Stanford cepa | Center for Education Policy Analysis

CEPA Working Paper No. 15-14

The Effects of Class Size in Online College Courses: Experimental Evidence

AUTHORS

Eric Bettinger
Stanford University

Christopher Doss
Stanford University

Susanna Loeb
Stanford University

Eric Taylor
Harvard University

ABSTRACT

Class size is a first-order consideration in the study of education production and education costs. How larger or smaller classes affect student outcomes is especially relevant to the growth and design of online classes. We study a field experiment in which college students were quasi-randomly assigned to either a large or a small class. All classes were conducted online. Large classes had, on average, ten percent more students than small classes (mean increase 10.2 percent, standard deviation 4.0 percent). The experiment was conducted at DeVry University, one of the nation's largest for-profit post-secondary institutions, and included over 100,000 student course enrollments in nearly 4,000 classes across 111 different undergraduate and graduate courses. We examine class-size effects on student success in the course and on student persistence in college. We find little evidence of effects on average or for a range of specific course types. Given the large sample, our estimates are quite precise, eliminating the likelihood of any meaningful effect of class size changes of the magnitude in question.

VERSION

May 2015

Suggested citation: Bettinger, E., Doss, C., Loeb, S., & Taylor, E. (2015). The Effects of Class Size in Online College Courses: Experimental Evidence (CEPA Working Paper No.15-14). Retrieved from Stanford Center for Education Policy Analysis: http://cepa.stanford.edu/wp15-14

The Effects of Class Size in Online College Courses: Experimental Evidence

Eric Bettinger, Christopher Doss, Susanna Loeb, and Eric Taylor[1]
Stanford University

May 2015

Abstract: Class size is a first-order consideration in the study of education production and education costs. How larger or smaller classes affect student outcomes is especially relevant to the growth and design of online classes. We study a field experiment in which college students were quasi-randomly assigned to either a large or a small class. All classes were conducted online. Large classes had, on average, ten percent more students than small classes (mean increase 10.2 percent, standard deviation 4.0 percent). The experiment was conducted at DeVry University, one of the nation's largest for-profit post-secondary institutions, and included over 100,000 student course enrollments in nearly 4,000 classes across 111 different undergraduate and graduate courses. We examine class-size effects on student success in the course and on student persistence in college. We find little evidence of effects on average or for a range of specific course types. Given the large sample, our estimates are quite precise, eliminating the likelihood of any meaningful effect of class size changes of the magnitude in question.

[1] Center for Education Policy Analysis, 520 Galvez Mall, CERAS Building, Stanford, CA 94305 ebettinger@stanford.edu, cdoss@stanford.edu, sloeb@stanford.edu, and erictaylor@stanford.edu. We are grateful to DeVry University, and especially Aaron Rogers, Ryan Green, and Earl Frischkorn. We thank Victor Lavy, Sue Dynarski, as well as seminar participants at CESifo Institute for helpful comments on earlier drafts. The research reported here was supported in part by the Institute of Education Sciences, U.S. Department of Education, through Grant R305B090016 to Stanford University. The opinions expressed are those of the author and do not necessarily represent views of the Institute or the U.S. Department of Education or DeVry University.

1. Introduction

Class size is a perennial issue in the economics of education. It has implications for both the cost and production of education. In K-12 education, the effects of class size have been vigorously debated (e.g. Mosteller 1995 Hanushek 2002, Krueger 2002, Hoxby 2000, Krueger and Whitmore 2001, Angrist and Lavy 1999, Leuven Oosterbeek, Ronning 2008, Gary-Bobo, Mahjoub 2013, Woessman and West 2006, Dynarski, Hyman, and Schanzenbach 2011, Chetty, Friedman, Hilger, Saez, Schazenbach, and Yagan 2010). Additionally, several states have attempted to regulate class size as a means of improving K-12 education (e.g. Morgan-Hart Class Size Reduction Act in California). Class size in college has received some, though far less, attention (Bettinger and Long 2010; De Girogi et al, 2012; Bandiera et al, 2010).

Our setting focuses on class size in virtual classrooms. Online college courses are becoming more and more important in higher education. About one-third of students in the US take at least one course online during their college career, and the proportion of students who take at least one course online has tripled over the past decade (Allen and Seaman 2013). According to the US Department of Education (2014), in the fall of 2012 12.5% of students in Title IV institutions exclusively enrolled in distance education classes and another 13.3% enrolled in some, but not all, distance education classes. More than 70 percent of public colleges now have some programs that occur completely online. With the growth of the online education space has come a spectrum of courses that vary in accessibility and size. Many higher educational institutions have chosen to create virtual classrooms, an online analog of the traditional brick-and-mortar classroom. Unlike the “Massive Open Online Courses” (MOOCs) which have garnered so much attention, these virtual classes are only available to students enrolled in the university and are approximately the same size as traditional classes.

Class size online presents a different set of financial and educational production challenges than class size for in-person classrooms. For example, the cost of adding an additional student is often negligible in online settings. No new desk and no new equipment are needed. The new costs are only incurred through additional staff time necessary to provide education to a larger group of students. Additionally, whereas class size might affect students through peers or congestion (e.g. Lazear 2002) in traditional classrooms, interactions are substantially different in an online setting where discussion boards are the primary forum through which peers interact. While online courses present an opportunity to reduce higher education costs, attempted cost saving through increased class size may affect students' educational outcomes. No research that we know of sheds light on this question.

In estimating the effect of class size on students, selection issues are pervasive. These selection issues are likely more pronounced at the collegiate level, as compared to the K-12 level, because students have substantial flexibility in choosing their course of study. Class size varies across fields and average student outcomes also vary across fields, unlikely driven only by class sizes differences. Similarly, professors of perceived higher quality can attract more students and their outcomes could vary because of both size and professor quality. Higher achieving students in turn might have preferences for small classes making it difficult to separate student quality from class size effects.

To measure the effects of college class size while addressing these empirical difficulties, we analyze over 100,000 student course enrollments at DeVry University in 2013-14. DeVry serves over 130,000 undergraduates, or about 5 percent of the for-profit college market, placing it among the 8-10 largest for-profit institutions in 2010. Two-thirds of undergraduate courses occur online, the other third occur at nearly 100 physical campuses throughout the United States.

Devry's online model focuses on recreating the features of a traditional classroom in an online context. The classes are open only to students enrolled in the university, and the average class size of 30 students is on par with class size in conventional in-person classes at the university.

In the fall of 2013, DeVry conducted a class size field experiment. The field-experiment created quasi-random variation on two dimensions: class size and peer quality. Over the span of four academic terms, more than 4,000 course sections (classes) of 111 courses were randomly assigned, within course and term blocks, to either "large" or "small" conditions. Large classes had 34 student, on average, compared to 31 in small classes. Students were not randomly assigned to sections, however, the DeVry registration process makes student assignment to sections quasi-random (plausibly ignorable) conditional on student registration date and time. Sections are not differentiated in the registration system except by section number, and students (almost) always register for the first section listed with an open seat; thus two students with sufficiently similar registration times were effectively randomly assigned to "large" or "small" sections. Variation in peer quality arose, as we describe in Section 3, from the normal patterns of student registration at the university. In short, in some "large" sections the added students lowered the average academic ability of students in the section, while in others average ability did not change. Thus the primary treatment-control contrast we study combines class size and peer quality in the "treatment" but, as described later, we test whether the effects are driven by one or the other.

We find, after addressing issues of selection, that small changes in online college class sizes have no discernable effect on student learning or persistence in college. This result is consistent across different types of courses where one might expect a meaningful exception. For example, classes which require substantial faculty work or courses where increased class size might crowd out meaningful interactions between faculty and students theoretically could generate meaningful class size effects. We find, however, that even in these courses no meaningful class size effect is evident.

The paper is organized as follows. Section 2 includes the background on class size and online schooling. Section 3 presents our data and methodology. Section 4 presents our baseline results. Section 5 presents heterogeneity results. Section 6 presents robustness checks and Section 7 concludes.

2. Background on Class Size

To date, most of the research on class size has focused on K-12 settings. As Hanushek (2002) and Krueger (2002) have illustrated, the quality of existing evidence and conclusions vary across studies. The most cited study is the Tennessee STAR experiment where students from kindergarten through third grade were assigned to either small (~15) or large (~22) class sizes. Multiple studies of STAR provide some evidence that small classes generate positive impacts on short-run (Mosteller 1995 on academic achievement), medium-run (Krueger and Whitmore 2001 on college attendance and SAT scores), and long-run outcomes (Dynarski, Hyman and Schazenbach 2011 on college completion; Chetty et al. 2010 on earnings). Some scholars have suggested that the STAR findings are the result of Hawthorne effects which led teachers to behave differently under experimental conditions rather than true class-size effects (Hoxby 2000, Ding and Lehrer 2010).

Other studies use a variety approaches to adjust for selection in the estimation of class-size effects with inconsistent findings. Angrist and Lavy (1999) exploit variation in class size in Israel arising from the application of Maimonides' rule to estimate the impact of class size.

Other papers use similar rules to identify impacts of class size (Gary-Bobo and Mahjoub, 2013) generally finding that small classes benefit students. In contrast, Hoxby (2000) and Leuven et al (2008) use natural variation in cohort size (Leuven et al also use maximum class size rules) to estimate impacts of class size and find no impact of class size on student outcomes.

On the collegiate level there is some evidence that larger classes negatively affect student outcomes. Bettinger and Long (2010) use natural variation in cohort size in college to estimate the impacts of class sizes in introductory courses and find positive impacts of small classes. Johnson's (2010) use of hierarchical linear modelling and Bandiera, Larcinese, and Rasul (2010) use of student fixed effects similarly find that small classes produce higher student outcomes. These results are further supported by De Giorgi, Pellizzari, and Woolston (2012) who leverage random assignment of students to teachers in an Italian university.

The findings to date, which are most often but not consistently positive for small classes, may or may not generalize to virtual college courses. Most studies focus on K-12 education, though the previous collegiate studies and a handful of other papers (Dillon, Kokkelenberg and Christy 2002, Kennedy and Siegfried 1997, Becker and Powers 2001) address higher education. None of the studies focus on virtual settings, even though those setting are increasingly common for students. The studies also define large and small classes differently. STAR compares class sizes of 15 and 22 students, a difference of 7 students. In Angrist and Lavy's (1999) study, the typical comparison was between classes of 39 and 27.[2] Changes in class size of approximately 10 percent for virtual classes have substantial implications for cost in the online setting but the extant research does not provide clear evidence on the likely effects of such a change on students.

[2] At the closest discontinuity to the mean, the comparisons were between classes of 39 and 27. At the 90th percentile, the comparisons were between classes of 39 and 30.

Why Might Class Size Matter?

An increase in class size is likely to affect many, but not all, aspects of a class. In this setting the large and small versions of a class are structured in identical ways. Students in each type of class must listen to the same lectures, complete the same assignments, and take the same assessments. The increase in class size will therefore not affect a professor's effort in preparing for a class. In contrast, there are many mechanisms through which an increase in the number of students can affect classroom dynamics and a professor's workload in *teaching* a class.

Lazear's disruptive peer model is the best-known model of how class-size affects classroom dynamics. In the model, each peer has a probability of creating a disruption which could impede the learning of other students. As class sizes increase, the probability of a disruption occurring increases. Hence, large class sizes, on average, lead to more disruptions and less learning. Lazear's model, while developed with traditional K-12 classrooms in mind, can still be applicable to higher education classrooms and virtual settings in particular. In Devry's online courses, typical face-to-face classroom conversations are replaced with asynchronous postings in discussion boards. Students may struggle to get personalized responses to their postings if congestion crowds out productive discussion. Bettinger, Loeb and Taylor (2015) test the effects of the extent (length and frequency) of peer posts in these discussion boards and finds evidence of congestion. Even though peers who post more do better in the courses themselves, their peers do worse.

In teaching the class the professor can respond to the increased workload in two ways. First, they have the option to work harder to meet the demand of the additional students without compromising the quality of education each student obtains. The other option is to put forth the same amount of effort but spread it across more students thereby decreasing the educational input each student receives. This is especially salient in this setting given that there are no teaching assistants with which the professors can share the additional burden.

‘Essentially parts of a professor’s job is a public good for students in the class. For example in face-to-face classrooms, learning partially occurs as students typically ask and respond to questions from the professor. In these virtual classrooms students ask and respond to questions via the online discussion board posts. As the number of students in a class grows, so do the number of posts, and each new one has a cost associated with it. Professors may respond to a smaller proportion of student statements, thereby decreasing the probability that an individual student will have their question answered, or the professor may respond less fully thereby degrading the quality of responses a student receives. Similarly as the number of students grow, the probability that a professor will initiate a conversation with a student via the discussion board decreases.

Competition for the professor’s time outside of class can similarly increase as class size grows. In brick-and-mortar classrooms particularly motivated students can attend office hours or make appointments with a professor to get more individualized attention. With more students, the professor can spend less time with any one student. Similarly, particularly motivated students in our setting can email the professor independently for more individualized attention. The ability of the professor to comprehensively answer the larger number of emails associated with a larger class decreases.

The same mechanisms are at play in completing assignments. Many assignments in virtual classes require feedback from professors (e.g. multiple drafts of a term paper, comments on exams, etc…). Such feedback may be lower quality in larger classes if professors devote less time to each piece of work. Generally, larger classes may limit the time that a professor allocates to any one student, which, in turn, might reduce student engagement or affect the quality of feedback and thus student learning.

The effects of class size changes could vary by the characteristics of the class. In particular, classes that require professors to spend more time one-on-one with students, through answering questions or grading assessments may see greater class size effects than other classes. Classes requiring projects, laboratories, and papers may require more feedback from professors than do other classes. The quantity and rigor of feedback may change with increasing class size more strongly for these courses. Class size effects also could differ by academic discipline if the likelihood that students contact professors depends on the discipline or the assignment structure of the class. Class size effects may differ depending on the extent of peer interaction, as well. Peers could have a greater effect on classes that required more interaction such as projects, laboratories, and perhaps computer technology. This potential peer mechanism is especially relevant in this study where the class size increase and the quality of the one's peers were intimately related.

3. Treatment, Setting, and Methodology

We analyze data from a field-experiment conducted by DeVry University to study whether changes in class size affect student outcomes in online college courses. In particular, we estimate the effect of increasing online class sizes on student learning, as measured by course grades; and on student persistence in college, measured by enrollment the next term and the number of credits attempted the next term. We investigate heterogeneity in these effects by course type (science and mathematics courses, social science courses, humanities courses, etc.), and by the types of course assignments (courses that require projects, laboratories, both, or neither).

The Treatment and Assignment to Treatment

The field-experiment created quasi-random variation on two dimensions: class size and peer quality. Over the span of four academic terms, more than 4,000 course sections (classes) of 111 courses were randomly assigned, within course and term blocks, to either "treatment" or "control".[3] Class size was directly manipulated. The university set higher enrollment caps for treatment sections—on average, 34 students compared to 31 students in control sections. While students were not randomly assigned to sections (treatment), the DeVry registration process makes student assignment to sections quasi-random conditional on registration date and time. Variation in peer quality arose, as we describe in the next few paragraphs, from the normal patterns of student registration at the university. In short, in some treatment sections the added students lowered the average academic ability of students in the section, while in others average ability did not change. Thus the primary treatment-control contrast we study combines class size and peer quality in the "treatment", but, as described later, we test whether the effects are driven by one or the other.

Registration for online courses at DeVry follows a few simple rules and patterns, ignoring for a moment the experimental manipulation. Students register online for courses and sections. The enrollment window starts six months before the term begins and ends a few days into the eight-week term. If demand exceeds the University's projections, additional sections are added. During registration, online course sections have no differentiating characteristics: meeting time and location are irrelevant, class sizes are identical, and professors are not identified. These

[3] DeVry divides its academic calendar into six eight-week terms called "sessions". The experiment was conducted during the July, September, and November 2013; and January 2014 terms.

features generate a simple but strong pattern: section 1 fills up with students first, then section 2 fills, and so on. Very few students choose to deviate from this pattern. Observable student characteristics vary by registration date. Notably, students with higher prior GPAs register earlier (a correlation of about 0.30 in any given term), thus generating important between-section variation in mean prior GPA.

During the four experimental terms, student registration began exactly as described in the previous paragraph. All sections within each of the participating courses had the same class size. Then on a set day, before the start of the term, DeVry administrators changed the class size cap of all odd numbered sections. This "cap change" or "class size assignment" day was six weeks out for the November and January terms (two-thirds of the sample) and two weeks out for the July and September terms. For nine out of ten courses, class sizes were increased in the odd numbered sections. A university administrator simply increased the enrollment cap in those sections, and new registrants began filling those slots following the pattern in the previous paragraph. In one out of ten courses, class sizes were decreased in the odd numbered sections. The administrator arbitrarily removed students from those sections and enrolled them in new sections.[4] In the final weeks before the term, additional students registered filling the now large and small sections.

The change in enrollment caps created differences between sections in class size, but also created differences between sections in peer characteristics. Consider the courses for which class size was increased in the odd numbered sections. Absent the experiment, students in section 1 and section 2 would have experienced the same class size and same distribution of peer quality. During the experiment, section 1 had both more students and the added students were likely less-

[4] The selection of students to be moved was arbitrary, but not, strictly speaking, random. The administrator who carried out this task had only one consideration when selecting whom to move: students with a hold on their account for financial or academic reasons could not be moved.

academically-prepared students. The added students were drawn from students who registered in the final weeks before the term began, students who registered only after the enrollment caps had been raised. By contrast, students in the last sections created, say m and $m - 1$, experienced similar peer characteristics even though section m had more students. Because of the importance of the assignment date, we distinguish in our analysis between those students who registered prior to the class size assignment date and those who registered subsequently. Among "incumbent" students (i.e. students registered prior to the class size assignment date), our only identifying assumption is that there is randomness locally around the registration time. For late registrants, many were quasi-randomly assigned because the administrator reassigned them somewhat arbitrarily. About 49 percent of students registered late enough to see a choice between sections, and their choice of sections may be endogenously influenced by the class size.[5]

Data and Descriptive Statistics

Table 1 presents descriptive statistics for our data and the experimental sample. Classes, on average, contained 32 students; though enrollment ranged between 16 and 40 students. Just under half of the sample attended small classes (48.7 percent). The most common classes were in business and management (35.5 percent) with a substantial number in humanities (17.1 percent) and science and mathematics (15.1 percent). We also classify courses based on their requirements for students. Approximately a quarter of courses require just projects and another quarter requires just laboratories. Another 14.7 percent require both and 35.8 percent require

[5] The potential endogeneity of late registrants' choice was not communicated by DeVry University. Rather, it was inferred from knowledge of the registration process. A screenshot of the registration window is shown in Figure 2 of the Appendix. In this window "Cap" indicates the student enrollment cap for a section, "Act" indicates the number of seats in the section taken at the time of registration, and "Rem" indicates the number of seats in the section remaining at the time of registration. In order to take class size into consideration a late registrant would need to infer the meanings of these labels.

neither. We hypothesize that courses with project or lab requirements might be more affected by class size changes than would other classes.

The "large" sections, on average, contained 34 students, in contrast to "small" sections, which contained 31 students, a difference of about ten percent. This percent difference, however, ranged across courses from 2.9 to 25 percent. Appendix Figure 1 shows the distribution of class size changes across students and indicates that the most common changes were approximately thirteen, nine percent and seven percent, respectively. Table 1 also shows that for the few measures of student characteristics available such as prior GPA, the students in big classes who registered before the size assignment day (cap change day) look almost identical to those in smaller classes. These samples will be our primary samples of interest for the study, though we will also look at the effects on students who registered after the size assignment day.

Methods and Identification Strategy

We make use of the assignment process in combination with a fixed effects strategy in order to estimate the effect of class size on a variety of student outcomes. The key insight is that while student characteristics covary with the time at which they enroll in a course, because assignment to a "large" classroom was quasi-randomly assigned, there will be variation in class size among students who registered for the same course at similar times. Moreover, in a short enough time window that class size variation should be spread among students with similar observed and unobserved characteristics.

The following regression model summarizes the approach:

$$Y_{isct} = \beta_0 + T_{isct}\beta_1 + X_{isct}\beta_2 + \alpha_{sct} + \varepsilon_{isct}$$

Here Y_{isct} represents the outcome of interest for student i enrolled in course c during session s who registered at time t. DeVry calls their eight-week academic terms "sessions". The primary outcomes of interest are three: (i) the grade the student received in the course (measured on 0-4 scale), (ii) student persistence to the next session/term on the extensive margin (an indicator = 1 if any enrollment the next session), and (iii) persistence on the intensive margin (number of credits attempted in the next session, including zeros). Persistence data are not available for the last (January 2014) term in the data.

The right hand side variables include the treatment variable, T_{isct}, which is the intended class size of the section to which the student was assigned. We defined this treatment variable in two ways: (i) a binary indicator for "small" section (compared to "large") and (ii) the log of intended class size. Our preferred specification is the log class size because it lends itself to a percent increase interpretation, which depends on the size of a "small" class and the increase in class size, whereas the binary indicator hides any heterogeneity in treatment. However, given the lack of clarity about the form of the effect, we use both approaches. X_{isct} represents a vector of pre-treatment student characteristics that includes prior cumulative GPA at DeVry, an indicator for being a new student, an indicator for missing prior GPA if not a new student, and an indicator for previously having failed a course.

All estimates include α_{sct} which represents fixed effects defined by the intersection of course, term, and registration date-time group. To create registration groups we, first, order students by the date-time they registered for the course; and, second, divide students into either (i) groups of 15 students with sequential registration date-time within course by term blocks, or (ii) vigintiles of registration date-time within course by term blocks. In the paper we report estimates using approach (i) and approach (ii), but the choice of 15 students or 20 quantiles is somewhat arbitrary. As described more below, the results presented are robust defining the fixed effects with larger or smaller groups of students, and more of fewer quantiles.

The identifying assumption is that within these groupings students are similar on all observable and unobservable characteristics and were effectively randomly assigned to different class sizes based on the scheme explained above. This identifying assumption can be partially tested by looking at the covariate balance across class size within course-by-session-by-registration-group fixed effects. Table 2 provides results of that test. Panel A looks at the entire sample and Panels B and C disaggregate the sample by those who registered before the cap on the sections was changed, and those who registered after the cap change, respectively. Only one estimated coefficient in 48 tests any evidence of imbalance. Moreover, the difference is quantitatively small: the average students in 10 percent larger class would have a prior GPA 0.009 points lower on a 4 point scale.

As stated earlier, students who registered after the class caps were changed, and therefore typically had weaker prior academic outcomes, were added to previously full sections. By enlarging these previously full sections with later registrants, class sizes were not only increased, but lower performing peers were mixed with higher performing peers. There could therefore be a class size effect *and* a peer effect that would differentially affect a student based on their prior academic success. To test whether students who registered before and after the cap were differentially affected by this treatment, we interacted a binary indicator for registering after the cap changed with the treatment.

$$Y_{isct} = \beta_0 + T_{isct}\beta_1 + X_{isct}\beta_2 + PostAssignment_{isct}\beta_3$$
$$+ T_{isct} * PostAssignment_{isct}\beta_4 + \alpha_{sct} + \varepsilon_{isct}$$

4. Main Results

Table 3 presents main results from four models, and two different fixed effects strategies. Panels A and C show the effect of enrolling in a small class, as described in Section 2. In all cases the point estimates are quantitatively small and statistically insignificant. Small class sizes do not seem to affect student grades, the probability that they enroll in the next term, nor the number of credits they attempt in the next term.

Panels B and D show estimates allowing the effect of class size to differ for students registering before or after the class size assignment day. Recall that for students registering before the cap change, the treatment-control contrast includes both an increase in the number of peers and reduction in the quality of peers. Broadly speaking, these results suggest the estimated effects do not differ for students who registered before and after the class size cap change, with the possible exception of effects on course grade. Estimates for persistence outcomes are again quantitatively small and not statistically significant. The estimates in Panel D for course grade suggest students registering before the cap change may have been worse-off in larger classes populated with less academically skilled peers; however, the effects are not robustly statistically significant and are not seen in Panel B. Moreover, the grade point estimates in Panel D are still educationally small, as we discuss in the next paragraphs.

Given the broadly null effects shown in Table 3, we next detail the size of class size effects we would be able to rule out. Table 4 reports the 95 percent confidence interval for the effect of a 10 percent increase class size. Recall that 10 percent was the average change in our sample. We focus this exercise on the results in Table 3 Panel D, which show the largest point estimates and lease precision. Confidence intervals are reported in student standard deviation units.

For students who registered before the class size assignment day, we can rule out a 0.03-0.04 standard deviation reduction in course grade. In our data a one standard deviation difference in course grade is about 1.3 grade points, or about the difference between a "B" and a "C-". Thus a 0.04 effect can be thought of as about one-sixth of the difference between a "B" and a "B-". Similarly we can rule out a grade improvement of 0.006. Switching to student who register after the size assignment day, we can reject effect sizes outside the range -0.02 to 0.02. Confidence intervals for persistence outcomes are of a similar order of magnitude.[6] In short, changes in class sizes had small effects on outcomes in these online college courses.

To contextualize these effect sizes in a different way, we can easily rule out effects typically found in the higher education class size literature. For example, Bandiera, Larcinese, and Rasul (2010) found that a one standard deviation increase in university class size reduced end of year test scores by 0.108 student standard deviations. Similarly, De Giorgi, Pellizzari, and Woolston (2012) found that a standard deviation increase in class size (approximately 20 students in classes that on average contained 131 students) reduced grades by 0.140 standard deviations. Our estimated effects are also significantly smaller than the effects of other higher education interventions on grades. In their 2009 paper Angrist, Lang, and Oreopoulos found that assigning college students a fellowship and support services had an effect size on fall grades of

[6] It is possible that a professor that teaches more than one section in a given session may substitute effort and time between sections as class size increases. We test this proposition in three ways. First, we interact the total number of students a professor taught in the current session with the class size treatment variable. We also separate the total number of students taught in online classes and offline classes in case the substitution effect differed by the types of classes each professor taught. Second, we estimate Table 3 Panel D separately for a sample of professors that only taught one section in the current session and professors that taught more than one section. Third, we calculate the average section size across all courses taught by a given professor and instrument it with the treatment indicator for small classes. The results are presented in the appendix, Tables A1-A3. There is no consistent significant result found in any model indicating that professors do not seem to be substituting time or effort across classes as a result of the increase in class size.

0.227 standard deviations, most of which was concentrated on women where the effect size was 0.346 standard deviations. The confidence intervals in Table 4 preclude estimates close to the magnitude of estimates found in any of these studies.

The findings are especially salient because they are at odds with the theoretical prediction that larger classes and lower-achieving peers generate negative effects. For students who registered before the cap changed and had stronger previous academic records, we would expect both the increased class size and the introduction of lower performing peers to negatively affect their outcomes. The negative effects should be reinforcing. Later registrants would also potentially be negatively affected by larger class sizes. The estimates suggest that both the class size and the peers had negligible effects on the students.[7]

5. Heterogeneity of Results

The effect of class size on student outcomes need not be constant across all classes. To test whether there is any heterogeneity in the class size effect we segmented the sample on two dimensions, and estimated effects within each characteristic group. First, we divided the sample by academic discipline and separated courses into subject groupings: business and marketing, computer science and technology, general education, humanities, science and mathematics, and social sciences. Second, we divided the sample by the types of assignments each class required of the students: project only, laboratories only, both projects and laboratories, or neither projects nor laboratories.

[7] We emphasize the theoretical predictions here in that our results are at odds with them. One could imagine other stories that could explain the null findings. For example, if there exists some mechanism by which large classes improve outcomes, it could offset negative peer effects. Given that theoretically positive impacts of large classes are implausible, we have not accentuated the potential offset. We did test for the possibility of positive impacts in one way. In the assignment mechanism, the peer effects should diminish in their importance as the contrast between incumbents and late registrants decreases. Hence, early sections should have both peer and class size impacts while later sections should have only class size impacts. We find no difference in the estimated impacts when we allow for different effects for early and late sections.

Table 5 shows the covariate balance exercise for each of these different sub-samples. Though there are a few covariates that are significantly different than zero at the 5 percent level, the number of tests is large. The most imbalanced sub-sample is for those classes that require both project and laboratory oriented courses. Even in this sample we cannot reject the null hypothesis that all coefficients are jointly equal to zero (p=0.2363 for the 15 student group fixed effect approach). As we show in Tables 6 and 7 we fail to find a significant effect of class size on student outcomes despite these imbalances.

Tables 6 and 7 show the estimated effects of class size increases on courses in different disciplines and with different assignment types, respectively. These models are analogous to Table 3 Panel D where student outcomes are regressed on the log of intended class size interacted with an indicator variable for registering after the class size cap was changed. In almost all cases there is no significant effect of class size on student outcomes. Almost all categories follow a familiar pattern where there is a small, negative insignificant effect on students who registered before the class cap was changed, and a small, insignificant positive effect on those who registered after the cap was changed. One notable exception is in the social sciences where we find a consistently significant, positive effect on students who registered after the cap size was changed. In the 15 student group fixed effect model the total effect for a ten percent increase in class size is to increase student grades by 0.057 grade points (on a scale from 0-4). There is no effect, however, on enrollment in the next term.

In total, there is little effect heterogeneity by subject or assignment type. Figure 1 summarizes the point estimates on all three student outcomes with the 15 student group fixed effect model. Effects are separated for those who registered before and after the class size cap was changed. All three outcomes are precisely measured in the full sample. The point estimates remain close to zero for all other subsamples, though due to the smaller sample size, the confidence intervals can be large in some cases. Nevertheless, there is little to no evident class size effect, except perhaps for students who registered after the cap size changed in social science courses.

6. Robustness Checks

The estimates in Table 3 are quite robust to alternative choices for registration group size in the course-by-term-by-registration-group fixed effects. Table 3, and all other tables, shows results for registration groups of 15 consecutively registering students and for vigintile groups of registration date and time. To illustrate the stability of our results relative to these choices, Figure 2 shows point estimates and standard errors for a range of choices: 5 to 50 student groups and 10 to 50 quantiles.

Some may question the appropriateness of course grade as an outcome. Grades are to a certain extent at the discretion of the professor but they are consequential for students. Degrees are only awarded after students have passed their required classes, and GPA predicts persistence in college (Desjardins, Ahlburg, and McCall 1999) and predicts post-college earnings (Jones and Jackson 1990; Loury and Garman 1995). From this perspective, the effect of class size on grades is relevant. Our null results, however, could arise by construction if professors were assigning grades based on a forced distribution. For example, if professors give a certain percentage of their classes As, Bs, Cs, and failing grades regardless of their composition and size, then an increase in class size will do nothing to affect aggregate grades. DeVry works against this practice by providing detailed grading rubrics for all courses. However, we test in two ways the possibility that professors grade on the same curve in all classes. First, to see whether there is a university-wide grading rule we simply create a histogram of the average section grade. If professors were asked to grade on a university-wide forced distribution, then the average section grades would be similar. Figure 3 shows that there is a distribution in the average section GPA across the university, with a variance of 0.282 grade points. We thus find little evidence for a university-wide forced distribution.

This result, however, does not preclude the possibility that a significant number of professors engage in such grading schemes, which would attenuate the effect of class size on grade outcomes. To explore grading practices further, we regress a section's average outcome grade on the section's average prior cumulative GPA. We would expect that overall average prior GPA would predict current year GPA if grading were not curved within section. If, however, professors were using a forced grading distribution scheme, this relationship should disappear after introducing professor fixed effects into the model. Table 8 shows the results for average GPA, and the average probability of receiving an A or above, B or above, and C or above. Without any fixed effects, there is a positive and significant relationship between the two variables, as expected. After including professor fixed effects the point estimates remain relatively stable and significant, thus indicating that most, if not all, professors do not grade on a forced distribution.

To account for the possibility that professors may change the forced distribution between sessions or between courses, a third model includes session-by-course-by professor fixed effects. Once again the point estimates remain relatively stable though their significance disappears. We do not believe the lack of significance stems from the professors grading on a forced distribution. Rather, there are few professors who teach more than one section of a given course in a given session. The sample size is therefore effectively reduced, and the standard errors increase as a result. The stable point estimates in concert with the larger standard errors account for the lack of significance of the point estimates. Overall then, there is no evidence that a forced distribution of grades is at play that would account for our null results. It is also worth noting that our null grade results are also consistent with the null results for the persistence measures, which would also be unaffected by a forced grade distribution. Overall, our initial conclusion remains that there is little effect of class size on student outcomes.

7. Conclusions and Policy Implications

In this paper we present evidence of the effect of online class size on a variety of student outcomes. For online classes with an average of 30 students, increasing the class size ten percent does not significantly affect student grades, enrollment in the next term, or credits attempted the next term. We can confidently rule out effect sizes similar to the prior literature in both K-12 and in higher education.

At first one may find these results unsurprising. Some may argue that online classes are meant to make education accessible to more people by removing physical barriers and reducing the dependence on communication with the professor and peers. The increased temporal and physical flexibility comes at the cost of interacting as closely with others. In this view the mechanisms posited by Lazear by which class size could affect student outcomes would be muted, allowing classes to grow without changing the quality of the product. Indeed this mechanism is the hope behind MOOCs. The goal for DeVry, however, was not to create MOOCs which have average completion rates of about 7 percent (Jordan 2014), but rather to recreate the traditional classroom delivered over the internet. These online classes are meant to provide temporal and physical flexibility for the students, while preserving the interaction among students and between students and faculty. Students are required to post and respond to other posts, while professors actively engage in online conversations. If class size could affect online performance, it would be in these types of classes where disruptive or congestive effects as predicted in Lazear's model are possible (see Bettinger, Loeb and Taylor 2015).

As it turns out, these interactions may be muted enough in virtual classrooms to accommodate financially relevant changes in class size without degrading outcomes for students. Even in this context it is more difficult for students to distract each other while the professor is lecturing, and students have more flexibility in reading peer posts than in listening to peer discussion during the class. Even in classes where peer-to-peer interactions would be plausibly more important, such as classes that include projects and laboratories, students surprisingly are not affected by the ten percent change in class size. The online classes may have reduced professor-student interactions, or made them more efficient, such that increasing class sizes may not necessarily reduce professor availability. For example professors may be able to answer all questions and to respond to all posts, whether there are 30 or 33 students in the class.

These results have policy and financial implications. Virtual classrooms have the advantage of making courses more widely available to students who have difficulty with the location or timing of in-person classrooms. Prospective students who live too far from a university, or whose schedules do not fit with traditional school schedules, can use online classes to access previously unavailable content. Virtual courses also do not have physical plant limitations that can limit the number of people who can enroll in a course at a time. Thus, schools can respond to changes in demand more quickly in the online setting. As long as the institution can preserve the quality of the product, online courses have the potential to broaden access to higher education. This study gives the first evidence that small but meaningful increases in class sizes in the online context may not degrade the quality of the class.

As we stated earlier, DeVry offers two-thirds of its courses online. They can dramatically increase the access to their offerings simply by adding two to five students to each of their thousands of online sections. If these results are generalizable to other for-profit institutions with a large online presence, such as the University of Phoenix, these already large gains in access to higher education can be amplified even further. The sheer number of online sections means that small increases in class sizes can result in large increases in aggregate enrollment – all plausibly with little or no reduction in student outcomes.

These results also have financial implications for universities and students. A 10 percent increase in class size is equivalent, in most all cases, to a 10 percent reduction in the number of course sections DeVry must offer and staff with a professor. The associated cost savings would come primarily from professor compensation; the marginal cost of adding a section is negligible. Assuming the section wage rate is constant for the class size increases in this range, and assuming that enrollment is not a function of class size, a 10 percent class size increase would save the university 10 percent in salary expenses. DeVry could use the cost saving to improve the quality of classes in other ways through higher wages for professors or greater investments in curriculum and instruction. A caveat to this analysis is that this study can only estimate the short-term effects of increasing class size and not the long term effects. In the long run, larger classes may affect who DeVry can hire as instructors and their compensation thereby eroding some of these financial gains.

Though the class size increases in this study are significant in financial terms, the main limitation of the study is that we test small changes in class sizes relative to most educational proposals. The results are unlikely to be applicable to large changes in class size such as doubling or tripling current numbers. It is worth noting, that the changes we assess, while small, are not significantly smaller in magnitude than some changes in Angrist and Lavy's (1999) seminal class size paper. Though Maimonides' rule in Israel generated large swings in class size, the authors concluded that smaller schools drove their results. For fifth graders in their discontinuity sample, larger classes in their instrumental variable estimation had 10.7 students, 4.4 students, and 1.1 more students in the first, second, and third enrollment segments respectively. The second and third wave of enrollment segments have class size changes of similar magnitude to our study. In addition the authors provide a *per pupil* effects size of 0.036 SD for the full sample, 0.071SD for the discontinuity sample, and 0.017-0.019SD for the fourth grade sample. Table 4 illustrates that the largest effect size in the 95% confidence interval of our experiment is a decrease in student outcomes by 0.035SD. When dividing this by the average increase in class size of 3 students, the analogous per pupil effect size is 0.012SD. Therefore though the Israeli study had some class size swings of the order of magnitude of our study, the greatest possible effect of our intervention is smaller than the smallest significant effect seen in their study.

References

Allen, I. E., & Seaman, J. (2013). *Changing Course: Ten Years of Tracking Online Education in the United States*. Sloan Consortium. Newburyport, MA. Retrieved From: http://eric.ed.gov/?id=ED541571

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., Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholasticachievement. *Quarterly Journal of Economics*, *114*, 533-575.

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.

Becker, W. E., & Powers, J. R. (2001). Student performance, attrition, and class size given missing student data. *Economics of Education Review*, *20*(4), 377-388.

Bettinger, E., Loeb, S., & Taylor, E. (2015). Remote but Influential: Peer Effects in Online Higher Education Classrooms. Center for Education Policy Analysis Working Paper.

Chetty, R., Friedman, J. N., Hilger, N., Saez, E., Schanzenbach, D. W., & Yagan, D. (2010). *How does your kindergarten classroom affect your earnings? Evidence from Project STAR* (No. w16381). National Bureau of Economic Research.

De Giorgi, G., Pellizzari, M., & Woolston, W. G. (2012). Class size and class heterogeneity. *Journal of the European Economic Association*, *10*(4), 795-830.

Desjardins, S.L., Ahlburg, D.A., McCall, B.P. (1999). An event history model of student departure. *Economics of Education Review, 18*(3), 275-390.

Dillon, M., Kokkelenberg, E. C., & Christy, S. M. (2002). The effects of class size on student achievement in higher education: applying an earnings function. *Cornell Higher Education Research Institute (CHERI)*, 15.

Ding, W., & Lehrer, S. F. (2010). Estimating treatment effects from contaminated multi-period education experiments: the dynamic impacts of class size reductions. *The Review of Economics and Statistics*, *92*(1), 31-42.

Dynarski, S., Hyman, J. M., & Schanzenbach, D. W. (2011). *Experimental evidence on the effect of childhood investments on postsecondary attainment and degree completion* (No. w17533). National Bureau of Economic Research.

Gary-Bobo, R. J., & Mahjoub, M. B. (2013). Estimation of Class-Size Effects, Using "Maimonides' Rule" and Other Instruments: The Case of French Junior High Schools *Annales d'Economie et de Statistique*, (111-112).

Hanushek, E. A. (2002). Evidence, politics, and the class size debate. *The class size debate*, 37-65, Retrieved From: http://hanushek.stanford.edu/sites/default/files/publications/ Hanushek%202002%20ClassSizeDebate.pdf

Hoxby, C. M. (2000). The effects of class size on student achievement: New evidence from population variation. *Quarterly Journal of economics*, 1239-1285.

Johnson, I.Y. (2010). Class size and student performance at a public research university: A cross-classified model. *Research in Higher Education, 51(8), 701-723*

Jones, E. B., & Jackson, J. D. (1990). College grades and labor market rewards. *Journal of Human Resources*, 253-266.

Jordan, K. (2014). Initial trends in enrolment and completion of massive open online courses. *The International Review of Research in Open and Distributed Learning*, *15*(1).

Kennedy, P. E., & Siegfried, J. J. (1997). Class size and achievement in introductory economics: Evidence from the TUCE III data. *Economics of Education Review*, *16*(4), 385-394.

Krueger, A. B. (2002). Understanding the magnitude and effect of class size on student achievement. *The class size debate*, 7-35.

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.

Lazear, E. (1999). *Educational production* (No. w7349). National bureau of economic research.

Lazear, E. (2001). "Educational Production." The Quarterly Journal of Economics, 16(3), 777- 803.

Leuven, E., Oosterbeek, H., & Rønning, M. (2008). Quasi‐experimental Estimates of the Effect of Class Size on Achievement in Norway. *The Scandinavian Journal of Economics*, *110*(4), 663-693.

Loury, L. D., & Garman, D. (1995). College selectivity and earnings. *Journal of Labor Economics*, 289-308.

McKeachie, W. J. (1986). *Teaching and learning in the college classroom: A review of the research literature* (Vol. 86). University of Michigan Press.

Mosteller, F. (1995). The Tennessee study of class size in the early school grades. *The future of children*, 113-127.

Siegfried, J. J., & Kennedy, P. E. (1995). Does pedagogy vary with class size in introductory economics?. *The American Economic Review*, 347-351.

Topp, Robert F. (1998) "Education of an Undergraduate: Action Before Reaction." *College and University,* 45(2), pp. 123-127.

US Department of Education. (2014). Enrollment in distance education courses, by state: Fall 2012. Retrieved from: http://nces.ed.gov/pubs2014/2014023.pdf

Wöessmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review*, *50*(3), 695-736.

A. Course grade

B. Enrolled the next term

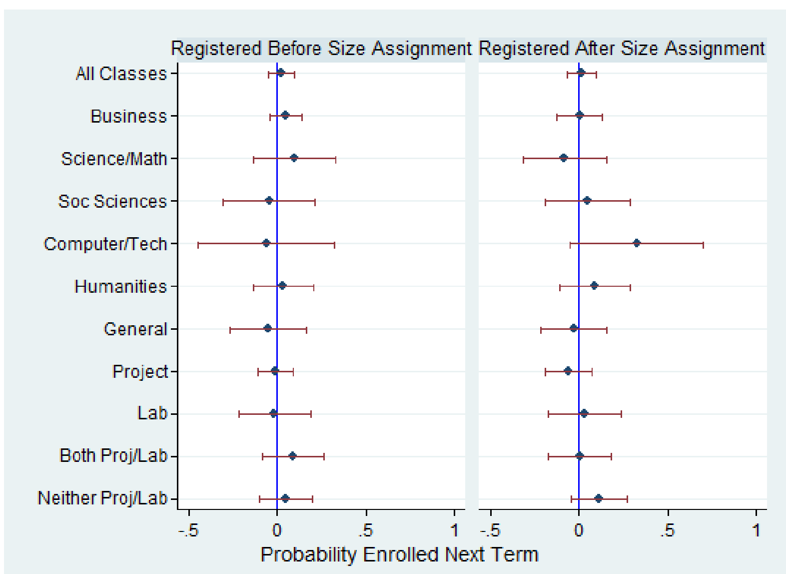


Figure 1—Effect of class size overall and by course characteristics:
Coefficient on log of intended class size and 95 percent confidence intervals

C. Units enrolled the next term

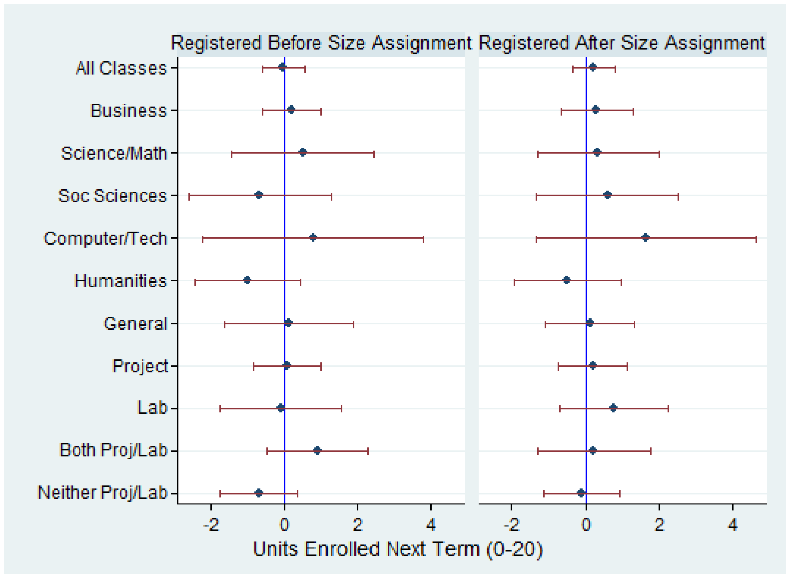
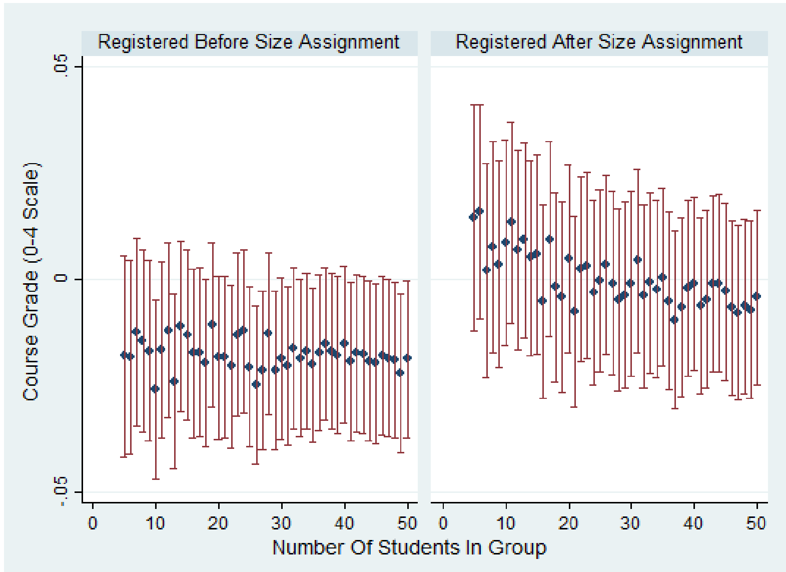


Figure 1 (cont.)—Effect of class size overall and by course characteristics:
Coefficient on log of intended class size and 95 percent confidence intervals

Note: Point estimates and confidence intervals depicted in the figure are taken from OLS (Panel A and C) or LPM (Panel B) regressions using student-by-course observations. Estimation is identical to Table 3 Panel D, but carried out separately for each sub-sample defined by row headers. The pair of estimates in each row come from a single regression. Course-by-term-by-registration-group fixed effects with registration-group is defined by 15 students with sequential registration date-time. Standard errors allow for clustering in sections.

A. Registration-group defined by N students with sequential registration date-time



B. Registration-group defined by quantiles of registration date-time

Figure 2—Class size effect on course grade, estimated with varying registration-group size in course-by-term-by-registration-group fixed effects: Coefficient on log of intended class size and 95 percent confidence intervals

Note: Point estimates and confidence intervals depicted in the figure are taken from OLS regressions using student-by-course observations. Estimation is identical to Table 3 Panel D, but with varying definitions of registration-group in the course-by-term-by-registration-group fixed effects. The pair of estimates for after and before size assignment, for a given x-axis values, come from a single regression. Standard errors allow for clustering in sections.

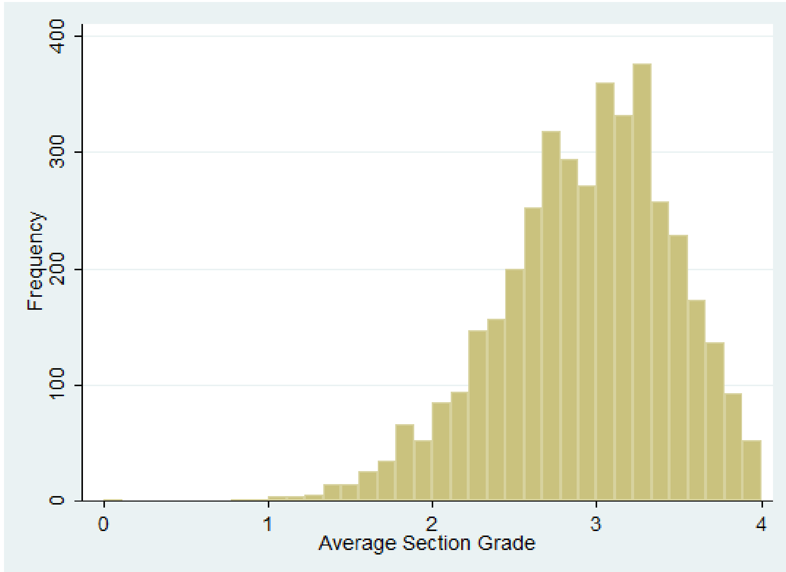


Figure 3—Histogram of average section grade

Table 1--Descriptive Statistics

	All student-by-course observations			Registered before class size assignment		Registered after class size assignment	
	All sections	Large sections	Small sections	Large sections	Small sections	Large sections	Small sections
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Student-by-course observations	102,209	52,462	49,747	26,652	27,046	25,810	22,701
(A) Section characteristics, student-weighted mean (st. dev.)							
Assigned to be "small" section	0.487	0	1	0	1	0	1
Intended class size	32.144	33.625	30.583	33.77	30.886	33.476	30.221
	(4.48)	(4.32)	(4.10)	(4.13)	(3.90)	(4.50)	(4.29)
ln(intended class size)	3.459	3.506	3.41	3.511	3.421	3.501	3.397
	(0.15)	(0.14)	(0.15)	(0.13)	(0.14)	(0.15)	(0.16)
Subject							
Business/management	0.355						
Science/mathematics	0.151						
Computer/technology	0.094						
Social sciences	0.128						
Humanities	0.171						
General	0.101						
Assignment characteristics							
Projects only	0.261						
Labs only	0.235						
Both projects and labs	0.147						
Neither projects nor labs	0.358						
(B) Student characteristics, mean (st. dev.)							
Cumulative prior GPA	3.173	3.173	3.175	3.254	3.254	3.069	3.055
	(0.78)	(0.79)	(0.78)	(0.74)	(0.74)	(0.83)	(0.84)
Failed a class previously	0.221	0.221	0.222	0.205	0.208	0.236	0.238
New student	0.091	0.092	0.090	0.036	0.034	0.150	0.156
Missing prior GPA	0.058	0.058	0.058	0.020	0.020	0.097	0.102

Note: Authors' calculations. Courses included in the class size trial. July, September, November 2013 and January 2014 sessions.

Table 2--Covariate balance

	Prior GPA (0-4, F=0 and A=4)		Failed a prior class (binary)		New student (binary)		Missing prior GPA (binary)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A) All students								
Small class (binary)	0.004	0.008	0.000	-0.001	-0.001	0.000	-0.001	0.000
	(0.006)	(0.005)	(0.003)	(0.003)	(0.001)	(0.002)	(0.001)	(0.001)
ln(intended class size)	-0.060	-0.089+	0.017	0.023	0.004	-0.005	0.011	0.010
	(0.053)	(0.051)	(0.026)	(0.025)	(0.015)	(0.016)	(0.013)	(0.014)
Student-by-course observations	87,006	87,006	102,209	102,209	102,209	102,209	102,209	102,209
(B) Students who registered before size assignment								
Small class (binary)	-0.001	0.003	0.001	0.000	0.001	0.001	-0.001	0.000
	(0.007)	(0.007)	(0.004)	(0.003)	(0.001)	(0.001)	(0.001)	(0.001)
ln(intended class size)	-0.005	-0.048	0.000	0.015	-0.013	-0.010	0.010	0.005
	(0.066)	(0.064)	(0.035)	(0.033)	(0.012)	(0.014)	(0.010)	(0.011)
Student-by-course observations	50,721	50,721	53,698	53,698	53,698	53,698	53,698	53,698
(C) Students who registered after size assignment								
Small class (binary)	0.010	0.013	-0.002	-0.002	-0.003	-0.001	0.000	0.000
	(0.009)	(0.009)	(0.004)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)
ln(intended class size)	-0.132	-0.131	0.040	0.030	0.021	-0.001	0.011	0.011
	(0.087)	(0.083)	(0.039)	(0.036)	(0.028)	(0.029)	(0.026)	(0.026)
Student-by-course observations	36,285	36,285	48,511	48,511	48,511	48,511	48,511	48,511
Course-by-term-by-registration-group fixed effects, registration-group defined by…								
15 students grouped by date-time	√		√		√		√	
vigintiles of date-time		√		√		√		√

Note: Each cell reports an estimate from a separate OLS or LPM regression using student-by-course observations. Dependent variables are described in the column headers. Each regression includes a treatment measure and fixed effects defined by the intersection of course, term, and registration date-time group. The treatment variables "small class" and "ln(intended class size)" are based on the class size assigned by DeVry University to each section, not the actual enrollment. Registration date-time group is defined by, first, ordering students by the date-time they registered for the course; and, second, dividing students into either (i) groups of 15 students with sequential registration date-time, or (ii) vigintiles of registration date-time. Standard errors allow for clustering in sections.

+ indicates $p < 0.10$, * < 0.05, ** < 0.01

Table 3--Effect of class size on student achievement and persistence (intent to treat estimates)

	Course grade (0-4, F=0 and A=4)				Enrolled next term (binary)		Units enrolled next term (0-20)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A)								
Small class	-0.002	0.005	-0.008	0.000	-0.003	-0.001	-0.021	-0.015
	(0.009)	(0.009)	(0.011)	(0.011)	(0.003)	(0.003)	(0.022)	(0.021)
(B)								
Small class	0.002	0.009	0.004	0.012	-0.003	-0.002	-0.015	-0.002
	(0.011)	(0.011)	(0.014)	(0.014)	(0.004)	(0.004)	(0.030)	(0.029)
Small class	-0.009	-0.008	-0.025	-0.025	0.001	0.001	-0.013	-0.026
* registered after size assignment	(0.015)	(0.016)	(0.019)	(0.020)	(0.006)	(0.006)	(0.043)	(0.042)
(C)								
ln(intended class size)	-0.036	-0.100	0.020	-0.067	0.017	0.003	0.104	0.059
	(0.088)	(0.088)	(0.115)	(0.115)	(0.029)	(0.028)	(0.220)	(0.213)
(D)								
ln(intended class size)	-0.130	-0.213*	-0.151	-0.241+	0.022	0.027	-0.022	0.068
	(0.103)	(0.103)	(0.133)	(0.134)	(0.036)	(0.035)	(0.294)	(0.279)
ln(intended class size)	0.190	0.229+	0.335+	0.342*	-0.008	-0.048	0.252	-0.026
* registered after size assignment	(0.137)	(0.134)	(0.179)	(0.167)	(0.052)	(0.048)	(0.388)	(0.364)
Course-by-term-by-registration-group fixed effects, registration-group defined by…								
15 students grouped by reg. date-time	√		√		√		√	
vigintiles of reg. date-time		√		√		√		√
July, Sept. and Nov. terms only			√	√	√	√	√	√
Student-by-course observations	102,209	102,209	67,858	67,858	67,858	67,858	67,858	67,858

Note: Each column within each panel reports estimates from a separate OLS or LPM regression using student-by-course observations. Dependent variables are described in the column headers. The treatment variables "small class" and "ln(intended class size)" are based on the class size assigned by DeVry University to each section, not the actual enrollment. In addition to the covariates shown above, Panels B and D also include a main effect for "after size assignment," an indicator = 1 if the student registered after DeVry made the class size assignments to course sections. All regressions include prior cumulative GPA, in indicator = 1 if the student has failed a course previously, an indicator = 1 if the student is new to DeVry, and an indicator = 1 if prior GPA is missing for a continuing student. Prior GPA is set to zero for new students and students missing prior GPA. All regressions include fixed effects defined by the intersection of course, term, and registration date-time group. Registration date-time group is defined by, first, ordering students by the date-time they registered for the course; and, second, dividing students into either (i) groups of 15 students with sequential registration date-time, or (ii) vigintiles of registration date-time. Outcome for "next term" are not observed for the January term, and Jan. observations are only included in Columns 1 and 2. Standard errors allow for clustering in sections.
+ indicates $p < 0.10$, * < 0.05, ** < 0.01

Table 4--Impact of a 10 percent increase in class size, 95 percent confidence intervals in student standard deviation units

	Course grade (0-4, F=0 and A=4)		Enrolled next term (binary)		Units enrolled next term (0-20)	
	(1)	(2)	(3)	(4)	(5)	(6)
Students who registered *before* class size assignment	[-0.028,0.006]	[-0.035,-0.001]	[-0.014,0.026]	[-0.012,0.027]	[-0.021,0.020]	[-0.017,0.022]
Students who registered *after* class size assignment	[-0.013,0.022]	[-0.016,0.019]	[-0.018,0.025]	[-0.026,0.015]	[-0.013,0.029]	[-0.019,0.022]
Course-by-term-by-registration-group fixed effects, registration-group defined by…						
15 students grouped by reg. date-time	√		√		√	
vigintiles of reg. date-time		√		√		√

Note: 95 percent confidence interval for a 10 percent increase in class size calculated with the estimates in Table 3 Panel D. Reported here in student standard deviation units. Course grade intervals use the full sample estimates.

Table 5--Covariate balance by course characteristic sub-samples
Treatment measure = ln(intended class size)

	Prior GPA (0-4, F=0 and A=4)		Failed a prior class (binary)		New student (binary)		Missing prior GPA (binary)		Obsv.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(A) Course subject									
Business/management	-0.043	-0.091	-0.018	0.010	0.039+	0.024	0.015	0.004	36,278
	(0.077)	(0.077)	(0.048)	(0.047)	(0.021)	(0.022)	(0.011)	(0.013)	
Computer science/technology	0.277	0.200	-0.165	-0.127	-0.051	-0.024	-0.048	0.004	9,584
	(0.269)	(0.273)	(0.112)	(0.111)	(0.065)	(0.071)	(0.076)	(0.077)	
General classes	-0.472*	-0.181	0.053	0.025	0.011	-0.006	-0.015	-0.021	10,311
	(0.230)	(0.213)	(0.058)	(0.053)	(0.071)	(0.073)	(0.049)	(0.050)	
humanities	0.008	-0.126	0.057	0.051	-0.049+	-0.060*	-0.002	-0.001	17,514
	(0.117)	(0.115)	(0.060)	(0.063)	(0.029)	(0.029)	(0.031)	(0.028)	
Science/mathematics	-0.231+	-0.193	0.121+	0.082	-0.017	-0.014	0.019	0.030	15,420
	(0.136)	(0.124)	(0.064)	(0.062)	(0.026)	(0.025)	(0.037)	(0.037)	
Social sciences	0.114	0.077	-0.049	-0.005	0.024	0.010	0.068	0.054	13,102
	(0.168)	(0.151)	(0.069)	(0.057)	(0.029)	(0.028)	(0.051)	(0.050)	
(B) Course assignment types									
Project only	-0.053	-0.073	-0.003	0.009	0.021	0.003	0.019	0.011	26,653
	(0.091)	(0.089)	(0.043)	(0.040)	(0.031)	(0.036)	(0.022)	(0.024)	
Laboratory only	-0.103	-0.101	0.068	0.044	-0.028	-0.023	0.014	0.038	24,023
	(0.128)	(0.117)	(0.060)	(0.057)	(0.026)	(0.026)	(0.034)	(0.035)	
Both projects and laboratories	-0.256*	-0.249*	0.103	0.106+	-0.007	-0.018	-0.014	-0.012	14,993
	(0.120)	(0.114)	(0.067)	(0.061)	(0.033)	(0.030)	(0.021)	(0.020)	
Neither projects nor laboratories	0.073	-0.002	-0.037	-0.018	0.012	0.003	0.013	0.000	36,540
	(0.095)	(0.096)	(0.048)	(0.047)	(0.023)	(0.025)	(0.027)	(0.028)	
Course-by-term-by-registration-group fixed effects, registration-group defined by…									
15 students grouped by reg. date-time	√		√		√		√		
vigintiles of reg. date-time		√		√		√		√	

Note: Each cell reports an estimate from a separate OLS or LPM regression using student-by-course observations. Estimation identical to Table 2, but carried out separately for each sub-sample defined by course subject or assignment characteristics. Standard errors allow for clustering in sections.
+ indicates p < 0.10, * < 0.05, ** < 0.01

Table 6--Effect of class size on student achievement and persistence, by course subject (intent to treat estimates)

		Course grade (0-4, F=0 and A=4)				Enrolled next term (binary)		Units enrolled next term (0-20)		Obsv.
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Business/ marketing	ln(intended class size)	-0.177	-0.187	-0.056	-0.145	0.046	0.036	0.206	0.354	36,278
		(0.157)	(0.159)	(0.195)	(0.194)	(0.047)	(0.048)	(0.408)	(0.427)	
	ln(intended class size)	0.268	0.254	0.212	0.379	-0.041	-0.021	0.098	0.124	
	* registered after size assign.	(0.212)	(0.211)	(0.279)	(0.264)	(0.074)	(0.073)	(0.630)	(0.668)	
Computer/ technology	ln(intended class size)	-0.032	-0.247	0.500	0.431	-0.063	-0.047	0.802	0.532	9,584
		(0.385)	(0.363)	(0.570)	(0.547)	(0.195)	(0.192)	(1.538)	(1.571)	
	ln(intended class size)	0.262	0.336	0.443	0.178	0.392	0.314	0.865	0.495	
	* registered after size assign.	(0.596)	(0.627)	(0.862)	(0.883)	(0.284)	(0.293)	(2.202)	(2.218)	
General	ln(intended class size)	-0.41	-0.577+	-0.646	-0.950*	-0.052	-0.111	0.133	-0.506	10,311
		(0.360)	(0.343)	(0.433)	(0.425)	(0.110)	(0.099)	(0.894)	(0.782)	
	ln(intended class size)	0.48	0.565	0.983+	1.041*	0.023	0.098	-0.009	0.785	
	* registered after size assign.	(0.435)	(0.380)	(0.525)	(0.473)	(0.140)	(0.123)	(1.032)	(0.894)	
Humanities	ln(intended class size)	-0.261	-0.181	-0.436	-0.248	0.032	0.053	-1.002	-0.956	17,514
		(0.260)	(0.267)	(0.332)	(0.347)	(0.086)	(0.089)	(0.732)	(0.747)	
	ln(intended class size)	-0.176	-0.162	-0.132	-0.426	0.058	-0.036	0.506	-0.049	
	* registered after size assign.	(0.312)	(0.314)	(0.390)	(0.393)	(0.126)	(0.120)	(0.844)	(0.905)	
Science/ mathematics	ln(intended class size)	0.025	0.058	-0.244	-0.065	0.096	0.08	0.51	0.222	15,420
		(0.257)	(0.248)	(0.382)	(0.377)	(0.117)	(0.108)	(0.980)	(0.872)	
	ln(intended class size)	-0.162	-0.301	0.489	0.303	-0.178	-0.176	-0.157	0.185	
	* registered after size assign.	(0.374)	(0.362)	(0.534)	(0.542)	(0.171)	(0.161)	(1.323)	(1.215)	
Social sciences	ln(intended class size)	0.165	-0.167	-0.091	-0.282	-0.046	-0.08	-0.66	-0.869	13,102
		(0.311)	(0.275)	(0.420)	(0.386)	(0.132)	(0.110)	(0.985)	(0.790)	
	ln(intended class size)	0.762*	0.739*	1.108*	0.939+	0.095	0.09	1.259	1.545	
	* registered after size assign.	(0.388)	(0.367)	(0.549)	(0.516)	(0.168)	(0.150)	(1.354)	(1.156)	
Course-by-term-by-registration-group fixed effects, registration-group defined by…										
15 students grouped by reg. date-time		√		√		√		√		
vigintiles of reg. date-time			√		√		√		√	
July, Sept. and Nov. terms only				√	√	√	√	√	√	

Note: Each column within each panel reports estimates from a separate OLS or LPM regression using student-by-course observations. Estimation identical to Table 3, but carried out separately for each sub-sample defined by course subject. Outcome for "next term" are not observed for the January term, and Jan. observations are only included in Columns 1 and 2. Standard errors allow for clustering in sections.

+ indicates p < 0.10, * < 0.05, ** < 0.01

Table 7--Effect of class size on student achievement and persistence, by course assignment characteristics (intent to treat estimates)

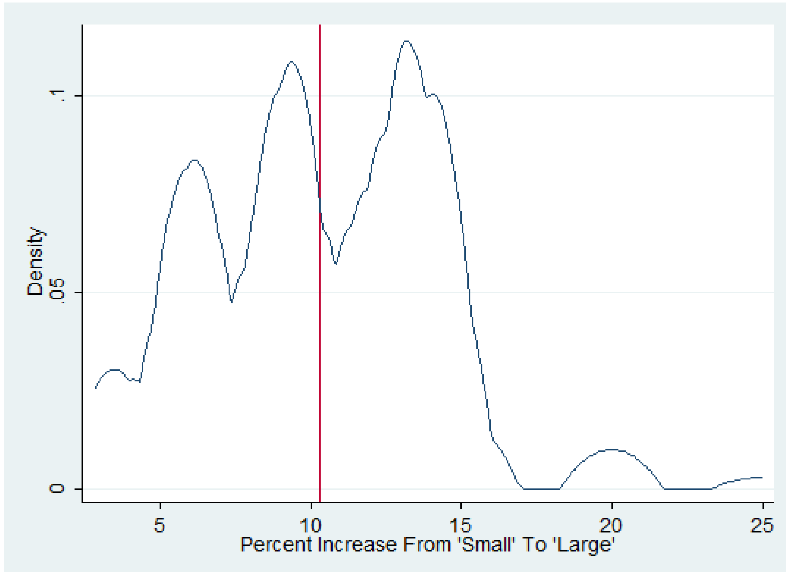
		Course grade (0-4, F=0 and A=4)				Enrolled next term (binary)		Units enrolled next term (0-20)		Obsv.
		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Projects only	ln(intended class size)	-0.2	-0.302+	-0.226	-0.395+	-0.01	-0.041	0.073	-0.2	26,653
		(0.173)	(0.174)	(0.210)	(0.213)	(0.052)	(0.051)	(0.469)	(0.444)	
	ln(intended class size)	0.419+	0.381+	0.672*	0.610*	-0.049	0.005	0.135	0.362	
	* registered after size assign.	(0.245)	(0.227)	(0.302)	(0.277)	(0.081)	(0.075)	(0.649)	(0.596)	
Labs only	ln(intended class size)	-0.197	-0.212	-0.093	0.017	-0.018	0.009	-0.088	-0.091	24,023
		(0.228)	(0.218)	(0.326)	(0.322)	(0.103)	(0.097)	(0.845)	(0.784)	
	ln(intended class size)	0.075	0.035	0.29	0.201	0.047	-0.006	0.871	0.733	
	* registered after size assign.	(0.320)	(0.315)	(0.456)	(0.461)	(0.151)	(0.141)	(1.141)	(1.079)	
Both projects and labs	ln(intended class size)	-0.138	-0.086	-0.002	-0.098	0.087	0.079	0.91	0.593	14,993
		(0.247)	(0.239)	(0.366)	(0.357)	(0.088)	(0.081)	(0.695)	(0.633)	
	ln(intended class size)	0.147	0.025	0.213	0.223	-0.083	-0.074	-0.682	-0.015	
	* registered after size assign.	(0.318)	(0.300)	(0.432)	(0.401)	(0.113)	(0.105)	(0.935)	(0.840)	
Neither projects nor labs	ln(intended class size)	0.032	-0.042	-0.192	-0.182	0.049	0.086	-0.686	0.018	36,540
		(0.201)	(0.207)	(0.256)	(0.268)	(0.077)	(0.075)	(0.537)	(0.531)	
	ln(intended class size)	-0.011	0.043	0.065	-0.002	0.068	-0.08	0.597	-0.553	
	* registered after size assign.	(0.251)	(0.263)	(0.331)	(0.337)	(0.107)	(0.098)	(0.676)	(0.670)	
Course-by-term-by-registration-group fixed effects, registration-group defined by…										
15 students grouped by reg. date-time		√		√		√		√		
vigintiles of reg. date-time			√		√		√		√	
July, Sept. and Nov. terms only				√	√	√	√	√	√	

Note: Each column within each panel reports estimates from a separate OLS or LPM regression using student-by-course observations. Estimation identical to Table 3, but carried out separately for each sub-sample defined by course assignment characteristics. Outcome for "next term" are not observed for the January term, and Jan. observations are only included in Columns 1 and 2. Standard errors allow for clustering in sections.
+ indicates p < 0.10, * < 0.05, ** < 0.01

Table 8--The distribution of assigned grades within sections and prior grades of students in the section

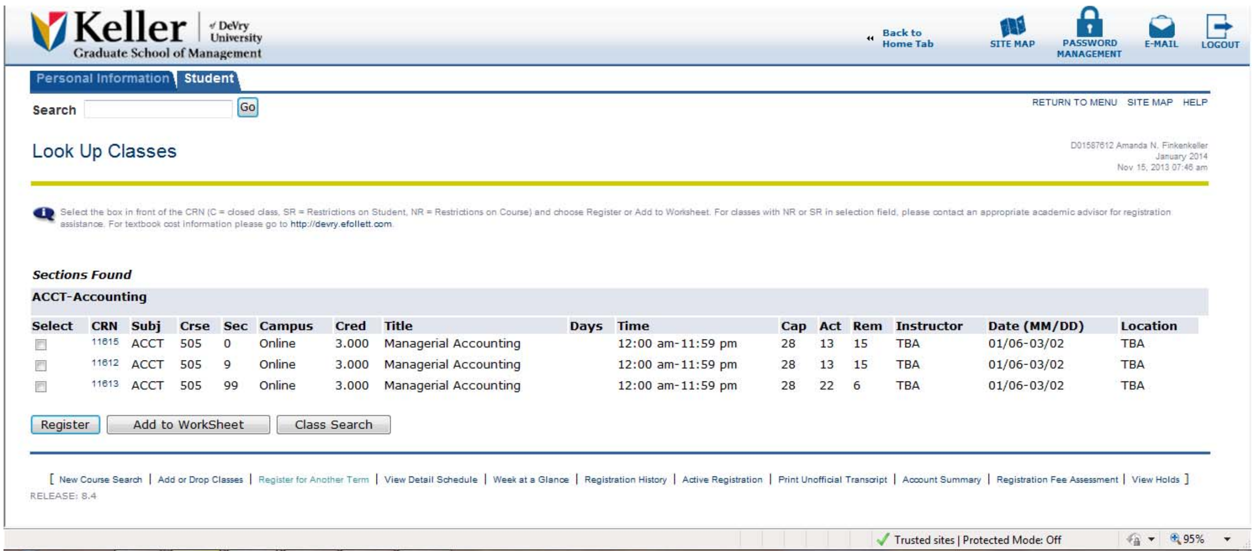
	Section mean course grade (0-4)			Proportion of students in the section receiving a course grade of								
				A or higher			B or higher			C or higher		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Section mean prior GPA	0.410*	0.324*	0.399	0.118+	0.100+	0.111	0.123*	0.094+	0.120	0.099**	0.077*	0.100
	(0.163)	(0.162)	(0.330)	(0.066)	(0.053)	(0.098)	(0.049)	(0.050)	(0.119)	(0.030)	(0.034)	(0.066)
Professor fixed effects		√			√			√			√	
Professor-by-course-by-term fixed effects			√			√			√			√

Note: Each column reports estimates from a separate OLS regression using 3,994 course-by-section observations. Dependent variables are described in the column headers. All independent variables are listed above: the mean of prior cumulative GPA for the section, and fixed effects. Standard errors in parentheses.
+ indicates p < 0.10, * < 0.05, ** < 0.01



Appendix Figure 1—Variation in magnitude of class size increase

Note: Density plot of course-by-term observations. The x-axis measure the percentage increase in class size for the "large" class relative to the "small" class for a given course in a given term. The vertical line marks the average increase in class size.



Appendix Figure 2—Image of information shown to students at the time of course registration

Appendix Table A1--Effect of class size on student achievement and persistence (intent to treat estimates), class size interacted with professor's total assigned student load across sections

	Course grade (0-4, F=0 and A=4)				Enrolled next term (binary)		Units enrolled next term (0-20)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A) Intended class size interacted with total professor student load								
ln(intended class size)	-0.046	-0.109	-0.001	-0.073	0.015	0.000	0.098	0.076
	(0.089)	(0.089)	(0.116)	(0.117)	(0.029)	(0.029)	(0.224)	(0.216)
Professor's total students in all sections	0.002	0.002	-0.001	0.004	-0.002	-0.001	-0.009	0.009
	(0.011)	(0.011)	(0.012)	(0.012)	(0.003)	(0.003)	(0.025)	(0.024)
ln(intended class size) * professor's total students	0.000	0.000	0.001	0.000	0.000	0.000	0.002	-0.003
	(0.003)	(0.003)	(0.003)	(0.003)	(0.001)	(0.001)	(0.007)	(0.007)
(B) Intended class size interacted with total professor online and in-person student load								
ln(intended class size)	-0.047	-0.110	-0.010	-0.079	0.013	-0.001	0.086	0.067
	(0.089)	(0.089)	(0.116)	(0.117)	(0.029)	(0.029)	(0.224)	(0.216)
Professor's total students in online sections	0.011	0.007	0.020	0.023+	0.003	0.003	0.013	0.035
	(0.012)	(0.013)	(0.013)	(0.013)	(0.003)	(0.003)	(0.027)	(0.026)
ln(intended class size) * professor's total students online	-0.002	-0.001	-0.005	-0.006	-0.001	-0.001	-0.004	-0.010
	(0.004)	(0.004)	(0.004)	(0.004)	(0.001)	(0.001)	(0.008)	(0.008)
Professor's total students in in-person sections	-0.013	-0.007	-0.046*	-0.034	-0.011+	-0.009	-0.055	-0.043
	(0.019)	(0.017)	(0.022)	(0.021)	(0.006)	(0.006)	(0.047)	(0.044)
ln(intended class size) * professor's total students in-person	0.005	0.003	0.014*	0.010+	0.003+	0.002	0.015	0.012
	(0.005)	(0.005)	(0.006)	(0.006)	(0.002)	(0.002)	(0.013)	(0.013)
Course-by-term-by-registration-group fixed effects, registration-group defined by…								
15 students grouped by reg. date-time	√		√		√		√	
vigintiles of reg. date-time		√		√		√		√
July, Sept. and Nov. terms only			√	√	√	√	√	√
Observations	102,209	102,209	67,858	67,858	67,858	67,858	67,858	67,858

Note: Each column within each panel reports estimates from a separate OLS or LPM regression using student-by-course observations. Estimation identical to Table 3 Panel C, but with the added covariates shown above. "Total students" variables are sums across all courses and sections taught by the professor in the current term, regardless of whether the course was part of the class-size experiment. Outcome for "next term" are not observed for the January term, and Jan. observations are only included in Columns 1 and 2. Standard errors allow for clustering in sections.

+ indicates p < 0.10, * < 0.05, ** < 0.01

Appendix Table A2--Effect of class size on student achievement and persistence, by number of sections professor taught (intent to treat estimates)

	Course grade (0-4, F=0 and A=4)				Enrolled next term (binary)		Units enrolled next term (0-20)	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(A) Taught by a professor teaching only one section in the current term								
ln(intended class size)	-0.083	-0.313+	-0.019	-0.292	0.082	0.077	0.501	0.336
	(0.188)	(0.177)	(0.249)	(0.236)	(0.067)	(0.063)	(0.541)	(0.500)
ln(intended class size)	0.056	0.226	0.051	0.202	(0.019)	(0.079)	0.090	0.096
* registered after size assignment	(0.247)	(0.225)	(0.316)	(0.284)	(0.090)	(0.081)	(0.649)	(0.618)
Student-by-course observations	48,429	48,429	32,024	32,024	32,024	32,024	32,024	32,024
(B) Taught by a professor teaching two or more sections in the current term								
ln(intended class size)	-0.330*	-0.222	-0.308+	-0.259	-0.064	-0.037	-0.544	-0.435
	(0.142)	(0.145)	(0.182)	(0.190)	(0.055)	(0.051)	(0.421)	(0.391)
ln(intended class size)	0.383*	0.189	0.622*	0.391+	-0.054	-0.077	0.022	-0.380
* registered after size assignment	(0.193)	(0.188)	(0.253)	(0.231)	(0.083)	(0.072)	(0.587)	(0.512)
Student-by-course observations	53,780	53,780	35,834	35,834	35,834	35,834	35,834	35,834
Course-by-term-by-registration-group fixed effects, registration-group defined by…								
15 students grouped by reg. date-time	√		√		√		√	
vigintiles of reg. date-time		√		√		√		√
July, Sept. and Nov. terms only			√	√	√	√	√	√

Note: Each column within each panel reports estimates from a separate OLS or LPM regression using student-by-course observations. Estimation identical to Table 3, but carried out separately for two sub-samples. The sample in Panel A includes all sections where the professor taught only one section during the term. The sample in Panel B includes all sections where the professor taught two or more sections during the term. Outcome for "next term" are not observed for the January term, and Jan. observations are only included in Columns 1 and 2. Standard errors allow for clustering in sections.

+ indicates p < 0.10, * < 0.05, ** < 0.01

Appendix Table A3--Professor student load, 2SLS estimates

	Course grade (0-4, F=0 and A=4)	
	(1)	(2)
(A)		
Average actual class size across all sections taught	0.002	-0.004
by the student's professor in the current term	(0.006)	(0.006)
F-statistic for excluded instrument in first stage	115.38	108.67
(B)		
Average actual class size across all sections taught	-0.001	-0.006
by the student's professor in the current term	(0.007)	(0.007)
Average actual class size across all sections	0.007	0.005
* student registered after class size assignment	(0.011)	(0.011)
F-statistic for excluded instruments in first stage	58.01	54.45
Course-by-term-by-registration-group fixed effects, registration-group defined by…		
15 students grouped by reg. date-time	√	
vigintiles of reg. date-time		√
Student-by-course observations	102,209	102,209

Note: Each column within each panel reports estimates from a separate 2SLS regression using student-by-course observations. The dependent variables are described in the column headers. The covariate of interest is the average actual class size across all sections (including different courses, and both online and in-person sections) taught by student i's professor during session s. In Panel A "average class size" is instrumented with an indicator = 1 if the section was a "small class" as assigned by the experiment. In Panel B both "average class size" and the interaction term "average class size" * "student registered after class size assignment" are instrumented with the "small class" indicator and the interaction term "small class" * "student registered after class size assignment." The regressions in Panel B also include a main effect for the indicator "student registered after class size assignment. "All regressions include prior cumulative GPA, in indicator = 1 if the student has failed a course previously, an indicator = 1 if the student is new to DeVry, and an indicator = 1 if prior GPA is missing for a continuing student. Prior GPA is set to zero for new students and students missing prior GPA. All regressions include fixed effects defined by the intersection of course, term, and registration date-time group. Registration date-time group is defined by, first, ordering students by the date-time they registered for the course; and, second, dividing students into either (i) groups of 15 students with sequential registration date-time, or (ii) vigintiles of registration date-time. Standard errors allow for clustering in sections.
+ indicates $p < 0.10$, * < 0.05, ** < 0.01