Variable: Final Exam Essay Questions Score				
Modified-Tablet Class	-0.03	-0.03	-0.05	-0.04
	(0.08)	(0.08)	(0.07)	(0.07)
GPA at start of course			0.80***	0.77***
			(0.08)	(0.08)
Composite ACT				0.01
				(0.01)
Demographics		X	X	X
Observations	466	466	466	466

Notes: This table reports estimates from a regression of exam scores on an indicator for being assigned to a classroom that permits modified-tablet usage. The sample used to construct this table excludes students in classrooms where laptops and tablets are permitted without restriction. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. All estimates include instructor fixed effects, class hour fixed effects, semester fixed effects, and (instructor) x (semester) and (class hour) x (semester) interactions. Demographic controls include indicators for female, white, black, hispanic, prior military service, athlete, and a linear term for age at the start of the course. Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respecitvely.

Table 6. Reduced Form Estimates by Subgroup
Dependent Variable: Final Exam Multiple Choice and Short Answer Score

	A. By Demographic Groups			
	Women	Men	Nonwhite	White
	(1)	(2)	(3)	(4)
Laptop/Tablet Class	0.02	-0.21***	-0.22**	-0.18**
	(0.14)	(0.06)	(0.10)	(0.07)
Observations	131	580	244	467
	B. By Baseline GPA			
	Bottom Third of Distribution	Middle Third of Distribution	Top Third of Distribution	
	(1)	(2)	(3)	
Laptop/Tablet Class	-0.10	-0.17*	-0.25**	
	(0.12)	(0.10)	(0.10)	
Observations	236	237	238	
	C. By ACT Score			
	Bottom Third of Distribution	Middle Third of Distribution	Top Third of Distribution	
	(1)	(2)	(3)	
Laptop/Tablet Class	-0.05	-0.22**	-0.24**	
	(0.09)	(0.10)	(0.11)	
Observations	271	221	219	
	D. By Predicted Exam Score			
	Bottom Third of Distribution	Middle Third of Distribution	Top Third of Distribution	
	(1)	(2)	(3)	
Laptop/Tablet Class	-0.07	-0.13	-0.21**	
	(0.11)	(0.10)	(0.10)	
Observations	235	234	242	

Notes: This table reports estimates from a regression of exam scores on an indicator for being assigned to a classroom where laptops or tablets are permitted for the subgroups identified in each column heading. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. All estimates include instructor fixed effects, class hour fixed effects, semester fixed effects, (instructor) x (semester) fixed effects, (class hour) x (semester) fixed effects, linear terms for baseline GPA, composite ACT, baseline age, and indicators for female, white, black, hispanic, prior military service, and athlete. Predicted exam scores are constructed using the method described in Abadie et al. (2013). Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respecitvely.

Table 7. 2SLS and OLS Estimates of Laptop and Tablet Usage
Dependent Variable: Final Exam Multiple Choice and Short Answer Score

	Endogenous Variable: Any Laptop/Tablet Usage			Endogenous Variable: Average Laptop/Tablet Usage		
	First Stage	2SLS	OLS	First Stage	2SLS	OLS
	(1)	(2)	(3)	(4)	(5)	(6)
A. All Classrooms in Sample						
Laptop/Tablet Usage	0.62***	-0.28***	-0.11*	0.42***	-0.42***	-0.16**
	(0.03)	(0.09)	(0.06)	(0.02)	(0.13)	(0.08)
Observations	711	711	711	711	711	711
B. Laptop Classrooms and Non-Computer Classrooms						
Laptop/Tablet Usage	0.79***	-0.23***	-0.19***	0.56***	-0.32***	-0.24***
	(0.03)	(0.08)	(0.07)	(0.02)	(0.11)	(0.09)
Observations	507	507	507	507	507	507
C. Modified-Tablet Classrooms and Non-Computer Classrooms						
Modified-Tablet Usage	0.41***	-0.41**	-0.04	0.24***	-0.68**	-0.13
	(0.04)	(0.17)	(0.09)	(0.03)	(0.29)	(0.14)
Observations	466	466	466	466	466	466

Notes: This table reports 2SLS and OLS estimates of computer usage on exam scores. In the "First Stage" and "2SLS" columns, an indicator for being assigned to a classroom that permits laptop or tablet usage instruments for actual laptop or tablet usage. Laptop and tablet usage was recorded during three lessons each semester of the experiment. In columns 1 through 3, actual computer usage is coded as an indicator variable for ever using a laptop or tablet during one of the three lessons where computer usage was recorded. In columns 4 through 6, computer usage is coded as the average usage rate over the three lessons where computer usage was recorded (e.g. a student who uses a computer during one of three lessons has an average usage rate of one-third). Column 3 reports OLS estimates from a regression of exam scores on an indicator for ever using a laptop or tablet during the semester while column 6 reports OLS estimates from a regression of exam scores on average laptop or tablet usage rates during the semester. Estimates in panel A include students in all classrooms of the experiment, estimates in panel B exclude students in classrooms where only modified-tablet usage was permitted, and estimates in panel C exclude students in classrooms where laptop and unrestricted tablet usage was permitted. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. All estimates include instructor fixed effects, class hour fixed effects, semester fixed effects, (instructor) x (semester) fixed effects, (class hour) x (semester) fixed effects, linear terms for baseline GPA, composite ACT, baseline age, and indicators for female, white, black, hispanic, prior military service, and athlete. Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respecitvely.

Appendix Table 1. Reduced Form Estimates With and Without Instructor Fixed Effects

	A. DV: Multiple Choice		B. DV: Short Answer		C. DV: Essay Questions	
	(1)	(2)	(1)	(2)	(1)	(2)
Laptop/Tablet Class	-0.14**	-0.12*	-0.21***	-0.20**	0.02	0.01
	(0.06)	(0.06)	(0.06)	(0.09)	(0.06)	(0.17)
GPA at start of course	0.89***	0.89***	0.93***	0.91***	0.70***	0.59***
	(0.07)	(0.08)	(0.07)	(0.06)	(0.07)	(0.09)
Composite ACT	0.055***	0.050***	0.051***	0.052***	0.030***	0.039***
	(0.012)	(0.011)	(0.012)	(0.014)	(0.010)	(0.014)
Instructor Fixed Effects	X		X		X	
R^2	0.479	0.451	0.424	0.381	0.508	0.191
Observations	711	711	711	711	711	711

Notes: This table reports estimates from a regression of exam scores on an indicator for being assigned to a classroom that permits either laptops or tablets with and without instructor fixed effects. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. Estimates reported in column (1) include instructor fixed effects, class hour fixed effects, semester fixed effects, (instructor) x (semester) fixed effects, (class hour) x (semester) fixed effects, indicators for female, white, black, hispanic, prior military service, and Division I athlete as well as linear terms for age at the start of the course, GPA at the start of the course, and composite ACT score. Estimates reported in column (2) exclude instructor and (instructor) x (semester) fixed effects. Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respecitvely.

Appendix Table 2. 2SLS Estimates of Laptop or modified-tablet Usage by Subgroup
Dependent Variable: Final Exam Multiple Choice and Short Answer Score

	A. By Demographic Groups			
	Women	Men	Nonwhite	White
	(1)	(2)	(3)	(4)
Laptop/Tablet Usage	0.03	-0.35***	-0.34**	-0.30***
	(0.17)	(0.10)	(0.14)	(0.11)
Observations	131	580	244	467

	B. By Baseline GPA		
	Bottom Third of Distribution	Middle Third of Distribution	Top Third of Distribution
	(1)	(2)	(3)
Laptop/Tablet Usage	-0.18	-0.30*	-0.39***
	(0.19)	(0.16)	(0.14)
Observations	236	237	238

	C. By ACT Score		
	Bottom Third of Distribution	Middle Third of Distribution	Top Third of Distribution
	(1)	(2)	(3)
Laptop/Tablet Usage	-0.08	-0.39**	-0.35**
	(0.14)	(0.17)	(0.16)
Observations	271	221	219

	D. By Predicted Exam Score		
	Bottom Third of Distribution	Middle Third of Distribution	Top Third of Distribution
	(1)	(2)	(3)
Laptop/Tablet Usage	-0.10	-0.23	-0.34**
	(0.15)	(0.16)	(0.15)
Observations	235	234	242

Notes: This table reports 2SLS estimates of the effects of laptop or modified-tablet usage on academic performance for the subgroups identified in each column heading. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. An indicator for being assigned to a classroom where computer usage is allowed is an instrument for ever using a computer during the semester. All estimates include instructor fixed effects, class hour fixed effects, semester fixed effects, (instructor) x (semester) fixed effects, (class hour) x (semester) fixed effects, linear terms for baseline GPA, composite ACT, baseline age, and indicators for female, white, black, hispanic, prior military service, and athlete. Predicted exam scores are constructed using the method described in Abadie et al. (2013). Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respecitvely.

Appendix Table 3. Comparison of Standard Errors for Reduced Form Estimates

	Robust Standard Errors	Conventional Standard Errors	Clustered Standard Errors	Group Means
	(1)	(2)	(3)	(4)
	A. DV: Final Exam Multiple Choice and Short Answer Score			
Computer or modified-tablet Classroom	-0.175***	-0.175***	-0.175***	-0.174***
	(0.057)	(0.057)	(0.035)	(0.050)
P-Value	0.0022	0.0023	0.0000	0.0015
Clusters (classrooms)			50	
Residual Deg-of-Freedom	679	679	49	27
Observations	711	711	711	50
	B. DV: Final Exam Multiple Choice Score			
Computer or modified-tablet Classroom	-0.137**	-0.137**	-0.137***	-0.137***
	(0.060)	(0.061)	(0.033)	(0.048)
P-Value	0.0232	0.0244	0.0001	0.0081
Clusters (classrooms)			50	
Residual Deg-of-Freedom	679	679	49	27
Observations	711	711	711	50
	C. DV: Final Exam Short Answer Score			
Computer or modified-tablet Classroom	-0.206***	-0.206***	-0.206***	-0.203**
	(0.063)	(0.064)	(0.059)	(0.085)
P-Value	0.0011	0.0013	0.0011	0.0236
Clusters (classrooms)			50	
Residual Deg-of-Freedom	679	679	49	27
Observations	711	711	711	50
	D. DV: Final Exam Essay Questions Score			
Computer or modified-tablet Classroom	0.016	0.016	0.016	0.015
	(0.057)	(0.059)	(0.062)	(0.082)
P-Value	0.7817	0.7882	0.8006	0.8549
Clusters (classrooms)			50	
Residual Deg-of-Freedom	679	679	49	27
Observations	711	711	711	50

Notes: This table reports estimates from a regression of exam scores on an indicator for being assigned to a classroom that permits either laptops or tablets. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. Estimates reported in columns (1) - (3) include instructor fixed effects, class hour fixed effects, semester fixed effects, (instructor) x (semester) fixed effects, (class hour) x (semester) fixed effects, linear terms for baseline GPA, composite ACT, baseline age, and indicators for female, white, black, hispanic, prior military service, and athlete. Group means are constructed by first regressing the outcome on an indicator variable for each classroom while controlling for individual level covariates, then by regressing the estimated classroom fixed effects on a dummy variable indicating if the classroom is a laptop or modified-tablet classroom, weighting by classroom size, and controlling for instructor fixed effects, class hour fixed effects, semester fixed effects, (instructor) x (semester) fixed effects, and (class hour) x (semester) fixed effects. Standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respecitvely.

Appendix Table 4. Reduced Form Estimates by Semester

	Full Sample	Spring Semester, AY2014-2015			Fall Semester, AY2015-2016	
	DV: Final Exam	DV: Final Exam	DV: Final Exam	DV: TUCE Pre-Exam	DV: Final Exam	DV: Computer Class Questions
	(1)	(2)	(3)	(4)	(5)	(6)
Laptop/Tablet Class	-0.18***	-0.17*	-0.15	0.13	-0.15**	0.003
	(0.06)	(0.10)	(0.11)	(0.15)	(0.07)	(0.093)
TUCE Pre-Exam Sample			X	X		
Observations	711	252	203	203	459	459

Notes: This table reports estimates of the effects of being assigned to a classroom that permits laptop or modified-tablet usage on the outcomes specified in the heading of each column. Final exam scores are scores derived from multiple choice and short answer questions on the final exam, excluding questions from lessons where all classrooms mandated computer use. TUCE Pre-Exam scores are derived from a pre-exam, modeled after the Test of Understanding in College Economics, administered to classrooms during the spring semester of the 2014-2015 academic year. "Computer Class Questions" are scores derived from 6 final exam questions that tested students' understanding of personal finance concepts, where students in all classrooms were required to use computers. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. Estimates in column 1 are from the full sample. Estimates in column 2 are from all students who took the course in the spring semester of the 2014-2015 academic year. Estimates in columns 3 and 4 are from students who took the course in the spring semester of 2014-2015 academic year and who had a valid pre-exam score on file. Estimates in columns 5 and 6 are from all students who took the course in the fall semester of the 2015-2016 academic year. All estimates include instructor fixed effects, class hour fixed effects, semester fixed effects, (instructor) x (semester) fixed effects, (class hour) x (semester) fixed effects, linear terms for baseline GPA, ACT score, baseline age, and indicators for female, white, black, hispanic, prior military service, and Division I athlete. Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respecitvely.