Economics of Education Review 56 (2017) 118–132

Contents lists available at ScienceDirect

Economics of Education Review

journal homepage: www.elsevier.com/locate/econedurev

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 S. Walker

United States Military Academy, 607 Cullum Road, West Point, NY 10996, USA

A R T I C L E I N F O

Article history:
Received 5 July 2016
Revised 8 December 2016
Accepted 9 December 2016
Available online 14 December 2016

A B S T R A C T

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 0.18 standard deviations 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.

Published by Elsevier Ltd.

1. Introduction

Internet-enabled classroom technology is nearly universal at all levels of education in the United States. Between 1994–2005, the percentage of U.S. public school classrooms with Internet access increased from 3 to 94%, 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 remains 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% 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] Moreover, this proliferation of technology and Internet access for educational purposes is not confined to the U.S. Reports show that high school students in many European countries have the same or higher computer access as students in the U.S.[3]

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.[4]

As institutions, including the one in this study,[5] push for faster and continuous access to wireless Internet to support the proliferation of web-enabled educational resources, it is unclear how permitting computers in the classroom impacts student performance. Many Internet-enabled innovations may enhance the learning environment, including group activities and providing immediate feedback to instructors through question responses. In a traditional classroom, where computers and tablets are only used to take notes, benefits may include the ability to take notes faster, store copies of notes, and carry the notes with you at all times without carrying various notebooks. Consistent with this, survey evidence

☆ 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 Josh Angrist, David Autor, Daphna Bassok, Rekha Balu, Perry Bolding, Bill Skimmyhorn, and the USMA economics program, especially those instructors participating in our study.

* Corresponding author.
E-mail addresses: susan.carter@usma.edu (S.P. Carter), kyle.greenberg@usma.edu (K. Greenberg), michael.walker2@usma.edu (M.S. Walker).

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

[2] See U.S. Department of Education (2016), page 6.

[3] See Bulman and Fairlie (2016) for a review of the growth of Internet-access and computer technology in both the U.S. and various developed and developing countries.

[4] 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.

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

http://dx.doi.org/10.1016/j.econedurev.2016.12.005
0272-7757/Published by Elsevier Ltd.

from 2006 suggests that students enjoy having computers in the classroom (Barak, Lipson, & Lerman, 2006). Outside of the classroom, increased connectivity on college campuses provides opportunities for students and teachers to collaborate outside of the classroom. Personal computers provide convenient options for student research via university library-enabled online search engines. Electronic textbooks also allow students to easily carry their entire curriculum with them at all times. "Enhanced" textbooks offer students the capability to watch embedded videos, follow hyperlinks to pertinent articles on the Internet.[6]

Despite the potential benefits of computers in the classroom, other evidence suggests that permitting computers negatively impacts student learning. Several studies suggest that potential distractions from web-surfing, e-mail checking, and electronic chatting with friends have negative effects. Fried (2008) finds that students report increased multitasking when laptops are in the classroom. Kraushaar and Novak (2010) and Grace-Martin and Gay (2001) monitor activity usage and find negative correlation between non-class related activity on their computer and performance. Relatedly, multiple laboratory-style studies demonstrate negative effects of laptop multi-tasking on test performance (e.g., Hembrooke & Gay, 2003; Sana, Weston, & Cepeda, 2013). While distractions and multi-tasking are one potential channel through which computers may negatively impact performance, another potential channel is that students recall less information when they are required to take notes with computers rather than by hand, as suggested in a lab experiment by Mueller and Oppenheimer (2014).[7]

This paper contributes to the debate surrounding the impact of computers in the classroom through an experiment that randomly allowed students to access their laptop and tablet 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.

Our experiment expands upon the previous research in a number of dimensions. In contrast to laboratory studies that demonstrate a negative impact of computers on learning (e.g., Hembrooke & Gay, 2003; Mueller & Oppenheimer, 2014; Sana et al., 2013), 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, some of the negative effects found in lab experiments could partially be attributed to experimental designs that require 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).[8] Our research design intentionally seeks to limit the influence of such artificial behaviors by not requiring students to use computers. 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, an investigation of the effects of technology in the context of an actual course is an important extension of laboratory research.

Our randomized control trial design also allows us to improve upon previous research done in real classrooms as well. First, we are able to control for selection into computer usage and avoid the problems associated with students self-reporting 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.

The results of our study suggest that permitting computing devices in the classroom reduces final exam scores by 0.18 standard deviations. 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. Our results also indicate that the negative impact of computers occurs in classrooms that permit laptops and tablets without restriction and in classrooms that only permit modified-tablet usage.

The results of our study are consistent with other recent research on the impact of technology on classroom performance that is causal in nature. Beland and Murphy (2016) exploit variation in school mobile phone policies to find that banning mobile phones is associated with a 0.07 standard deviation increase in exam scores among UK high school students. Elsewhere, Patterson and Patterson (2016) instrument for computer usage in classes that allow laptops with the laptop policies from students' other classes during the day and find that computer usage reduces academic performance among undergraduate students at a private liberal arts college. Our experiment complements these studies by directly investigating the potential impact of a teacher's decision to permit or restrict laptops and tablets in their classrooms.[9]

Section 2 of this paper provides background on West Point for the purposes of generalization, and Section 3 discusses our experimental design. Sections 4 and 5 discuss our empirical framework, sample selection, data, and evidence of successful random assignment. Section 6 presents the results of our regression analysis, Section 7 discusses additional robustness checks, and Section 8 concludes.

2. Background on west point

The United States Military Academy at West Point, NY, is a 4-year undergraduate institution with an enrollment of approximately 4400 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.

[6] 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 assignments and completion rates. 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.

[7] Other studies report lower satisfaction for students with laptops (e.g., Wurst, et al., 2008) and lower student-reported understanding of material (Fried, 2008). These could potentially be the result of bot.

[8] 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.

[9] In fact, anecdotal evidence suggests that professors and teachers are increasingly banning laptop computers, smart phones, and tablets from their classrooms. See, for example, Gross (2014), "This year, I resolve to ban laptops from my classroom," *Washington Post*, available from https://www.washingtonpost.com.

Table 1
Comparison of West Point 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)	Davidson College, NC (Ranked 9)	Washington and Lee, VA (Ranked 14)	Colorado College, CO (Ranked 25)	United States Military Academy	All	Pop between 1000 & 10,000	All	Pop between 1000 & 10,000
Full-time degree seeking undergrads										
Undergraduate population	4414	2014	1765	1876	2036	4591	9215	4645	2841	3213
student to faculty ratio	7:1	7:1	10:1	8:1	10:1					
% Female	17%	51%	51%	50%	53%	17%	56%	56%	56%	58%
Non-resident aliens	1%	7%	6%	4%	6%	1%	3%	2%	3%	3%
hispanic	11%	12%	7%	4%	9%	10%	12%	10%	12%	12%
Black / AA, non-hispanic	9%	7%	6%	2%	2%	8%	14%	17%	16%	15%
White, non-hispanic	67%	56%	69%	83%	66%	69%	59%	59%	54%	56%
American indian	1%	0%	1%	0%	0%	1%	2%	2%	1%	1%
Asian, non-hispanic	6%	11%	6%	3%	5%	6%	4%	3%	4%	3%
Pacific islander	0%	0%	0%	0%	0%	1%	1%	1%	0%	0%
Two or more races	3%	7%	4%	2%	8%	4%	3%	3%	2%	2%
Race unknown	2%	0%	2%	2%	3%	1%	3%	4%	7%	7%
% from out of state	93%	88%	77%	86%	82%					
Freshman profile										
ACT composite										
25th Perc	26	31	28	30	28	27	20	19	20	21
75th Perc	31	34	32	33	32	30	25	24	25	26
SAT critical reading										
25th Perc	570	680	610	660	620	580	459	442	468	471
75th Perc	690	790	720	730	730	695	565	544	577	578
SAT Math										
25th Perc	590	670	620	660	630	600	474	453	477	480
75th Perc	700	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.

West Point's student composition is unique primarily due 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; however 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), to complete required military education courses, and to take a rigorous academic course load. These requirements likely lead to a student body that is more 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.).[10] 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.[11] In panel A, we show gender, race, and home location breakdowns for West Point relative to four 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 students accounting for only 17% 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 1000 and 10,000, all 4-year schools (including private non-profit and private for-profit), and all 4-year schools with a population between 1000 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.

3. 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, as depicted in Fig. 1. 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 lap-

[10] Nominations may also occur if someone has prior military service, has a parent with military service, or from the U.S. Vice President and/or the Secretary of the Army.

[11] See U.S. News and World Report (2016), available at http://colleges.usnews.rankingsandreviews.com/best-colleges, accessed 29 April 2016.

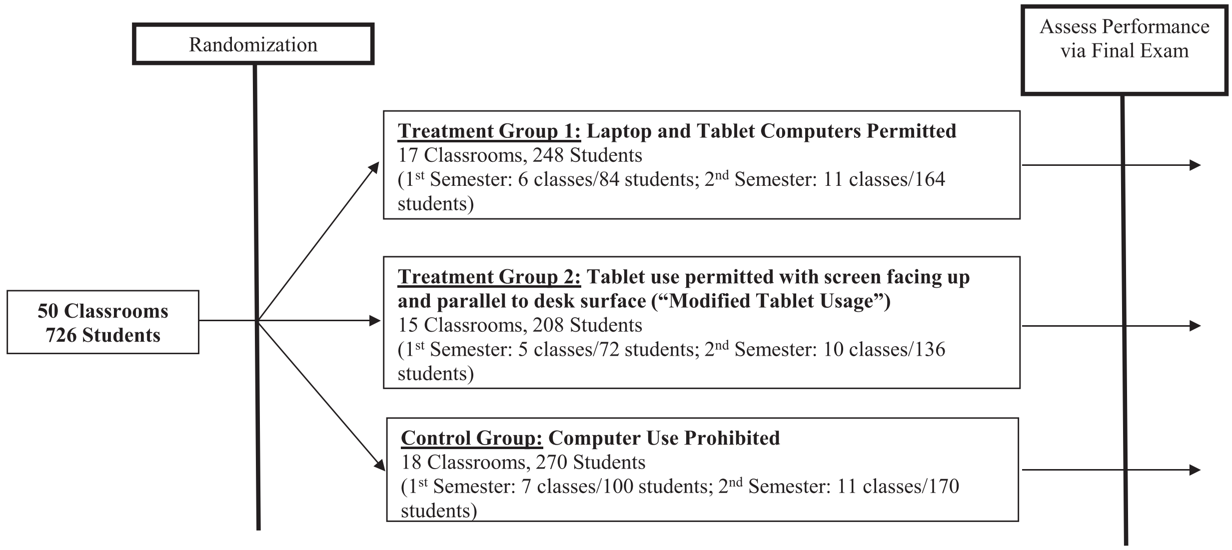

Fig. 1. Experimental design.

tops 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 up to 30 different sections of the course per semester, 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.[12]

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. 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's rank in the graduating class, the greater the student's chance of receiving his or 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.[13] 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.

4. Empirical framework

To compare outcomes between students assigned to classrooms that permitted laptop or tablet usage and students assigned to classrooms that prohibited computer usage, we estimate the following model of undergraduate academic achievement:

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

[12] It is important to note that West Point professors do not have teaching assistants. West Point policy also forbids students from using mobile phones during any period of instruction.

[13] 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.

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-professor fixed effects and semester-class fixed effects, we compare students within the same semester while also controlling for unobserved mean differences in academic performance across professors and across class-hours.[14] As laptops 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.

Since the treatment in this experiment varies at the classroom level, it would typically be appropriate to cluster standard errors on classrooms. Considering both the level of experimental variation and the relatively small number of clusters in this experiment, we instead report robust standard errors as well as wild-bootstrap p-values using the finite-cluster correction technique discussed in Cameron, Gelbach, and Miller (2008) in our main tables. In the appendix, we compare our results with clustered standard errors and confirm that our conclusions are robust to clustering. In fact, inference based on robust standard errors is usually more conservative than inference based on clustered standard errors, even after incorporating asymptotic refinements. We discuss this in more detail and explore alternative standard error estimates in Section 7.

Estimates of Eq. (1) capture the effect of being in a class that allows technology, not the causal effect of actually using a computer. We do not to use Z_{jht} from Eq. (1) as an instrument for actual computer usage because we are concerned that computer usage by a student's peers may provide a strong enough distraction to influence her own performance.[15] Alternatively, teachers may respond differently to classrooms where students use computers. Both scenarios would violate the exclusion restriction required for instrumental variables estimates to be consistent. Systematic under-reporting (or over-reporting) of computer use would also cause instrumental variable estimates to be biased upwards (or downwards).

5. Data & descriptive statistics

5.1. Sample selection, student characteristics, & covariate balance

Our sample consists of students enrolled in West Point's Principles of Economics course during either 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 students enrolled in classrooms of professors who chose not to participate in the experiment, resulting in a final sample of 50 classrooms and 726 students (see Fig. 1).[16]

Columns 1 through 3 of panel A, Table 2 report baseline 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 together 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.[17]

Subsequent columns of panel A, 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 0.61, 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 (panel A) 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 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 other differences.

5.2. Computer usage in the treatment and control groups

We asked professors to record which students used computers on three separate occasions during the semester. Panel B of Table 2 compares observed computer usage in the control group to observed usage in both treatment groups. Over 80% of students assigned to classrooms that permitted laptops or tablets without restriction used a computing device during at least one of these three classes during the semester, which we henceforth define as any computer usage. In contrast, just under 40% of students in modified-tablet classrooms used a tablet at some point during the

[14] Each student only has one observation in the data for the analysis that follows. The within semester comparison is critical for two reasons. First, the students participating in the experiment spanned two separate class years, which may have been subject to different admissions policies and/or admissions personnel. Second, professors in charge of the 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.

[15] Empirical evidence of a "distraction effect" is mixed. Aguilar-Roca et al. (2012) randomly assign students to classrooms with "laptop-free" seating zones and observe no impact of the seating arrangements on student 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 able to view peer 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.

[16] 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.

[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/.

Table 2
Summary statistics and covariate balance.

	Control (1)	Treatment 1 (laptops/tablets) (2)	Treatment 2 (tablets, face up) (3)	Both treatments vs. control (4)	Treatment 1 vs. control (5)	Treatment 2 vs. control (6)
A. Baseline characteristics						
Female	0.17	0.20	0.19	0.03	0.06	0.00
				(0.03)	(0.04)	(0.04)
White	0.64	0.67	0.66	0.02	0.02	0.02
				(0.04)	(0.04)	(0.05)
Black	0.11	0.10	0.11	−0.02	−0.02	−0.03
				(0.03)	(0.03)	(0.04)
Hispanic	0.13	0.13	0.09	0.00	0.02	−0.03
				(0.03)	(0.03)	(0.03)
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	0.19	0.19	0.16	−0.02	0.00	−0.01
				(0.03)	(0.04)	(0.04)
Division I athlete	0.29	0.40	0.35	0.05	0.07*	0.04
				(0.04)	(0.04)	(0.05)
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
B. Observed computer (laptop or tablet) use						
any computer use	0.00	0.81	0.39	0.62***	0.79***	0.40***
				(0.02)	(0.03)	(0.04)
Average computer use	0.00	0.57	0.22	0.42***	0.56***	0.24***
				(0.02)	(0.02)	(0.03)
Observations	270	248	208	726	518	478

Notes: Columns 1–3 of this table report mean characteristics of student in the control group (classrooms where laptops and tablets are prohibited), treatment group 1 (laptops and tablets permitted without restriction), and treatment group 2 (tablets are permitted if they are face up). Standard deviations are reported in brackets. Columns 4–6 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 4–6 include (instructor) x (semester) fixed effects and (class hour) x (semester) fixed effects. The reported *p*-values in Panel A are from a joint test of the null hypothesis that all coefficients are equal to zero. Observed computer usage, reported in panel B, was recorded during three lessons each semester of the experiment. Any computer use is an indicator variable for ever using a laptop or tablet during one of these three lessons. For example, a student who uses a computer during one of these three lessons has a value of one for any computer use and has an average usage rate of one-third. Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respectively.

semester and only one student (less than one percent) in the control group classrooms ever used a laptop or tablet.[18] These differences in computer usage across treatment arms are also evident when the outcome is average computer usage over the three days where professors recorded usage.[19] A comparison of the results in columns 5 and 6 of panel B suggests that requiring students to keep their tablets face-up on the desk substantially reduces the number of computing devices in the classroom. Although not reported in Table 2, *p*-values from tests of the hypothesis that observed computer use estimates are equal in both treatment arms are smaller than 0.001 using both the “any use” and the “average use” measures.

5.3. Measuring final exam performance

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.[20] Students had 210 min. to complete the exam in an online testing platform, which required the students to use a computer to answer questions.[21] The testing software automatically graded all multiple choice and short answer questions, but professors manually scored all essay responses.[22] 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. Attrition is not significantly correlated with assignment to either treatment group and the observable characteristics of the few

[18] 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.

[19] As an 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.

[20] 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.

[21] 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.

[22] 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.

students without exam scores are similar in both treatment groups and the control group.[23]

One concern with 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. Including professor fixed effects in our regression model accounts 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, 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.

The plots in Appendix Fig. A.1 indeed suggest that some instructors grade essay questions in a manner that ensures some of their students achieve a passing score. Panel A shows the distribution of overall exam scores, which combine multiple choice, short answer, and essay scores. There is a large mass of overall exam scores immediately to the right of 67%. The plot in panel B shows the distribution of final exam scores after removing the essay score and only including computer graded questions. Unlike panel A, there is no mass of grades to the right of 67%. These figures together suggest that instructors may grade essays in a manner that pushes some students over the passing threshold. If permitting computers reduces multiple choice and short answer scores, then it is possible that instructors will increase essay scores for students who were permitted access to computers as an endogenous response to this negative treatment effect.

To further investigate the possibility that grades reflect grader influence rather than academic achievement, Appendix Table A.1 compares the percentage of variation in test scores explained by professor fixed effects (the partial R^2 when adding professor fixed effects) for multiple choice, short answer, and essay questions. Column 1 of each panel reports estimates of Eq. (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^2 reported in columns 1 and 2 of panel A indicates that professor fixed effects explain 3% of the variation in multiple choice test scores ($0.48-0.45=0.03$). Similarly, professor fixed effects explain only 4% of the variation in short answer test scores. On the other hand, professor fixed effects explain 33% 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 Fig. A.1 and Appendix Table A.1 indicates that essay scores do not provide an accurate measurement of student achievement. 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.[24] For this particular outcome, the average score among all students in our sample was roughly 71.7% with a standard deviation of 9.2 percentage points.[25] The raw averages among students in the control group, treatment group 1, and treatment group 2 were 72.9%, 70.5%, and 71.4%, respectively. Throughout our 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.

6. Results

6.1. Effects of permitting computers on exam performance

A comparison of exam scores of students in either treatment arm to the scores of students assigned to classrooms where laptops and tablets were prohibited suggests that permitting computers negatively impacts scores. Panel A of Table 3 reports estimates of Eq. (1) where the outcome is the composite of a student's multiple choice and short answer scores, standardized to have a mean of zero and variance of 1. 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.[26] 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 scores. As seen in column 3 of panel A, the effect of being assigned to a classroom that permits computers is roughly 17% as large as the association between a one point reduction in baseline GPAs and final exam scores ($\frac{-0.19}{1.13}=0.17$). To put it another way, a student in a classroom that prohibits computers is on equal footing with her peer who is in a class that allows computers but who has a GPA that is one-third of a standard deviation higher than her GPA.[27]

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

[23] These results are available from the authors upon request.

[24] 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.

[25] Separately, the mean (standard deviation) for multiple choice, short answer, and essay scores is 73 (10), 70 (12), and 84 (12), respectively.

[26] 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.

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

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
R^2	0.05	0.24	0.52	0.54
Robust SE *P*-Val	0.010	0.005	0.001	0.002
Wild Bootstrap *P*-Val	0.000	0.000	0.000	0.000
B. Dependent variable: Final exam multiple choice score				
Laptop/tablet class	−0.18**	−0.17**	−0.16***	−0.15**
	(0.08)	(0.07)	(0.06)	(0.06)
Demographic controls		X	X	X
GPA control			X	X
ACT control				X
R^2	0.06	0.24	0.46	0.48
Robust SE *P*-Val	0.027	0.019	0.009	0.016
Wild Bootstrap *P*-Val	0.000	0.000	0.000	0.000
C. Dependent variable: Final exam short answer score				
Laptop/tablet class	−0.22***	−0.22***	−0.21***	−0.19***
	(0.08)	(0.07)	(0.06)	(0.06)
Demographic controls		X	X	X
GPA control			X	X
ACT control				X
R^2	0.06	0.18	0.42	0.43
Robust SE *P*-Val	0.007	0.004	0.001	0.002
Wild Bootstrap *P*-Val	0.006	0.016	0.000	0.008
D. Dependent variable: Final exam essay questions score				
Laptop/tablet class	0.02	0.02	0.03	0.03
	(0.07)	(0.06)	(0.06)	(0.06)
Demographic controls		X	X	X
GPA control			X	X
ACT control				X
R^2	0.33	0.38	0.50	0.51
Robust SE *P*-Val	0.785	0.766	0.642	0.548
Wild Bootstrap *P*-Val	0.757	0.775	0.627	0.509

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 estimates are from a sample of 711 students who took the final exam. All scores are standardized to have a mean of 0 and a standard deviation of 1 for each semester. All estimates include (instructor) x (semester) fixed effects and (class hour) x (semester) fixed effects. 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. The reported *P*-values are from the null hypothesis that the effect of being assigned to a classroom that permits laptops or tablets equals zero. Wild bootstrap *p*-values with classroom-level clusters are constructed from the procedure describe in Cameron et al. (2008). Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respectively.

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
R^2	0.08	0.28	0.54	0.57
Robust SE *P*-Val	0.003	0.007	0.005	0.005
Wild Bootstrap *P*-Val	0.000	0.000	0.000	0.000
B. Dependent variable: Final exam multiple choice score				
Computer class	−0.25***	−0.20**	−0.16**	−0.15**
	(0.10)	(0.009)	(0.07)	(0.07)
Demographic controls		X	X	X
GPA control			X	X
ACT control				X
R^2	0.08	0.27	0.48	0.50
Robust SE *P*-Val	0.009	0.023	0.025	0.029
Wild Bootstrap *P*-Val	0.000	0.000	0.000	0.000
C. Dependent variable: Final exam short answer score				
Computer class	−0.25***	−0.21**	−0.18**	−0.17**
	(0.09)	(0.09)	(0.07)	(0.07)
Demographic controls		X	X	X
GPA control			X	X
ACT control				X
R^2	0.08	0.21	0.44	0.46
Robust SE *P*-Val	0.008	0.016	0.017	0.019
Wild Bootstrap *P*-Val	0.008	0.020	0.022	0.028
D. Dependent variable: Final exam essay questions score				
Computer class	−0.03	−0.01	0.02	0.02
	(0.08)	(0.08)	(0.07)	(0.07)
Demographic controls		X	X	X
GPA control			X	X
ACT control				X
R^2	0.32	0.37	0.50	0.51
Robust SE *P*-Val	0.705	0.912	0.801	0.755
Wild Bootstrap *P*-Val	0.549	0.811	0.721	0.641

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 consists of 507 students who took the final exam and were not in modified-tablet classrooms. All scores are standardized to have a mean of 0 and a standard deviation of 1 for each semester. See the notes from Table 3 for a list of controls included in each regression. The reported *P*-values are from the null hypothesis that the effect of being assigned to a classroom that permits laptops or tablets equals zero. Wild bootstrap *p*-values with classroom-level clusters are constructed from the procedure describe in Cameron et al. (2008). Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respectively.

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.

6.2. 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 panel A of Table 4 from −0.28 to -0.18σ. This is due to random differences in the composition of students between the first treatment arm and the control group. Although concerning, there are a few reasons to believe that the treatment effect suggested by the estimates in panel A of Table 4 is not purely the result of unobservable differences between students in treatment group 1 and students in the control group. First, the estimates in all columns of panel A are statistically indistinguishable. Second, including individual level covariates increases R^2 values from 0.08 to 0.57. Applying a bounding technique suggested by Oster (2015), which assumes that any residual omitted variable bias in our estimate of the treatment effect is equal to the change in coefficient estimates when individual covariates are included multiplied by the ratio of residual variation in the outcome to the change in R^2 when controls are included, still suggests that permitting computers negatively impacts academic performance.[28] By comparison, Oster (2015; Table 4) finds that among papers published in top economics journals between 2008 and 2013 where including additional controls attenuates coefficient estimates towards 0, only 42% of results from randomized data and 20% of results from non-randomized data survive this bounding exercise. Third, we present additional results in Section 7.2 which suggest that students in the first treatment arm performed as well as or better than students in control classrooms on performance measures that were not likely to be influenced by computer usage.

Table 5 reports estimates of Eq. (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. 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. We caution, however, that the minimum detectable difference in effects between the unrestricted computer use treatment and the modified tablet treatment is 0.15σ. Even though the results in Tables 4 and 5 suggest that both unrestricted computer usage and modified-tablet usage negatively impact exam scores, we cannot rule out large differences in the true magnitude of each effect.

Table 5
Modified-tablet classrooms vs. non-computer classrooms.

	(1)	(2)	(3)	(4)
A. Dependent variable: Final exam multiple choice and short answer score				
Computer class	−0.17*	−0.18**	−0.20***	−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
R^2	0.07	0.26	0.53	0.54
Robust SE *P*-Val	0.087	0.050	0.007	0.019
Wild Bootstrap *P*-Val	0.000	0.000	0.000	0.000
B. Dependent variable: Final exam multiple choice score				
Computer class	−0.15	−0.15*	−0.17**	−0.14*
	(0.10)	(0.09)	(0.08)	(0.07)
Demographic controls		X	X	X
GPA control			X	X
ACT control				X
R^2	0.07	0.26	0.48	0.49
Robust SE *P*-Val	0.141	0.100	0.027	0.057
Wild Bootstrap *P*-Val	0.000	0.000	0.000	0.000
C. Dependent variable: Final exam short answer score				
Computer class	−0.21**	−0.22**	−0.24***	−0.21**
	(0.10)	(0.09)	(0.08)	(0.08)
Demographic controls		X	X	X
GPA control			X	X
ACT control				X
R^2	0.11	0.22	0.43	0.45
Robust SE *P*-Val	0.032	0.016	0.004	0.010
Wild Bootstrap *P*-Val	0.000	0.000	0.000	0.000
D. Dependent variable: Final exam essay questions score				
Computer class	−0.01	−0.01	−0.03	−0.02
	(0.08)	(0.08)	(0.07)	(0.07)
Demographic controls		X	X	X
GPA control			X	X
ACT control				X
R^2	0.37	0.41	0.54	0.54
Robust SE *P*-Val	0.882	0.853	0.682	0.742
Wild Bootstrap *P*-Val	0.687	0.727	0.318	0.426

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 consists of 466 students who took the final exam and were not in classrooms where laptops and tablets were permitted without restriction. All scores are standardized to have a mean of 0 and a standard deviation of 1 for each semester. See the notes from Table 3 for a list of controls included in each regression. The reported *P*-values are from the null hypothesis that the effect of being assigned to a classroom that permits laptops or tablets equals zero. Wild bootstrap *p*-values with classroom-level clusters are constructed from the procedure describe in Cameron et al. (2008). Robust standard errors are reported in parentheses. ***,**, and * denote significance at the 1%, 5%, and 10% level, respectively.

6.3. Effects by subgroup

Appendix Table A.2 explores whether treatment effects vary by subgroups by conditioning the sample based on gender, race, baseline GPA, composite ACT scores, and predicted exam scores. Although differential treatment effects by subgroup are generally not statistically distinguishable, Appendix Table A.2 offers suggestive evidence that permitting computers is more detrimental to male students than female students and to students with relatively high ACT scores. Still, these differences are only significant at the 10% level, could be chance findings, and are worthy of future research to verify their robustness. The results reported in panel C suggest that the negative impact of computers is similar for students with relatively low baseline GPAs and for students with relatively high baseline GPAs. While this is qualitatively different from Beland and Murphy (2016), who find that the positive effects associated with banning mobile phones in UK high schools is concentrated in the lowest-achieving students, the size of the standard errors reported in panel C of Appendix Table A.2 do not allow us to rule out this possibility.

We use the method suggested by Abadie, Chingos, and West (2013) to further investigate whether computer and tablet usage is most harmful for students who would otherwise perform well in the absence of treatment. We first compute predicted exam scores for those in the *control* group, using the leave-out fitted values:

$$Y_k = \beta'_{(-i)}X_k + \varepsilon_k;\ \ k \neq i \tag{2}$$

Y_k is individual test score and X_k includes individual covariates. 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:[29]

$$\hat{Y}_i = \hat{\beta}'_{(-i)}X_i. \tag{3}$$

Panel E of Appendix Table A.2 reports estimates of Eq. (1) for students within the lower and upper-halves of the distribution of $\hat{Y}_i$. The treatment effect appears to be more negative for students with higher predicted exam scores, although the estimates reported in columns 1 and 2 are not statistically distinguishable. As a final check, Appendix Table A.3 reports quantile treatment effects. The results from this exercise suggest that prohibiting computers in the classroom causes a roughly 0.2 standard deviation reduction in the first, second, and third quartiles of the performance

[28] Specifically, this coefficient bound is calculated as: $\beta^* = \tilde{\beta} - [\dot{\beta} - \tilde{\beta}](\frac{1-\tilde{R}}{\tilde{R}-\dot{R}})$, where $\tilde{\beta}$ is the treatment effect when all covariates are included, $\dot{\beta}$ is the treatment effect when only basic controls (instructor-by-semester fixed effects and class hour-by-semester fixed effects) are included, $\tilde{R}$ is the R^2 when all controls are included, and $\dot{R}$ is the R^2 when only basic controls are included. Oster (2015) actually suggests a less conservative coefficient bound, calculated as : $\beta^* = \tilde{\beta} - [\dot{\beta} - \tilde{\beta}](\frac{1.3\tilde{R}-\tilde{R}}{\tilde{R}-\dot{R}})$, to account for the fact that some variation in the outcome is likely to be idiosyncratic. The results in Panel A of Table 5 imply a bounded coefficient of -0.09 using the more conservative method described above and a bounded coefficient of -0.15 using the bound suggested by Oster (2015).

[29] Note that Eq. (4) 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.

Table 6
Estimates by academic 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)
A. Laptop and modified-tablet classrooms vs. non-computer classrooms						
Laptop/tablet class	−0.18***	−0.17*	−0.15	0.13	−0.15**	0.00
	(0.06)	(0.10)	(0.11)	(0.15)	(0.07)	(0.09)
TUCE pre-exam sample			X	X		
Observations	711	252	203	203	459	459
B. Unrestricted laptop/tablet classrooms vs. non-computer classrooms						
Computer class	−0.18***	−0.23*	−0.24*	0.13	−0.15*	−0.01
	(0.07)	(0.13)	(0.14)	(0.17)	(0.08)	(0.11)
TUCE pre-exam sample			X	X		
Observations	507	181	154	154	326	326
C. Modified-tablet classrooms vs. non-computer classrooms						
Modified-tablet class	−0.17**	−0.13	−0.04	0.01	−0.15*	0.03
	(0.07)	(0.13)	(0.14)	(0.21)	(0.09)	(0.12)
TUCE pre-exam sample			X	X		
Observations	466	169	129	129	297	297

Notes: This table report 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, respectively.

distribution, but we cannot rule out substantial differences in the magnitude of these effects.

7. Robustness checks

7.1. 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 for this possibility, we constructed a pre-exam, modeled after the Test of Understanding in College Economics (TUCE)[30] 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.[31] In panel A, column 3, of Table 6, we reproduce estimates from Table 3 after restricting the sample to those individuals in the spring who took the TUCE exam. We again find that permitting computers decreases exam scores (coefficient = −0.15). In column 4, we use 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 panel A, Table 6 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 computers 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.[32] 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 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 panel A of column 6 reveals that, on portions of the test related to classes where all students were exposed to equal treatment of computer access, there is no difference in exam performance.

Panels B and C of Table 6 report the same results as panel A but after restricting the treatment group to classrooms that permitted unrestricted laptop or tablet usage (panel B) or to classrooms that permitted only modified-tablet usage (panel C). Although the sample sizes are too small for precise inference, columns 4 and 6 of panel B indicate that, relative to students in the control group, students in the first treatment group performed better on the TUCE and roughly equally as well on the personal finance questions. We interpret this as additional suggestive evidence that students assigned to classrooms in the first treatment arm did not systematically differ from students in the control group on unobservable dimensions.

[30] See Walstad, Watts, and Rebeck (2007).
[31] Approximately 50 students did not take the pre-exam during the spring semester.
[32] 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 was not tested on their final exam.

7.2. Clustering

Appendix Table A.4 compares robust, conventional, and clustered standard error estimates for the specification described by Eq. (1). We also report wild-bootstrap p-values using the technique described by Cameron et al. (2008) to ensure that inference accounts for the relatively small number of clusters in our experiment. Robust and conventional standard errors are nearly identical, but, surprisingly, clustered standard errors are substantially smaller. Estimates of the intraclass correlation coefficient (ICC), reported in Appendix Table A.4, are negative for our preferred outcome of multiple choice and short answer questions. As discussed in Giraudeau (1996), there is little theoretical basis for a negative ICC and the negative estimates of such strongly suggests that test scores exhibit little correlation within classrooms after including professor and class hour fixed effects.

To further substantiate the precision of our estimates, we construct estimates of Eq. (1) using the two-step grouped-data estimation procedure for models with microcovariates described by Angrist and Pischke (2009; pp. 313–314). In the first step of this procedure, we construct covariate-adjusted classroom effects by estimating:

$$Y_{ijht} = \mu_{jht} + \gamma' X_i + \eta_{ijht}. \tag{4}$$

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 Eq. (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:[33]

$$\hat{\mu_{jht}} = \kappa_{jt} + \lambda_{ht} + \pi Z^{jht} + \epsilon_{jht}. \tag{5}$$

Column 4, panel A, of Appendix Table A.4 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.[34] 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% level.

8. Conclusion

The results from our randomized experiment suggest that computer devices can reduce student's knowledge of the material gained during the semester. Our estimates imply that permitting computers or laptops in a classroom lowers exam scores by 0.18σ. Considering that the standard deviation of exam scores among students in our sample is 9.2, this amounts to about a 1.7 point reduction on a 100 point scale. Although many aspects of West Point differ from typical 4-year undergraduate institutions, there are reasons to believe that permitting computers in traditional lecture-style classrooms could have similar or 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.

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. 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 computers are available.

We find negative effects in both treatment arms of our experiment, 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 harm 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 students' ability to effectively take notes on a tablet may be limited.

Our results are similar to others which examine the effects of charter school, teacher value-added, and class size effects on test scores.[35] For example, Dobbie and Fryer (2011) find that attending charter schools in Harlem increases math and ELA test scores for elementary and middle schoolers by 0.2 standard deviations. The results in Rockoff (2004) reveal that a 0.2 standard deviation increase in teacher quality produces a 0.2σ increase in reading and math test scores. Similarly, having a teacher with 10 or more years of experience relative to a new teacher increases test scores by 0.17 standard deviations.[36] Finally our results are in-line, although slightly larger, than research on the effects of class size: Krueger and Whitmore (2001) estimate that those who attended a small class while in elementary school had 0.1 standard deviations (0.2σ for black students) higher ACT or SAT scores in high school using the Tennessee STAR project and Bandiera, Larcinese, and Rasul (2010) find that a standard deviation increase in class size decreases end of the year test scores at a university by 0.074 standard deviations.

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. Additionally, since all students at USMA have equal access to laptops and tablets outside of the classroom, we cannot test whether school-wide access to computers affects student outcomes.[37] Rather, our results relate only to classes where students

[33] 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.

[34] This follows the suggestion of Donald and Lang (2007). When both treatment arms are combined (panel A), there are 50 classrooms, 16 combinations of professor and semester, and 8 combinations of class hour and semester, resulting in a residual degrees of freedom of 50-15-7-1=27.

[35] We focus on effects on test scores, but there is a stream of literature that examines the impacts on GPA. (See Angrist, Oreopoulos, and Williams (2014) and Angrist, Lang, and Oreopoulos (2009) for examples of the effect of financial incentives. See Lyle (2007) and Carrell, Fullerton, and West (2009) for peer effects.

[36] Several other value-added studies find results of similar magnitude (between 0.10 and 0.15) for students in lower grade levels. See Chetty, Friedman, and Rockoff (2014), Rivkin, Hanusheck, and Kain (2005), Rockoff (2005), Aaronson, Barrow, and Sander (2007), and others.

[37] Bulman and Fairlie (2016) survey the literature on computer and internet access in schools, home computer usage, and computer aided instruction (CAI); the literature finds mixed results for each of these. For example, while not examining computer use in classrooms specifically, Belo, Ferreira, and Telang (2013) find that

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 our results, and the increasing emphasis of using technology in the classroom, additional research aimed at distinguishing between these channels is clearly warranted.

the spread of broadband access in Portugal led to middle school students performing worse on their exams. On the other hand, Carrillo, Onofa, and Ponce (2011) report a 0.30 standard deviation increase in mathematics test scores for primary students who received access to an Internet-enabled computer lab and adaptive learning software.

Appendix A

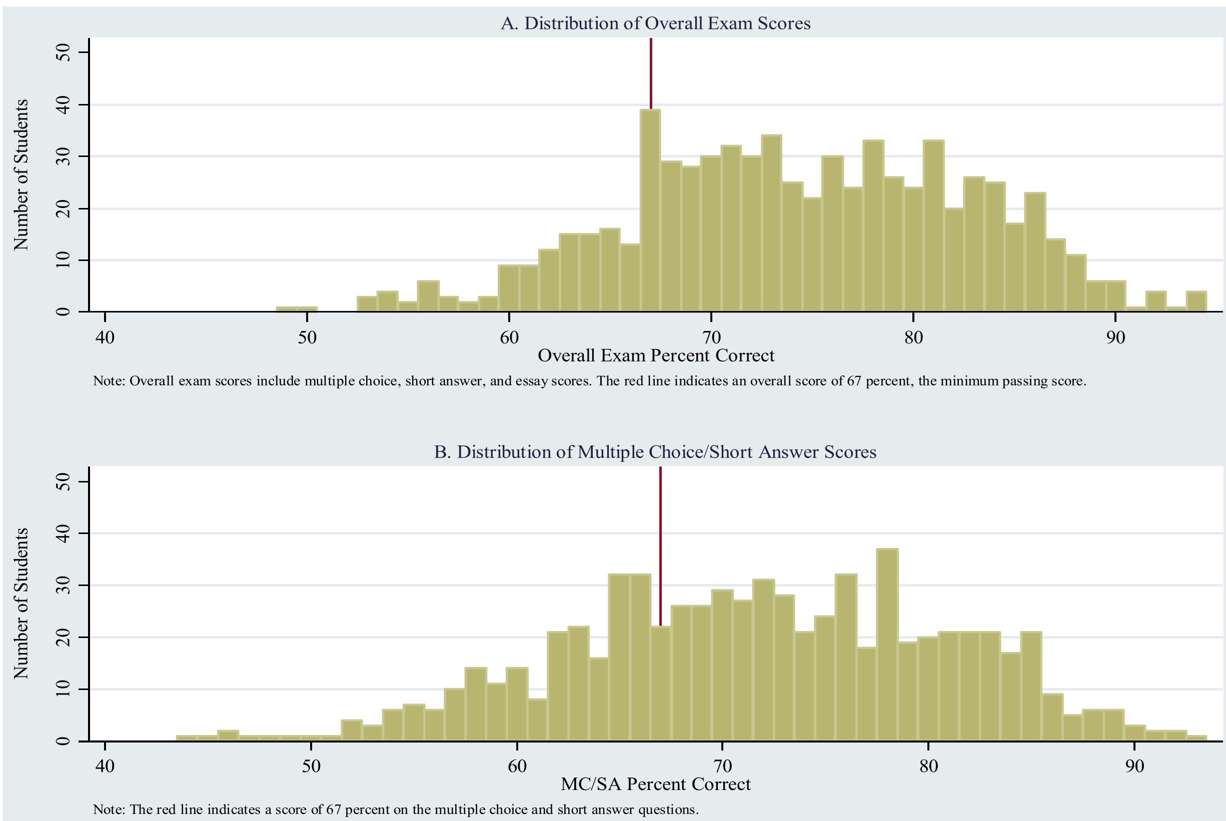

Fig. A.1. Distribution of exam scores.

Table A.1
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	−0.15**	−0.12*	−0.19***	−0.19**	0.03	0.04
	(0.06)	(0.06)	(0.06)	(0.09)	(0.06)	(0.18)
GPA at start of course	0.89***	0.89***	0.94***	0.91***	0.69***	0.58***
	(0.07)	(0.08)	(0.07)	(0.06)	(0.07)	(0.09)
Composite ACT	0.057***	0.052***	0.049***	0.050***	0.027***	0.036**
	(0.012)	(0.011)	(0.012)	(0.013)	(0.010)	(0.015)
Instructor fixed effects	X		X		X	
R^2	0.48	0.45	0.43	0.39	0.51	0.18
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, respectively.

Table A.2
Estimates by subgroup, dependent variable: Final exam multiple choice and short answer score.

	A. By gender		B. By race	
	women (1)	Men (2)	nonwhite (1)	White (2)
Laptop/tablet class	0.02	−0.21***	−0.22**	−0.18**
	(0.14)	(0.06)	(0.10)	(0.07)
Observations	131	580	244	467
Difference in Treatment Effects P-Value				
Robust Std. Errors	0.093		0.727	
Clustered Std. Errors	0.077		0.643	
	C. By baseline GPA		D. By ACT score	
	GPA ≤ 2.89 (1)	GPA > 2.89 (2)	ACT ≤ 28 (1)	ACT > 28 (2)
Laptop/tablet class	−0.19**	−0.19**	−0.07	−0.27***
	(0.09)	(0.08)	(0.08)	(0.08)
Observations	355	356	351	360
Difference in Treatment Effects P-Value				
Robust Std. Errors	0.968		0.082	
Clustered Std. Errors	0.966		0.040	
	E. By predicted exam score			
	Bottom half of distribution (1)	Top half of distribution (2)		
Laptop/tablet class	−0.13	−0.24***		
	(0.08)	(0.08)		
Observations	356	355		
Difference in Treatment Effects P-Value				
Robust Std. Errors	0.340			
Clustered Std. Errors	0.278			

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. The *P*-values reported in each panel are from the hypothesis test of equal treatment effects within each sub-group in the panel. We report *P*-values for these difference in coefficients tests using both robust standard errors and standard errors clustered at the classroom level. ***,**, and * denote significance at the 1%, 5%, and 10% level, respectively.

Table A.3
Quantile regression estimates, dependent variable: Final exam multiple choice and short answer score.

	Q = 0.25 (1)	Q = 0.5 (2)	Q = 0.75 (3)
A. All classrooms in sample			
Laptop/tablet usage	−0.23***	−0.20***	−0.18***
	(0.08)	(0.05)	(0.07)
Observations	711	711	711
B. Unrestricted laptop/tablet classrooms and non-computer classrooms			
Computer usage	−0.25***	−0.21***	−0.19**
	(0.06)	(0.06)	(0.09)
Observations	507	507	507
C. Modified-tablet classrooms and non-computer classrooms			
Modified-tablet usage	−0.23***	−0.13*	−0.12
	(0.08)	(0.08)	(0.09)
Observations	466	466	466

Notes: This table reports quantile regression estimates of computer usage on exam scores. 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, respectively.

Table A.4
Comparison of standard errors, dependent variable: Final exam multiple choice and short answer score.

	Robust standard errors (1)	Conventional standard errors (2)	Clustered standard errors (3)	Group means (4)
A. Laptop and modified-tablet classrooms vs. non-computer classrooms				
Laptop/tablet class	−0.18***	−0.18***	−0.18***	−0.18***
	(0.06)	(0.06)	(0.04)	(0.05)
P-Value	0.0020	0.0021	0.0000	0.0014
Wild Bootstrap *P*-Value			0.0000	
Interclass corr. coef. estimate			−0.029	
ICC std. error estimate			(0.015)	
Clusters (classrooms)			50	
Resid deg-of-freedom	679	679	49	27
Observations	711	711	711	50
B. Unrestricted laptop/tablet classrooms vs. non-computer classrooms				
Computer class	−0.18***	−0.18***	−0.18***	−0.18**
	(0.07)	(0.07)	(0.04)	(0.06)
P-Value	0.0055	0.0050	0.0001	0.0102
Wild Bootstrap *P*-Value			0.0000	
Interclass corr. coef. estimate			−0.031	
ICC std. error estimate			(0.017)	
Clusters (classrooms)			35	
Resid deg-of-freedom	476	476	34	13
Observations	507	507	507	35
C. Modified-tablet classrooms vs. non-computer classrooms				
Modified-tablet class	−0.17**	−0.17**	−0.17***	−0.17***
	(0.07)	(0.07)	(0.03)	(0.05)
P-Value	0.0192	0.0220	0.0000	0.0087
Wild Bootstrap *P*-Value			0.0000	
Interclass corr. coef. estimate			−0.051	
ICC std. error estimate			(0.018)	
Clusters (classrooms)			33	
Resid deg-of-freedom	434	434	32	10
Observations	466	466	466	33

Notes: This table reports estimates from a regression of exam scores on an indicator for being assigned to a classroom that permits laptops or tablets. Scores are standardized with mean 0 and standard deviation 1 for each semester. Estimates reported in columns 1–3 include (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) 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, respectively.

References

Aaronson, D., Barrow, L., & Sander, W. (2007). Teachers and student achievement in the Chicago public high schools. *Journal of Labor Economics, 25*(1), 95–135.

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

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, 1*(1), 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, 38*(3), 245–263.

Beland, L.-P., & Murphy, R. (2016). Ill communication: Technology, distraction, & student performance. *Labour Economics, 41*, 61–76.

Belo, R., Ferreira, P., & Telang, R. (2013). Broadband in school: Impact on student performance. *Management Science, 60*(2), 265–282.

Bulman, G., & Fairlie, R. W. (2016). Technology and education: Computers, software, and the internet. *National Bureau of Economic Research Working Paper No. w22237*, NBER.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *The Review of Economics and Statistics, 90*(3), 414–427.

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

Carrillo, J., Onofa, M., & Ponce, J. (2011). Information technology and student achievement: Evidence from a randomized experiment in Ecuador (IDP-WP-223). *IDB Working Paper Series*, IDB.

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.

Dobbie, W., & Fryer, R. G., Jr. (2011). Are high-quality schools enough to increase achievement among the poor? Evidence from the harlem children's zone. *American Economic Journal: Applied Economics, 3*(3), 158–187.

Donald, S. G., & Lang, K. (2007). Inference with difference-in-differences and other panel data. *The Review of Economics and Statistics, 89*(2), 221–233.

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

Giraudeau, B. (1996). Negative values of the intraclass correlation coefficient are not theoretically possible. *Journal of clinical epidemiology, 49*(10), 1205.

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). This year, I resolve to ban laptops from my classroom. The Washington Post, December 30.

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.

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

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, 89*(2), 289–299.

Mueller, P. A., & Oppenheimer, D. M. (2014). The pen is mightier than the keyboard: Advantages of longhand over laptop note taking. *Psychological Science*, 1–10. doi:10.1177/0956797614524581.

Oster, E. (2015). Unobservable selection and coefficient stability: Theory and evidence. *University of Chicago Booth School of Business Working Paper*.

Patterson, R., & Patterson, R. (2016). Laptop computers in college classrooms and academic performance. *Working Paper*.

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

Rockoff, J. E. (2005). Teachers, schools and academic achievement. *Econometrica, 73*, 417–458.

Rockoff, J. E. (2004). The impact of individual teachers on student achievement: Evidence from panel data. *American Economic Review: Papers and Proceedings, 94*(2), 247–252.

Sana, F., Weston, T., & Cepeda, N. J. (2013). Laptop multitasking hinders classroom learning for both users and nearby peers. *Computers & Education, 62*, 24–31.

U.S. Department of Education, 2016 National Education Technology Plan. (2016). "Future ready learning: Reimagining the role of technology in education." Washington, DC.

U.S. News and World Report. (2016). National Liberal Arts Colleges Rankings. Accessed May 1, 2016. http://colleges.usnews.rankingsandreviews.com/best-colleges/rankings/national-liberal-arts-colleges?int=a73d09.

Walstad, W. B., Watts, M., & Rebeck, K. (2007). *Test of understanding in college economics* (4th ed.). 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 Accessed May 1, 2016 http://nces.ed.gov/pubs2007/2007020.pdf.

Wurst, C., Smarkola, C., & Anne Gaffney, M. (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.