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

By

Susan Payne Carter

Kyle A. Greenberg

Michael S. Walker

March 2016

WORKING PAPER: PLEASE DO NOT CITE OR REDISTRIBUTE WITHOUT CONSENT

Abstract:

We present results from a study that prohibited computer devices in randomly selected classrooms of an introductory economics course at the United States Military Academy. Our results indicate that using a computer device reduces final exam scores by roughly one-fourth of a standard deviation. Through the use of two separate treatment arms, we uncover evidence that the negative impact of computer devices occurs not only in classrooms where laptops and tablets are permitted without restriction, but also in classrooms where students are only permitted to use tablets that must remain flat on the desk surface at all times. Computer usage appears to be most harmful to male students and to students who began the course with high cumulative GPAs.

* The views expressed in this report are those of the authors and do not necessarily reflect the official policy or position of the United States Military Academy, the Department of the Army, the Department of Defense, or the U.S. Government. 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

Americans can now speak of the ubiquitous presence of Internet-enabled classroom technology in the past tense. 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). Adoption of Internet-enabled technology in education has mirrored its incorporation and successful implementation in virtually every sector of the U.S. economy. While this technology undoubtedly aided organizational productivity by increasing worker efficiency and reducing barriers to information, its effect on student achievement in the classroom is unclear. Despite an ongoing debate surrounding its benefit for individual learning, improvement of classroom Internet access continues to serve as a major education 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 colleges and universities battle to attract the best students. Over the course of the past decade, students have become accustomed to near-constant Internet access at home and in the classroom, and, thus, 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 "needs," this increased connectivity also provides opportunities

---

[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 students and teachers to collaborate outside of the classroom, convenient options for student research via university library-enabled online search engines, and continuous access to web-based curriculum, 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, among other improvements over hardcopy books.[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.

However, 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 in the classroom. As institutions (including the location of the present study)[5] continue to push for ever faster and continuous access to wireless Internet and there is a 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, professors and teachings 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 report results from an experiment that randomly permits student access to laptop and tablet computers during an

---

[4] 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.

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, 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.

We find that using a computer device in the classroom over the course of an entire semester reduces student performance on final exam scores by approximately one-fourth of a standard deviation. This substantial negative effect occurs not only in classrooms where laptops and tablets are permitted without restriction, but also in classrooms where students are only permitted to use tablets that must remain flat on the desk surface at all times. Interestingly, we also find evidence that computer usage is most detrimental to male students and to students who entered the course with high cumulative GPAs.

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 experimental restrictions on student behavior by omitting required student tasks, 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, however, allows us to improve upon existing results, as we are able to control for selection into computer usage and avoid the problems associated with student self-reporting of computer activity. Furthermore, 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.

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, and Section VII 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

[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.

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 (whether it is 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 require 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 location breakdowns for West Point relative to five other schools ranked in the top 30 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 students accounting for only 17 percent of the undergrad

---

[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 for the full set of rankings.

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 the Congressional nomination requirement in the admissions process. 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, again, has a lower male to female ratio and a higher percentage of white students. Its average SAT and ACT scores are well above that of the average 4-year institution.

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. Since no students were required to use laptops or tablets in regular class meetings as part of the experiment, assignment to the first treatment will serve as an instrument for laptop/tablet usage and assignment to the second treatment group will serve as an instrument for modified-tablet usage in the analysis that follows.

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 (capped at 1:18 per class by Academy policy). Despite the large enrollment and small classes, 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. We were then able to able 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 number of required courses. As a result, 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 very little 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 their graduating class, the greater their chances of receiving their 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

---

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

their professor elects to report the incident, a student who misbehaves in class (whether it is arriving late, falling asleep, skipping class, or engaging in distracting behavior) will be disciplined by the officer in charge of their 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

The causal relation of interest for investigating the effect of laptop or tablet usage on academic performance is captured by ρ in the equation below:

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

Y_{ijht}, is the final exam score of student *i* who had professor *j* during class-hour *h* and semester *t*. D_{ijht} is an indicator variable that equals 1 if student *i* uses a laptop or a tablet and equals 0 otherwise. X_i is a vector of individual controls, the term α_{jt} includes fixed effects for each combination of professor and semester, β_{ht} includes fixed effects for each combination of class-hour and semester, and ε_i is the error term.[12]

If laptop or tablet usage was randomly assigned, ordinary least squares (OLS) estimates of ρ would capture the causal effect of computer usage on test scores. However, because students individually choose

[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.

[12] A richer estimating equation would include a fixed effect for each combination of professor, class-hour, and semester, thus ensuring that any variation in computer or tablet usage comes from students in the same classroom. We choose the less rich specification to match our instrumental variables estimates described later.

whether to use a laptop or tablet in the classroom, OLS estimates are likely to be biased by unobservable factors that are correlated with both computer usage and test scores. We therefore use assignment to a classroom that allows laptops or tablets as an instrumental variable (IV) for actual computer usage, thus exploiting the random assignment of classrooms to the two treatment arms described above.

The reduced-form equation for IV estimates is:

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

The coefficient π captures the impact of being assigned to a laptop computer- and/or tablet-permitted classroom (Z_{jht}) on academic performance. The inclusion of professor by semester fixed effects and class hour by semester fixed effects ensures that equation (2) identifies the impact of being assigned to a computer classroom by comparing students within the same semester while also controlling for unobserved mean differences in academic performance across professors and across class-hours.[13]

For IV estimates of ρ from equation (1) to be interpreted as causal, we must assume that assignment to a classroom that allows laptops and/or tablets only influences academic performance through a student's propensity to use a computer or tablet. This exclusion restriction might be violated if, for example, the computer usage by a student's peers provides a strong enough distraction to influence her own performance. Under the reasonable assumption that the potential distraction resulting from a neighbor's computer usage is weakly less deleterious than the effect of a student's personal computer usage on his own academic performance, we can interpret reduced form estimates of π in equation (2) and IV estimates of ρ in equation (2) as lower and upper-bounds, respectively, of the causal effect of a student's computer usage on his own academic performance.[14] Furthermore, we offer that any distraction effect in the West Point classroom is

[13] 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, leadership 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.

[14] Aguilar-Roca, et al (2012) randomly assign students into classrooms with "laptop-free" seating zones in a large-enrollment biology course. Although they observe no impact of the seating arrangements on student performance, they find that computer users tend to perform worse than students who elected not to use a computer and that student

minimized by 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.

V. DATA, STUDENT CHARACTERISTICS, AND COVARIATE BALANCE

Our sample consists of sophomore students who 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.[15] We further exclude students enrolled in classrooms of professors who only taught one class of the principles course, since professor fixed effects in our estimating equation prevents any variation in treatment arms among such classrooms. We also omit students enrolled in classrooms of professors who chose not to participate in the experiment, resulting in a final sample of 696 students.[16]

Columns 1 through 3 of Table 2 report descriptive statistics for students assigned to the control group, where computers 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

survey responses generally indicated a positive impact of the zoning rules on student attitudes. Given the magnitude of our parameter estimates, we would offer that the "distraction effect" remains a prime candidate for further research.

[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.

[16] Two professors informed the authors of their intention to not participate prior to the randomization of classrooms to treatment arms.

comprising 1 in 3 students. Average composite ACT scores are between 28 and 29 and average baseline (pre-treatment) grade point averages (GPAs) are between 2.8 and 2.9 for all three groups.[17]

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 differences in baseline characteristics are equal to zero, reported at the bottom of the column, is 0.579, 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. Similarly, average ACT scores among students assigned to tablet-only classrooms are 0.7 points lower than students assigned to classrooms in the control group. While these contrasts between treatment arms are likely to be chance findings, we control for baseline characteristics in our analysis below to ensure that our estimates are not confounded by the differences.

[17] 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/.

We derive outcomes in this experiment from a mandatory final exam in an introductory economics course. The final exam consisted of a combination of multiple choice, short answer, and essay questions, which were mapped directly to learning objectives in the course textbook and syllabus.[18] Students had 210 minutes to complete the exam in an online testing platform, which required the students to use a computer to answer questions.[19] The testing software automatically graded all multiple choice and short answer questions, but professors manually scored all essay responses.[20] Among students in our sample, the mean score on the final exam was roughly 75 percent (186 out of 250 possible points), with a standard deviation of 8.4 percentage points. 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. Notably, nearly all students in our sample sat for the final exam. Only 14 of the 696 students who began the semester did not have final exam scores, implying an attrition rate of roughly two percent.[21]

One potential concern with using final exam scores as an outcome is the possibility for professor grading tendencies to create bias in student test results. For example, 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 towards

[18] The final exam accounts for 25 percent of the total course points (250 of 1000). 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.

[19] 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.

[20] 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.

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

students who use computers in the classroom or if a professor's degree of grading leniency is influenced by a student's performance in the classroom or on other parts of the exam, then professor grading procedures could be correlated with assignment to one of our treatment arms. We do not believe this is a concern for multiple choice questions, as the online test platform renders an automatic, binary grade for multiple choice responses.[22] However, this is a much larger concern for essay questions, which are graded by professors who are able to observe a student's performance on previous questions of the exam prior to assigning a grade for each essay.

To further investigate this concern, Appendix Table 1 compares the percentage of variation in test scores explained by professor fixed effects (a partial R-squared of professor fixed effects) among multiple choice, short answer, and essay questions. Column 1 of each panel reports estimates of equation (2) 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 (0.483-0.454=0.029). Similarly, professor fixed effects explain less than 5 percent of the variation in short answer test scores, perhaps because all short answer questions are initially computer-graded and professors only have limited opportunities to offer partial credit.[23] 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 with multiple choice and short answer scores. Taken together, the evidence in Appendix Table 1 indicates

[22] In other words, the student answer for a multiple choice question is either fully correct or fully incorrect. A correct answer receives full credit, while an incorrect answer receives no credit.

[23] Additionally, short answer questions are either structured for standardized partial credit (e.g., a 3-part question is assigned a total of 6 points), or the course director provides guidance to standardize the assignment of partial credit (e.g., in a 7 point question, 4 points are associated with the appropriate shift of a demand curve).

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.

VI. RESULTS

We begin our analysis by comparing students in classrooms assigned to either treatment arm to students assigned to classrooms where laptop and tablet computers are prohibited. Column 1 of Table 3 reports first stage estimates analogous to equation (2) where the outcome (the endogenous variable) is an indicator that equals 1 if a student is ever observed using a laptop or a tablet and is 0 otherwise.[24] Here Z_{jht} (the instrument) indicates if the classroom is assigned to either treatment arm. For simplicity, we exclude all individual level covariates, X_i, which we add to subsequent two-stage least squares (2SLS) estimates. Just over 60 percent of students assigned to classrooms that permit computers or tablets actually use a computing device, as evidenced by the precisely estimated first stage of 0.62.

On average, exam scores among students in classrooms that permit laptops and tablets (treatment groups 1 and 2) are 0.21 standard deviations (hereafter σ) below the exam scores of students in classrooms that prohibit computers (control group). This can be seen by the reduced form, or intention-to-treat, estimates of equation (2) that are displayed in column 2 of panel A. 2SLS estimates, reported in column 3, indicate that using a laptop or tablet reduces exam performance by 0.34σ. The inclusion of demographic controls, ACT scores, and baseline GPAs has only a modest impact on 2SLS estimates, as seen in columns 4 and 5.[25] The statistically significant point estimate in column 5 of panel A suggests that computer usage lowers exam performance by roughly one-fourth of a standard deviation. By way of comparison, column 6

[24] We asked professors to record computer and tablet usage on three separate occasions during each semester. We chose to define the endogenous variable as ever being observed using a computer because some students were absent when their professors recorded computer usage. A few professors also only reported computer usage twice during a semester.

[25] The full set of controls for the regression estimates reported in column 5 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.

of Table 3 reports OLS estimates of equation (1). The point estimate of -0.15σ is smaller in magnitude than the corresponding 2SLS estimates, suggesting positive selection into computer usage (students who would have performed better on exams are more likely to use computers). However, this comparison should be made with caution, as 2SLS estimates are not statistically distinguishable from OLS estimates.

Subsequent panels of Table 3 report estimates for multiple choice scores, short answer scores, and essay scores. 2SLS estimates for multiple choice and short answer scores are similar, with both suggesting that laptop or tablet usage decreases academic performance by one-fourth of a standard deviation. In contrast, 2SLS estimates for essay scores, as seen in column 5 of Panel B, are close to 0 and significantly smaller in magnitude than multiple choice and short answer estimates. Our finding of a 0 effect for conceptual questions stands in contrast to previous research by Mueller & 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.

Since our instrument varies at the classroom level, it would normally be appropriate to cluster standard errors at this level as well. However, inference based on robust standard errors is more conservative than inference based on clustered standard errors, so we choose to report the former. To see this more clearly, Appendix Table 2 compares robust, conventional, and clustered standard error estimates for the reduced form specification described by equation (2). Robust and conventional standard errors are nearly identical, but, surprisingly, clustered standard errors are substantially smaller than robust standard errors. With only 48 classrooms in the experiment, clustered standard errors could potentially be 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 multiple choice questions, 0.002 for short answer questions, and 0.009 for essay 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 reduced form estimates of equation (2) using the two-step grouped-data estimation procedure for models with microcovariates, as described by Angrist and Pischke (2009).[26] In the first step of this procedure, we construct covariate-adjusted classroom effects by estimating:

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

In the second step, we regress the estimated classroom effects, $\widehat{\mu_{jht}}$, on classroom-level variables, with each observation (i.e. each classroom) weighted by classroom size:

(4) $$\widehat{\mu_{jht}} = \kappa_{jt} + \lambda_{ht} + \pi Z_{jht} + \epsilon_{jht}$$

Column 4 of Appendix Table 2 reports standard errors of π using this two-step method, with P-values based on inference from a t-distribution with 26 degrees of freedom.[27] Even this conservative method of inference indicates that the effect on our preferred outcome (multiple choice + short answer) is significant at the 1 percent level, while the effects on multiple choice and short answer scores are significant at the 1 and 5 percent levels, respectively.

Distinguishing Between Treatment Arms

We next explore whether the reduction in academic performance is concentrated in either classrooms that permit unrestricted computer usage or classrooms that permit only modified-tablet usage. Table 4 reports results from a comparison of students in classrooms where laptops and tablets are permitted without restriction (treatment group 1) to students in classrooms where computers are prohibited. The large, precisely estimated first stage of 0.80 indicates that computer usage is nearly ubiquitous in classes where computers are permitted. The 2SLS estimate with the full set of controls for our preferred outcome, seen in column 5 of Panel A, suggests that computer usage reduces academic performance by 0.22σ. It is worth noting that including demographic, ACT, and baseline GPA controls attenuates our 2SLS estimates from

[26] See *Mostly Harmless Econometrics*, pp. 313-314.

[27] This follows the suggestion of Donald and Lang (2007). With 48 classrooms, 15 combinations of professor and semester, and 8 combinations of class hour and semester, the residual degrees of freedom is 48-14-7-1=26.

0.35σ to 0.22σ (compare column 3 to column 5 of panel A). This is likely due to random differences in the composition of students between the first treatment arm and the control group. Importantly, however, the estimates reported in columns 3, 4, and 5 are statistically indistinguishable. The other results in Table 4 indicate that unrestricted computer usage reduces both multiple choice and short answer scores by roughly one-fifth of a standard deviation, but this treatment has no effect on essay scores, consistent with our results from Table 3. OLS estimates are also similar to 2SLS estimates, which is unsurprising considering the large first stage estimates.

Table 5 reports results from restricting our sample to students in classrooms where modified-tablet (must remain flat on desk) use is permitted (treatment group 2) to students in classrooms where computers are prohibited. For this analysis, assignment to a classroom that permits modified-tablet usage serves as an instrument for actual tablet usage. The first stage estimate of 0.38 is about half the size of the first stage estimate reported in Table 4, suggesting that requiring students to keep tablets face-up and flat on their desks substantially reduces technology usage in the classroom. The reduced form estimates reported in Table 5 are roughly two-thirds the magnitude of reduced form estimates reported in Table 4, but when scaled by their corresponding first stage estimates, the results in Table 5 actually suggest that modified-tablet usage has a more negative impact on academic performance than unrestricted computer usage. With the full set of controls, the point estimate in panel A of column 5 indicates that modified-tablet usage reduces academic performance by 0.41σ, which is almost double the effect of unrestricted computer use on academic performance (compare to column 5, panel A of Table 4). 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. It is also worth mentioning that 2SLS estimates are substantially more negative than OLS estimates, suggesting that students who choose to use tablets would otherwise perform better than students who choose not to use tablets.[28] Again, however, this observation is far from conclusive given the relative imprecision of 2SLS estimates.

Effects By Subgroups

Table 6 explores whether treatment effects vary by subgroups. The table reports 2SLS estimates of equation (1), where assignment to a classroom that allows unrestricted computer or modified tablet usage acts as an instrument for actual computer usage. The first two columns of panel A split our sample by gender while the next two columns split the sample by race. Interestingly, computer or modified tablet usage reduces male academic performance by 0.33σ but appears to have little effect on the academic performance of women. Even with the relatively small number of women in our sample, we are able to rule out equal effects between men and women. In contrast to the evidence of heterogeneous treatment effects for men and women, the estimates reported in columns 3 and 4 of panel A reveal similar effects for whites and nonwhites.

We also observe some evidence that computer usage is most harmful to students with relatively strong baseline academic performance. This can be seen in panel B of Table 6, which report 2SLS estimates for students who fall within the lower, middle, and upper-third of the distribution of baseline GPAs. Computer usage appears to lower the exam scores for students in the upper-tercile of baseline GPAs by 0.40σ. However, it only lowers exam scores by 0.17σ for students in the lowest tercile of baseline GPAs. Panel C reveals a similar pattern: computer usage has a small, statistically insignificant effect on students with relatively low ACT scores, but it reduces exam performance by 0.36σ for students within the middle and upper-tercile of the ACT distribution. A potential explanation for the relatively modest effects for students with low baseline GPAs and low ACT scores might be that students with relatively low baseline academic achievement find the course curriculum challenging, regardless of available enablers or

[28] However, 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.

distractions. Alternatively, these students may come from disadvantaged educational or home environments, leading to a lack of familiarity with distracting web applications and/or no pre-disposition for using technology in the classroom.

To further investigate whether computer and tablet usage is most harmful for students who would otherwise perform well in the absence of treatment, we show how treatment effects vary along a single index of expected test performance. We compute this index using leave-out fitted values to minimize the possibility that outcomes will be mechanically correlated with predicted performance, as suggested by Abadie, Chingos, and West (2013). More specifically, we use students in the control group to fit the following regression:

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

for test score Y_k and individual covariates X_k (gender, race, age, prior military service, Division I athlete, baseline GPA, and composite ACT score) in leave-out samples that omit each observation i. We then construct predicted exam scores using the following leave-out fitted values for all students:[29]

(6) $$\widehat{Y_i} = \hat{\beta}'_{(-i)} X_i.$$

Panel D of Table 6 reports results from this exercise for students within the lower, middle, and upper-third of the distribution of $\widehat{Y_i}$. Consistent with the results in panels B and C, the negative effect of computer or tablet usage appear to be strongest for students who have the highest predicted exam scores in the absence of treatment, although the point estimates in all three columns are statistically indistinguishable.

VII. CONCLUSION

The results from our randomized experiment suggest that computer devices have a substantial negative effect on academic performance. 2SLS estimates indicate that using a laptop or tablet device reduces final exam scores by 0.27 standard deviations. Comparing each treatment arm to the control group

[29] 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 computer or modified-tablet classrooms.

separately, we estimate that unrestricted laptop or tablet usage reduces test scores by 0.22 standard deviations while modified tablet usage reduces test scores by 0.41 standard deviations.

Although the estimated effects of our two treatments are statistically indistinguishable, our results suggest that 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. One explanation for this finding may be a greater propensity to access distracting web applications via the tablet computer, despite the professor's ability to monitor usage. The tablet computers used in this experiment use a mobile device operating system, which allows for cloud access to web applications typically used on smart phones. 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.

To put our results in perspective, Bandiera, Larcinese, and Rasul (2010) use a natural experiment to study the effect of class size changes in the UK on postgraduate students (equivalent to college seniors in the U.S). Results show that movement from an average class size to a class one standard deviation bigger would result in a 0.11 standard deviation reduction in test score.[30] Stinebrickner and Stinebrickner (2008) use an IV approach to study the effects of hours studying on grades. They find that an additional hour of studying increases the GPA by around 0.36 points. Angrist, Oreopoulos, and Williams (2014) measure the effect of a merit-based financial aid reward 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

[30] Their effects are concentrated in small to medium classes (below 20 to below 33).

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.

Technology is increasingly used in the classroom for experiments and teacher feedback (ex: clickers). We want to be clear that we cannot relate our results to a class where the laptop or tablet is used as a part of 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 or whether the increased availability of distractions for computer users (email, facebook, twitter, news, other classes, etc.) leads to lower grades. Given the magnitude of our results, additional research aimed at distinguishing between these channels is clearly warranted.

References

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

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.

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.

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.

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 from US News & World Report Education Rankings & Advice: http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/national-liberal-arts-colleges?int=a73d09

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.

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 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 30 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 (1)	Treatment 1 (Computers/Tablets) (2)	Treatment 2 (Tablets, face up) (3)	Both Treatments vs. Control (4)	Treatment 1 vs. Control (5)	Treatment 2 vs. Control (6)
Female	0.167	0.202	0.191	0.032	0.059	-0.001
				(0.030)	(0.036)	(0.040)
White	0.637	0.673	0.652	0.022	0.019	0.010
				(0.039)	(0.045)	(0.052)
Black	0.115	0.097	0.107	-0.024	-0.021	-0.010
				(0.026)	(0.029)	(0.036)
Hispanic	0.126	0.129	0.096	0.004	0.020	-0.030
				(0.028)	(0.033)	(0.035)
Age	20.12	20.15	20.15	0.03	0.05	0.07
	[1.06]	[1.00]	[0.97]	(0.08)	(0.09)	(0.11)
Prior Military Service	0.193	0.185	0.157	-0.017	-0.002	-0.011
				(0.031)	(0.036)	(0.040)
Division I Athlete	0.289	0.395	0.360	0.054	0.075*	0.033
				(0.036)	(0.044)	(0.047)
GPA at Baseline	2.87	2.82	2.88	-0.01	-0.05	0.02
	[0.52]	[0.54]	[0.51]	(0.04)	(0.05)	(0.05)
Composite ACT	28.78	28.30	28.23	-0.36	-0.37	-0.70**
	[3.21]	[3.46]	[3.28]	(0.26)	(0.31)	(0.34)
P Val (Joint χ^2 Test)				0.588	0.532	0.264
Observations	270	248	178	696	518	448

Notes: This table reports descriptive statistics of students in SS201 sections 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. All estimates 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. Results for Laptop and Modified-Tablet Classrooms vs. Non-Computer Classrooms

	First Stage	Reduced Form	2SLS			OLS
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Dependent Variable: Final Exam Multiple Choice and Short Answer Score					
Computer/Tablet Usage	0.62***	-0.21**	-0.34***	-0.32***	-0.27***	-0.15**
	(0.03)	(0.08)	(0.13)	(0.11)	(0.09)	(0.06)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	682	682	682	682	682	682
	B. Dependent Variable: Final Exam Multiple Choice Score					
Computer/Tablet Usage	0.62***	-0.18**	-0.29**	-0.27**	-0.23**	-0.13**
	(0.03)	(0.08)	(0.13)	(0.11)	(0.09)	(0.06)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	682	682	682	682	682	682
	C. Dependent Variable: Final Exam Short Answer Score					
Computer/Tablet Usage	0.62***	-0.20**	-0.33***	-0.32***	-0.28***	-0.15**
	(0.03)	(0.08)	(0.13)	(0.12)	(0.10)	(0.06)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	682	682	682	682	682	682
	D. Dependent Variable: Final Exam Essay Questions Score					
Computer/Tablet Usage	0.62***	0.00	-0.01	0.00	0.02	-0.07
	(0.03)	(0.07)	(0.11)	(0.10)	(0.09)	(0.06)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	682	682	682	682	682	682

Notes: This table reports 2SLS and OLS estimates of the effects of computer usage on academic performance. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. For 2SLS estimates, an indicator for being assigned to a classroom where laptop or tablet usage is allowed is an instrument for ever using a laptop or tablet (in either a modified or unrestricted manner) during the 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. Results for Laptop Classrooms vs. Non-Computer Classrooms

	First Stage	Reduced Form	2SLS			OLS
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Dependent Variable: Final Exam Multiple Choice and Short Answer Score					
Computer/Tablet Usage	0.80***	-0.28***	-0.35***	-0.28***	-0.22***	-0.20***
	(0.03)	(0.10)	(0.12)	(0.10)	(0.08)	(0.07)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	507	507	507	507	507	507
	B. Dependent Variable: Final Exam Multiple Choice Score					
Computer/Tablet Usage	0.80***	-0.25**	-0.31***	-0.24**	-0.19**	-0.18**
	(0.03)	(0.10)	(0.12)	(0.11)	(0.09)	(0.08)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	507	507	507	507	507	507
	C. Dependent Variable: Final Exam Short Answer Score					
Computer/Tablet Usage	0.80***	-0.25***	-0.31***	-0.26**	-0.21**	-0.17**
	(0.03)	(0.09)	(0.11)	(0.11)	(0.09)	(0.08)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	507	507	507	507	507	507
	D. Dependent Variable: Final Exam Essay Questions Score					
Computer/Tablet Usage	0.80***	-0.06	-0.07	-0.04	0.00	-0.05
	(0.03)	(0.08)	(0.10)	(0.10)	(0.09)	(0.07)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	507	507	507	507	507	507

Notes: This table reports 2SLS and OLS estimates of the effects of laptop or tablet usage on academic performance. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. For 2SLS estimates, an indicator for being assigned to a classroom that permits laptops and unrestricted tablet usage is an instrument for ever using a laptop or a tablet during the 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. Results for Modified-Tablet Classrooms vs. Non-Computer Classrooms

	First Stage	Reduced Form	2SLS			OLS
	(1)	(2)	(3)	(4)	(5)	(6)
	A. Dependent Variable: Final Exam Multiple Choice and Short Answer Score					
Computer/Tablet Usage	0.38***	-0.18*	-0.47*	-0.48*	-0.41**	-0.15
	(0.04)	(0.10)	(0.27)	(0.24)	(0.19)	(0.10)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	437	437	437	437	437	437
	B. Dependent Variable: Final Exam Multiple Choice Score					
Computer/Tablet Usage	0.38***	-0.16	-0.43	-0.43*	-0.37*	-0.15
	(0.04)	(0.10)	(0.28)	(0.25)	(0.20)	(0.11)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	437	437	437	437	437	437
	C. Dependent Variable: Final Exam Short Answer Score					
Computer/Tablet Usage	0.38***	-0.17*	-0.46*	-0.48*	-0.40*	-0.12
	(0.04)	(0.10)	(0.27)	(0.25)	(0.21)	(0.11)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	437	437	437	437	437	437
	D. Dependent Variable: Final Exam Essay Questions Score					
Computer/Tablet Usage	0.38***	-0.04	-0.11	-0.10	-0.10	-0.19*
	(0.04)	(0.08)	(0.22)	(0.20)	(0.18)	(0.10)
Demographics				X	X	X
ACT and baseline GPA					X	X
Observations	437	437	437	437	437	437

Notes: This table reports 2SLS and OLS estimates of the effects of modified tablet usage on academic performance. All scores have been standardized to have a mean of 0 and a standard deviation of 1 for each semester. For 2SLS estimates, an indicator for being assigned to a classroom where modified-tablet usage is allowed is an instrument for ever using a tablet during the 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. 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)
Computer/Tablet Usage	0.06	-0.33***	-0.30**	-0.31***
	(0.18)	(0.10)	(0.14)	(0.11)
Observations	125	557	236	446

	B. By Baseline GPA		
	Bottom Third of Distribution	Middle Third of Distribution	Top Third of Distribution
	(1)	(2)	(3)
Computer/Tablet Usage	-0.17	-0.27*	-0.40***
	(0.19)	(0.16)	(0.14)
Observations	226	228	228

	C. By ACT Score		
	Bottom Third of Distribution	Middle Third of Distribution	Top Third of Distribution
	(1)	(2)	(3)
Computer/Tablet Usage	-0.07	-0.36**	-0.36**
	(0.14)	(0.16)	(0.16)
Observations	262	210	210

	D. By Predicted Exam Score		
	Bottom Third of Distribution	Middle Third of Distribution	Top Third of Distribution
	(1)	(2)	(3)
Computer/Tablet Usage	-0.08	-0.26	-0.35**
	(0.14)	(0.17)	(0.15)
Observations	226	224	232

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. For 2SLS estimates, 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. (2014). 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)
Computer/Tablet Class	-0.14**	-0.12*	-0.17***	-0.18**	0.01	0.04
	(0.06)	(0.06)	(0.06)	(0.09)	(0.06)	(0.18)
GPA at baseline	0.89***	0.89***	0.97***	0.94***	0.70***	0.60***
	(0.07)	(0.08)	(0.07)	(0.06)	(0.07)	(0.09)
Composite ACT	0.057***	0.052***	0.046***	0.047***	0.033***	0.041***
	(0.012)	(0.012)	(0.012)	(0.013)	(0.010)	(0.015)
Instructor Fixed Effects	X		X		X	
R^2	0.483	0.454	0.446	0.401	0.529	0.209
Observations	682	682	682	682	682	682

Notes: This table reports reduced form estimates of the effects of being assigned to a computer class or to a modified tablet class on academic performance 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, linear terms for baseline GPA, CEER score, baseline age, and indicators for female, white, black, hispanic, prior military service, and athlete. 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. 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	-0.170***	-0.170***	-0.170***	-0.168***
	(0.056)	(0.057)	(0.036)	(0.051)
P-Value	0.0027	0.0028	0.0000	0.0027
Clusters (classrooms)			48	
Residual Deg-of-Freedom	651	651	47	26
Observations	682	682	682	48
	B. DV: Final Exam Multiple Choice Score			
Computer or Modified Tablet Classroom	-0.140**	-0.140**	-0.140***	-0.140***
	(0.060)	(0.061)	(0.034)	(0.050)
P-Value	0.0197	0.0207	0.0002	0.0090
Clusters (classrooms)			48	
Residual Deg-of-Freedom	651	651	47	26
Observations	682	682	682	48
	C. DV: Final Exam Short Answer Score			
Computer or Modified Tablet Classroom	-0.173***	-0.173***	-0.173***	-0.171**
	(0.062)	(0.063)	(0.058)	(0.081)
P-Value	0.0052	0.0058	0.0046	0.0455
Clusters (classrooms)			48	
Residual Deg-of-Freedom	651	651	47	26
Observations	682	682	682	48
	D. DV: Final Exam Essay Questions Score			
Computer or Modified Tablet Classroom	0.014	0.014	0.014	0.014
	(0.057)	(0.058)	(0.063)	(0.084)
P-Value	0.8024	0.8040	0.8206	0.8721
Clusters (classrooms)			48	
Residual Deg-of-Freedom	651	651	47	26
Observations	682	682	682	48

Notes: This table reports reduced form estimates of the effects of being assigned to a computer or modified tablet classroom on exam performance. 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, CEER score, 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 computer 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.