VIRTUAL CLASSROOMS: HOW ONLINE COLLEGE COURSES AFFECT STUDENT SUCCESS*

Eric Bettinger
Lindsay Fox
Susanna Loeb
Eric S. Taylor

Online college courses are a rapidly expanding feature of higher education, yet little research identifies their effects relative to traditional in-person classes. Using an instrumental variables approach and data from DeVry University, this study finds that taking a course online, instead of in-person, reduces student success and progress in college. Grades are lower both for the course taken online and in future courses. Students are less likely to remain enrolled at the university after an online class. These are local average treatment effects for students with access to both online and in-person options; for other students, online classes may be the only option for accessing college-level courses.

JEL No. I2, I23

* Bettinger, Fox, and Loeb: Stanford University, Center for Education Policy Analysis, 520 Galvez Mall, Stanford, CA 94305 (emails: ebettinger@stanford.edu, sloeb@stanford.edu, fox4@stanford.edu); Taylor: Harvard Graduate School of Education, Gutman Library 469, 6 Appian Way, Cambridge, MA 02138 (email: eric_taylor@harvard.edu). We greatly appreciate the support of DeVry University, especially Aaron Rogers, Ryan Green, and Earl Frischkorn. We also thank Tom Dee and seminar participants at UC Berkeley, Brigham Young University, CESifo, IZA, Mathematica Policy Research, Stanford University, University of Michigan, University of Stavanger, University of Texas Austin, Teachers College Columbia University, Texas A&M University, University of Uppsala, and University of Virginia for helpful discussions and comments. Financial support was provided by the Institute of Education Sciences, U.S. Department of Education, through Grant R305B090016 to Stanford University. The views expressed and any mistakes are those of the authors. A previous version of this paper was circulated with the title "Changing distributions: How online college courses alter student and professor performance."

Online college courses are a rapidly growing feature of higher education. One out of three students now takes at least one course online during their college career, and that share has increased threefold over the past decade (Allen and Seaman 2013). The promise of cost savings, partly through economies of scale, fuels ongoing investments in online education by both public and private institutions (Deming, Goldin, Katz, and Yuchtman 2015). Non-selective and for-profit institutions, in particular, have aggressively used online courses.

In this paper we estimate the effects of taking a given college course online, instead of in a traditional in-person classroom setting, on student achievement and progress in college. We examine both mean effects and how online courses change the variance of student outcomes. While online course-taking is both prevalent and growing, there remains relatively little evidence about how taking a course online, instead of in-person, affects student success in college. Evidence on this question from the for-profit sector is particularly scarce.

Our empirical setting has three advantageous features: the substantial scale of a large for-profit college, an intuitive counterfactual for each online course, and an instrument which combines two plausibly-exogenous sources of variation in whether students take a course online. The combination of these three features—and the resulting contributions to identification and generalizability—has not been possible in prior work.

We study students at DeVry University, a large for-profit college with an undergraduate enrollment of more than 100,000 students, 80 percent of whom are seeking a bachelor's degree. The average DeVry student takes two-thirds of her courses online. The remaining one-third of courses convene in conventional in-person classes held at one of DeVry's 102 physical campuses. The data for this paper cover more than four years of DeVry operations, including over 230,000 students enrolled in 168,000 sections of more than 750 different courses. [1]

DeVry University's approach to online education creates an intuitive, clear counterfactual. Each DeVry course is offered both online and in-person, and each student enrolls in either an online section or an in-person section. Online and in-person sections are identical in most ways: both follow the same syllabus and use the same textbook; class sizes are approximately the same; both use the same assignments, quizzes, tests, and grading rubrics. The contrast between online and in-person sections is primarily the mode of communication. In online sections, all interaction—lecturing, class discussion, group projects—occurs in online discussion boards, and much of the professor's "lecturing" role is replaced with standardized videos. In online sections, participation is often asynchronous while in-person sections meet on campus at scheduled times. In short, DeVry online classes attempt to replicate traditional in-person classes, except that student-student and student-professor interactions are virtual.

The contrast between online and in-person classes in the DeVry setting is, we think, consistent with intuitive definitions of "online" and "in-person" classes. We use these two labels throughout the paper as shorthand for this specific approach. Many other quite-different approaches to education are also commonly called "online education" or "online classes" (McPherson and Bacow 2015 provide a review), for example, massively open online courses (MOOCs). Our shorthand "online" should not read as broadly representative of all online education. However, the form of online education at DeVry is widely used in both the public and private sector.

To estimate the effects of taking a course online, instead of in-person, we use an instrumental variables approach. Our strategy makes use of two key

[1] This paragraph describes DeVry during the period we study—2009 to 2013. In recent years student enrollment has declined substantially, and many physical campuses have closed.

influences on students' course-taking behavior: (i) changes from term to term in which courses are offered in-person at each student's local campus, and (ii) the distance each student must travel to attend an in-person course at that local campus. Either of the two might be used as an instrument on its own. Distance has in fact often been used in studies of education, but with reservations (Card 2001, Xu and Jaggars 2013). Instead of using either alone, our instrument is the *interaction* of these two variables.[2] With the *interaction* serving as the excluded instrument, we control for the *main effects* of both variables in the first and second stages, following a strategy first proposed by Card (1995). The advantage of this interaction instrument is that it weakens the identifying assumptions. A causal interpretation of our estimates still requires the traditional assumptions—the instrument is conditionally independent of potential outcomes and can be excluded from the second stage—but the assumptions are more plausible because we can control for the main effects of variation in distance and course offerings.[3]

Our estimates provide evidence that online courses do less to promote student academic success and progression than do in-person courses. Taking a course online reduces student achievement, as measured by grades, in that course by about one-third of a standard deviation. Taking a course online also reduces student grades in future courses by one-eighth of a standard deviation, and reduces the probability of remaining enrolled a year later by over ten percentage points (over a base of 69 percent). Additionally, we find that student achievement outcomes are more variable in online classes, driven in part by a greater negative effect of online course-taking on students with lower prior GPA.

---

[2] The interaction of (i) an indicator $= 1$ if student i's home campus b offered course c on campus in a traditional in-person class setting during term t, and (ii) the distance between student i's residence and her home campus b. Results using either (i) or (ii) as the instrument are similar and available from the authors upon request.

[3] We further limit variation to within-course, with-home-campus, and within-major; control flexibly for secular trends; and control for prior achievement and other student observables.

Several plausible mechanisms could lead to poorer (improved) outcomes for students in online classes. Online courses substantially change the nature of interactions between students, their peers, and their professors. First, in online courses students can participate at any hour of the day from any place. That flexibility could allow students to better allocate time and effort, but could also be a challenge for students who have not learned to manage their own time. Chevalier, Dolton, and Luhrmann (2016), studying incentives and student effort, find evidence consistent with this hypothesis. Second, online courses change the constraints and expectations on academic interactions. Professors and students do not interact face-to-face; they interact only by asynchronous written communication. Thus students likely feel less oversight, generally, from their professors, and less pressure to respond to professors' questions. In the standard principal-agent problem, effort by the agent (student) falls as it becomes less observable to the principal (professor) (Jensen and Meckling 1976). Third, the role of the professor is quite different. Online classes standardize inputs, which traditionally vary between professors. For example, lectures are replaced with videos. Between-professor variation in student outcomes may shrink or may widen depending on how professors choose to use the time saved by no lecturing.

Our research contributes to two strands of literature. First, it provides substantial new evidence of the impact of online college classes— in particular, the impact for students in broad-access higher education institutions. Several prior studies randomly assign students to an online or in-person section of one course and find negative effects on student test scores (Figlio, Rush, and Yin 2013, Couch, Harmon, and Alpert 2014, Joyce et al. 2014) or, at best, null results (Bowen et al. 2014).[4] These studies are well-identified but each examines only a single course in economics or statistics, and college students at relatively-

[4] Using non-experimental methods, Brown and Liedholm (2002) and Coates et al. (2004) also find negative effects studying microeconomics principles courses.

selective public four-year colleges. We examine more than 700 courses, and students at a non-slective for-profit college, a population of particular interest for policy. At such colleges, online courses have grown most rapidly and are central to the institutions' teaching strategy. Several other quasi-experimental studies examine two-year community colleges and students taking a broad set of courses; the estimated effects of online course-taking are again negative.[5] Xu and Jaggers (2013, 2014) and Streich (2014b) use instrumental variables designs: distance from home to campus and availability of seats in in-person classes, respectively. A research design using either of these two instruments, on its own, requires relatively strong identifying assumptions for making causal claims. Our design substantially weakens identifying assumptions by combining two instruments.[6]

Second, our paper adds to the new and growing literature on private for-profit colleges and universities. Research on DeVry University and its peers is increasingly important to a complete understanding of American higher education. The for-profit share of college enrollment and degrees is large: nearly 2.4 million undergraduate students (full-time equivalent) enrolled at for-profit institutions during the 2011-12 academic year, and the sector granted approximately 18 percent of all associate degrees. For-profit colleges serve many non-traditional college students, who are often the focus of policy. Deming, Goldin, and Katz (2012) provide an overview of the for-profit sector.

---

[5] For comparison, one in three for-profit students takes all of her courses online, compared to one in ten community college students (McPherson and Bacow 2015).

[6] Additionally, Hart, Friedmann, and Hill (2014), makes a case for causal identification by focusing on within-course and within-student variation in whether a course is taken online or in-person. Streich (2014a) finds some evidence of positive effects on employment, though in years when the student is likely still enrolled in college. Finally, the widely-cited conclusion of a US Department of Education (2009) meta-analysis is that outcomes in online courses are better than traditional in-person courses; that meta-analysis, however, includes a wide variety of "online" courses. When the analysis is limited to studies of semester-length, fully-online courses there is no different in outcomes (Jaggers and Bailey 2010).

Our study is the first, of which we are aware, to estimate the effects of online courses among students at large for-profit colleges and universities. Our estimates complement a growing literature on labor market outcomes for for-profit college students. Deming et al. (forthcoming) report that graduates from mostly-online for-profit colleges are less likely to receive a callback for a job interview compared to observably similarly graduates from non-selective public colleges. For example, job applicants with a business degree were 22 percent less likely to be called back. Using a similar resume-audit design, Darolia et al. (2015) find for-profit graduates are no more likely to get a callback than are applicants without a college degree. These differences in hiring may or may not translate into differences in earnings (Turner 2012, Lang and Weinstein 2013, Cellini and Chaudhary 2014, Cellini and Turner 2016), but the differential for-profit earnings would need to be much larger to offset the higher costs (Cellini 2012). These poorer labor market outcomes may be in part because employers have learned what this study finds econometrically: students taking online courses have poorer academic achievement on average. However, a for-profit degree likely creates or signals other potential differences beyond a history of online course-taking.

I. Setting and Data

We study undergraduate, degree-seeking students taking courses at DeVry University. While DeVry began primarily as a technical school in the 1930s, today 80 percent of the University's undergraduate students are seeking a bachelor's degree; and most students major in business management, technology, health, or some combination. Each DeVry course is offered through both online classes and traditional in-person classes, though the availability of in-person classes varies over time and from campus to campus. For the average course, two-thirds of undergraduate classes occur online, and the other third occur at one of over 100 physical campuses throughout the United States. Students at DeVry are older and more often African-American or Latino(a) than students at public and non-profit colleges, though DeVry students are quite similar to students at other for-profit institutions (Appendix Table 1).

DeVry provided us with data linking students to their courses for all online and in-person sections of all undergraduate courses from Spring 2009 through Fall 2013. These data include information on over 230,000 students in more than 168,000 sections of 750 different courses. About one-third of the students in our data took courses both online and in-person. Table 1 describes the sample. Just under half of the students are female and average approximately 31 years of age, though there is substantial variability in age. Students in online courses are more likely to be female (55 percent vs. 35 percent) and older (33 years vs. 28 years).

The focus of this paper is on the level and variance of student outcomes. Ideally, we would like to know how much students learn in each course they take (whether online or in-person), but we have no direct measure of learning. Instead we examine the several different observed outcomes, which are imperfect correlates of learning: especially grades in the current course and future courses. There are reasons to be cautious about over-interpreting course grades. In many higher education institutions, course grades are subject to professor discretion, and professors may exercise that discretion differently in online and in-person classes. That discretion is a consideration in this paper, but DeVry's grading process permits less discretion than the typical selective college or university. For each course, DeVry's professors are asked to follow a common rubric for evaluating individual assignments and assigning grades. In many cases quizzes and tests are standardized across sections, whether online or in-person. Additionally, alongside course grades both for the target course and in future courses, we present results for persistence—a consequential outcome for students seeking a degree and one not influenced by professor subjectivity.

As shown in Table 1, the average grade was 2.8 (approximately B-), on the traditional zero (F) to four (A) scale. There is substantial variation in grades: the standard deviation is 1.3 (more than a full letter grade). Over 88 percent of students were still enrolled at DeVry University in the following semester or had completed their degree. The average student GPA in that following semester was 2.78. Over 69 percent were enrolled one year later or had completed their degree. These outcome means are consistently higher in the in-person classes than in the online setting. These differences could be the result of systematically different students enrolling in in-person classes, and in the next section we discuss our strategies for overcoming the selection bias.

II. Instrumental Variables Strategy

Our objective is to estimate the effect of taking a course online, instead of in a traditional in-person classroom, on student success in the course, success in future courses, and persistence in college. Calculating the difference in means or difference in variances is straightforward, but the decision to take a course in an online section or traditional section is likely endogenous—driven by unobservable information that could also influence each student's chances of success in the online versus traditional options. Our identification strategy is to use only variation in online versus in-person course-taking that arises because of two key influences on students' course-taking: (i) changes from term to term in which courses are offered in-person at each student's local campus, and (ii) the distance each student must travel to attend an in-person course at that local campus. Our instrument is the *interaction* of these two influences.

Our first student outcome is the grade, y_{ict}, received by student i in course c during term t. Each professor assigns traditional A-F letter grades, which we convert to the standard 0-4 point equivalents.[7] To estimate δ—the mean

[7] An A is 4 points, A- is 3.7, B+ is 3.3, B is 3, etc.

difference in course grades, y_{ict}, between students in online and traditional in-person classes—we first specify the following statistical model

$$(1) \qquad y_{ict} = \delta Online_{ict} + y_{i,\tau<t}\alpha + X_{it}\beta + \pi_c + \phi_t + \psi_{b(it)} + \rho_{p(it)} + \varepsilon_{ict},$$

where the indicator variable $Online_{ict} = 1$ if the course was taken online, and $= 0$ if it was taken in-person. Specification 1 includes several additional controls. First, we control for student i's prior grades, $y_{i,\tau<t}$, in all terms prior to term t. The vector $y_{i,\tau<t}$ includes two primary variables: (i) student i's prior grade point average (GPA) in all courses taken online, and separately (ii) her GPA in all courses taken in-person.[8] We also include observable student characteristics, X_{it} (gender and age); course fixed effects, π_c, for each of the 750 courses; major fixed effects, $\rho_{p(it)}$, for each of the 22 different degree programs; and a non-parametric time trend, ϕ_t, over the 27 terms in our data (spanning 4.5 years at 8 weeks per term). Finally, Specification 1 includes fixed effects for student i's "home campus" represented by $\psi_{b(it)}$. During the period of time we study, DeVry University operated 102 local campuses throughout the United States. We assign each student to one home campus, b, based on the physical distance between the student's home address and the local campus addresses, selecting as "home" the campus with the minimum distance.[9] Throughout the paper standard errors allow for clustering within campuses b.

[8] Each GPA is simply the weighted mean of prior course grades, where the weights are course credits. Because not all students have taken both online and in-person courses, $y_{i,\tau<t}$ also includes indicators for having previously taken any courses online, and any courses in-person. When online (in-person) GPA is missing we set it equal to the mean online (in-person) GPA.

[9] This distance is the straight-line distance. In addition to excluding students with missing address data, we exclude all students with international addresses and students whose addresses are not within the continental United States. The resulting sample contains 78 percent of the universe of undergraduate DeVry students over the time frame.

In response to selection bias concerns, we propose an instrumental variables strategy in which we instrument for $Online_{ict}$ in Specification 1 with the *interaction* of two variables: (i) $Offered_{b(i)ct}$, an indicator $= 1$ if student i's home campus b offered course c on campus in an in-person class setting during term t, and (ii) $Distance_{it}$, the distance in miles between student i's residence and her home campus b. With the *interaction* serving as the excluded instrument, we include *main effects* for $Offered_{b(i)ct}$ and $Distance_{it}$ in both the first and second stage. Card (1995) first proposed this interaction-instrument strategy.

To motivate the use of this interaction instrument, consider an alternative approach that uses the $Offered_{b(i)ct}$ indicator alone as the instrument. A strong first stage is quite plausible; the availability of an in-person section for course c at student i's local campus should reduce the probability that student i takes the course online. For identification, however, we would need to assume that local campus' course offering decisions are conditionally independent of potential outcomes. Our estimates would be biased if, for example, campuses are more likely to offer a course in-person during terms when the students who take the course are unobservably higher-achieving. This situation might occur because campus' decisions respond to explicit demand from such higher-achieving students. We discuss this and other potential threats in more detail as the paper progresses. By contrast, using the interaction instrument weakens the conditional independence assumption because we can control for the main effect of campus' in-person offering decisions, $Offered_{b(i)ct}$, in both stages. We can also control for the main effect of distance to that campus.

The interaction instrument still requires a strong first stage. A strong first stage is plausible. Assume the availability of an in-person class does reduce the probability of taking the course online; that effect should be heterogeneous with the reducing effect becoming weaker the further away from campus a student lives. We see this pattern empirically in the first-stage results presented in Table 2: availability of an in-person class reduces the probability of taking the course online by 22 percentage points for a student living next to the campus, but that reduction is smaller the further away from campus the student lives, about a 1.4 percentage point reduction for every 10 miles.

The interaction instrument has distinct advantages over using either $Offered_{b(i)ct}$ or $Distance_{it}$ alone as an instrument, but the interaction nevertheless uses the underlying variation in $Offered_{b(i)ct}$ and $Distance_{it}$. In the next two subsections we describe what gives rise to that variation.

II.a. Courses Offered In-person

Whether an in-person class option at the local campus is available to a student taking a particular course varies meaningfully. By contrast, an online class option is available every term for most all courses. DeVry divides its academic year into six eight-week terms. The typical course is, on average, offered in-person one out of every four terms at the typical campus (conditional on the campus ever offering the course in-person). The interquartile range is one out of thirteen terms to one out of three terms. About half of the total variation is within course and campus—differences from term to term in whether an in-person class is offered for a given course at a given campus.[10] In our analysis we limit identifying variation to this within-course and within-campus variation. The remainder of this section focuses on that residual variation and its causes.

Decisions about offering in-person classes are left largely to administrators at each local campus, especially decisions from term to term for a

[10] Consider observations on whether or not an in-person section is offered for course c at campus b in term t. Empirically, 25 percent of that variation is between courses, and an additional 21 percent is between campuses conditional on course. The remaining variation, 54 percent, is within course and campus over time.

given course.[11] The empirical data suggest that those decisions are not driven by current or past student demand. First, variation in in-person offerings is not explained, notably, by prior enrollment in the course at the campus. Enrollment in in-person sections during the prior year (six terms) explains just 0.1-0.6 percent of the variation.[12] Total enrollment—combining in-person and online enrollments for students assigned to the campus—explains similarly little of the variation. Moreover, the variation is not explained by the observable *characteristics* of students who enrolled in the course in prior terms, characteristics like GPA and whether they had taken an in-person course previously. Second, while more difficult to test, in-person offering decisions do not appear to be a function of demand in the current term t. Our partial test relies on a (supposed) University norm that in-person classes should be cancelled if fewer than five students enroll. If in-person offering decisions do respond to current demand we should see no classes with fewer than five students, or at least a discontinuous jump in the density of classes at five students. Empirically, there are many in-person classes with one to four students, and no evidence of a discontinuity at five students (or any other possible cutoff rule). A histogram is provided in Appendix Figure 1.

If not student demand, then what does explain the variation? First, the largest predictor we have identified is a form of seasonality. Nearly 25 percent of the variation in in-person offerings is explained by one pattern: often campuses offer in-person classes for a given course only every other term (e.g., in the 1st, 3rd, and 5th terms each year, or the 2nd, 4th, and 6th). This pattern suggests local campuses stick to established historical patterns. Second, anecdotally, university

[11] This has changed since our study period ended in 2013. Beginning in 2014 the University administration has gradually taken steps to centralize scheduling decisions. Personal correspondence in April 2016.

[12] The range covers a specification with a single regressor for the sum of enrollment in the prior year, and a specification with six regressors one for enrollment in each of six lagged terms.

administrators report that the availability of a professor interested in teaching the course also often determines in-person offerings.

To summarize, the data provide evidence that demand explanations are not first-order explanations for the variation in in-person offerings from term to term. The lack of empirical correlation with student demand and student characteristics is evidence consistent with conditional independence of in-person offering decisions and student potential outcomes. Variation is more readily predicted by seasonal, historical patterns, though there remains unexplained variation.

Two final remarks about the variation in in-person offerings: The first regards a potential exclusion restriction violation. If in-person class offerings are predictable, for example because of the seasonality described above, then campus b's offer of an in-person class for course c in term t may well influence student i's behavior in other courses, $d \neq c$, during terms, $\tau < t$, which precede term t. For example, a student might choose to take a pre-requisite course d online in term $t-1$ so that she can then take course c in-person in term t, even though she might have preferred to take course d in-person if time permitted. As we report in the results of this paper, taking course d online in term $t-1$ may well negatively affect her outcomes in course c during term t. The problem, in short, is that $Offered_{b(i)ct}$ is effectively assigned not in term t but in some prior term $\tau < t$. This threat to identification, like others, is partly addressed by the interaction instrument design. We also examine this threat empirically, as best we can, alongside other robustness checks after presenting the main estimates.

The second remark is that the variation in in-person offerings creates a missing data problem which raises a potential bias in our estimates. Consider the "never takers" in this setting: students who are unwilling take course c online in term t. These "never takers" would prefer to take course c during term t, but will only do so if there is an in-person class at their home campus b. Thus the "never takers" will be observed in our data only when $Offered_{b(i)ct} = 1$, and missing from the data when $Offered_{b(i)ct} = 0$. (By contrast, "always takers" and "compliers" are never missing from our data because courses are always offered online.) We describe the bias in detail in Appendix B. Two features of the bias are notable: First, the bias is, perhaps intuitively, proportional to the proportion of "never takers". This fact suggests an empirical test of the bias that we present in the section on robustness checks. Second, as a result of the bias, our estimates will likely understate the negative effects of taking a course online if the following assumption holds: if the missing "never takers" had been able to enroll in an in-person class as they desired, they would have higher grades in the course (persistence after the course) than the observed students, on average. While untestable, we think this assumption is plausible given the negative effects of online courses (reported here and in other papers), and the revealed preference for in-person courses among the missing "never takers".

II.b. Distance Between Student Residences and DeVry Campuses

The distance between student i's residence and the nearest DeVry campus b (which we call her "home campus") varies substantially. The median student lives 10 miles from the nearest DeVry campus, and the interquartile range is 5 to 28 miles. Ten percent of students live more than 100 miles from a campus making in-person course-taking quite unlikely. The complete CDF of $Distance_{it}$ is shown in Appendix Figure 2.

This distance is a function both of student i's choices about where to live and DeVry's choices about where to locate campuses. Both parties' choices may, relevant to identification, be partly influenced by potential outcomes. Indeed, there may be more scope for endogeneity in this setting than in other investigations using distance instruments: DeVry faces fewer constraints in changing campus locations compared to other colleges and universities, and college students generally may face fewer constraints to changing residences than other students. Our empirical strategy addresses, in large part, the concern by controlling explicitly for $Distance_{it}$. Nevertheless, our interaction instrument still makes use of the variation in $Distance_{it}$, and so we describe here what we know about the causes of that variation.

Students' movements to live closer to (further from) a campus *while they are enrolled* are not related to their prior academics or other observable characteristics. One in eight students changed residences during the period of our data. In half of moves (49 percent) students moved closer to a campus. There is no correlation between the distance moved and observables: GPA in online (in-person) classes, whether a student has taken online (in-person) classes, gender, and age all have correlations less than 0.03 in absolute value. Student moves may, nevertheless, be endogenous in unobservable ways; in the section on robustness tests we show our main results are unchanged if we fix student locations at their first residence in the data. However, we cannot say anything about students residence choices *before they enrolled*—choices which may have been based on the expectation of enrolling or influenced the probability of enrolling at DeVry.

DeVry's primary stated criteria when making campus decisions—opening, closing, and locations—are financial performance and local market competition.[13] These criteria suggest student demand may inform opening and closing decisions, Prior evidence from Cellini (2009) shows for-profit campus decisions do respond to demand. Empirically, however, this is not the case during the period of our data. Of the 102 campuses in our data, 10 opened during our study period. In Appendix Figure 3 we show average enrollment trends before and after a campus opening. We examine the enrollment of students near the new campus—students who will be assigned the new campus as their home campus because it is the

[13] Personal correspondence in April 2016.

closest campus—thus we observe enrollment before and after the actual opening. There is no trend up or down in enrollment of nearby students leading up to an opening, but enrollment does grow after an opening. The latter growth is true for enrollment in both in-person classes and online classes. DeVry may have opened these locations because they predicted potential enrollment, but the openings do not appear driven by prior demand trends. Additionally, during the period we study, DeVry closed just one campus despite widely declining enrollments.[14]

One final observation about the variation in distance before continuing the discussion of our instrumental variables strategy more generally: The relationship between distance from one's local campus and the probability of taking a class in-person may well be non-linear over the support in our data. Later in the robustness section we show that the pattern of results is generally robust to limiting estimation to students within 20 or 50 miles of a campus.

* * *

In addition to assessing effects on student grades, we use the same specification and interaction instrument strategy to estimate δ for other student outcomes. First, we examine the grades a student receives in subsequent classes, specifically student i's average grade in the next semester. We view future grades as an important complement to current grade. Future grades are drawn for both online and in-person settings, and are thus much less susceptible to concerns about differences is grading methods across mode of instruction or specific professor. Second, we turn to student persistence. We estimate the difference in the probability of enrolling at DeVry in the next semester and one year later.[15]

[14] This may surprise readers who have followed the for-profit-college market more recently. DeVry and its competitors have closed many more campuses in recent years (see, for example, Inside Higher Ed, May 7, 2015). As of this writing, DeVry University now lists just 55 campus locations. Falling enrollment, and its implications for profitability, may well have driven the closing of tens of campuses in 2014 and 2015.

[15] DeVry's academic calendar divides the year into six terms, with two consecutive terms equivalent to a semester in a more traditional calendar. We define "enrollment the next semester"

In addition to estimating the effect of taking an online course on *mean* outcomes, we also estimate the effect on the *variance* of outcomes. We estimate the variance effect, using an instrumental variables approach that builds on our approach for estimating the mean effect. First, we obtain the fitted residuals, $\hat{\varepsilon}_{ict}$, after two-stage least squares estimation of Specification 1. Second, we repeat the identical two-stage least squares regression used to estimate Specification 1, except that the dependent variable, y_{ict}, is replaced with the squared residuals $\hat{\varepsilon}_{ict}^2$. This is analogous to the familiar steps in FGLS or tests for heteroskdasticity.

Interpreting our estimates—for mean differences or variance differences—as causal effects of taking a course online, instead of in-person, requires the assumptions (i) that our instrument, $(Offered_{b(i)ct} * Distance_{it})$, is conditionally independent of potential grades and persistence outcomes; and (ii) that our instrument affects outcomes only by changing the probability that a course is taken online. The preceding discussion in this section provides some evidence on these assumptions. As one additional piece of evidence consistent with these assumptions, Appendix Table 2 shows a series of covariate tests where we estimate Specification 1 replacing the outcome variable with one student characteristic moved from the right hand side: gender, age, prior GPA in online and in-person courses. In all cases, the "effect" of taking a course online on these pre-treatment outcomes is not statistically different from zero.

III. Results

III.a. Effects on Student Grades and Persistence

Taking a course online, instead of in-person, reduces student success and progress in college. Table 3 reports local average treatment effect estimates (LATE) using the interaction instrument strategy described in the previous

---

as enrollment during either term $t + 1$ or $t + 2$ or both, and "enrollment one year later" as enrollment during either term $t + 5$ or $t + 6$ or both.

section. (Ordinary least squares estimates of Specification 1 are similar and are shown in Appendix Table 3.) The estimated effect of taking a course online is a 0.44 grade point drop in course grade, approximately a 0.33 standard deviation decline. Put differently, students taking the course in-person earned roughly a B- grade (2.8) on average while their peers in online classes earned a C (2.4). Results presented in Appendix Table 4 suggest this effect is true across the grade distribution, not simply at the top or bottom. Additionally, taking a course online reduces a student's GPA the following term by 0.15 points.

Grades are certainly an imperfect measure of what students actually learn in a course, and so we are cautious about over-interpreting the effects on grades. Nevertheless, these results likely do provide some signal about learning. Assume, reasonably we believe, that what students learn in one class affects their performance in future classes. If grades partly reflect actual learning, we should expect a poorer grade in one course to result in poorer grades in future courses, which is exactly what we see empirically. Moreover, the negative effect should be larger if the future course explicitly builds on knowledge gains in the first course. We test this prediction in two ways. First, we restrict the outcome measure to next term GPA for courses in the same subject area as the treated course. The point estimate is, as predicted, larger at -0.42 (standard error 0.068). Similarly, second, we restrict the outcome measure to GPA in future courses for which the treated course is a pre-requisite. Again the point estimate is larger -0.32 (standard error 0.148).[16] By contrast, if the negative estimate for course grade only reflected differences in grading standards between online and in-person, not differences in

[16] Both effects are estimated just as the effect in Table 3 Column 2, except with the respective restrictions of future courses that contribute to the outcome GPA measure. For the "same subject" estimate, subject is defined by course codes. For the "pre-requisites" estimate, course A is a pre-requisite to course B if 85 percent or more of students took course A before taking course B. The sample sizes are much smaller for these estimates—835,913 for "same subject" and 156,275 for "pre-requisites"—which partly reflects selection into future courses. As a result we use the results for GPA in all future courses as our main outcome.

learning, then we would not have expected this pattern of differences in future courses. Finally, however, even if grades do not reflect learning, grades do affect students' progress towards earning a degree.

Returning to Table 3, the estimates in Columns 3 and 4 suggest that taking a course online, instead of in-person, increases the probability that the student will drop out of school. In the semester after taking an online course, students are about 9 percentage points less likely to remain enrolled. On average 88 percent of DeVry students remain enrolled.[17] One year later, the reduction in enrollment has not grown or shrunk appreciably. This estimate suggests effects of online classes on persistence occur quickly, largely affecting decisions to reenroll the next semester. While our setting is quite different, it is useful to compare other estimates of effects on college persistence. For example, the literature on financial aid often finds that $1000 in financial aid increases persistence rates by about three percentage points (Bettinger 2004) and college mentorship increases persistence rates by five percentage points (Bettinger and Baker 2013). Additionally, as reported in Appendix Table 4, taking a course online also negatively affects the intensive margin of future enrollment; online students who do reenroll take fewer credits in future semesters.

III.b. Robustness Checks

In this section we present several robustness tests to address identification and bias questions raised in previous sections. First, the relationship between distance from one's local campus and the probability of taking a course in-person may well be non-linear over the support in our data. To test the sensitivity of our results to such non-linearities, we re-estimate our results limiting the sample to students who live within a plausible commuting distance: within 20 miles (the 70th percentile), and within 50 miles (the 83rd percentile). Table 4 Panels A and

[17] If a student graduates we set the outcome variable to 1 counting them as "enrolled" in this analysis.

B provide the results. The effects of taking an online course remain negative, though the losses are smaller and we lose precision in the restricted samples.

Next we address the possibility that our instrument could affect student choices and behavior in other courses, $d \neq c$, during terms, $\tau < t$, which precede term t. As described in greater detail earlier, this mechanism could lead to a potential exclusion restriction violation. We offer two related robustness checks. First, assume our instrument affects current outcomes partly by affecting academic experiences in prior terms. In that case we would, expect a correlation between our instrument and prior GPA, and thus omitting the prior GPA controls from our regression would add bias to our estimates. In Table 4 Panel C we show that omitting prior GPA controls does not change our results. Second, Panel D shows the estimates are also robust when we limit the sample to only students in their first term at DeVry. For new students there is much less scope for the instrument to affect prior academic choices and experiences.[18]

The third concern to address is the potential bias arising because of missing observations on "never takers", as we described in detail earlier. One important characteristic of this missing data bias, as shown in the Appendix B, is that the bias is proportional to the proportion of students who are "never takers." Thus the bias should be reduced if we exclude "never takers" from the estimation sample. We cannot directly observe students who would be "never takers" for course c at campus b in term t, but we have a plausible predictor for such students. In Table 4 Panel E, we report results where the estimation sample excludes students who have never taken an online course in a previous term. The results remain negative and statistically significant for all outcomes. The estimated negative effects on current and future grades are larger, which is

[18] There is some scope, of course. New students may have chosen when to start based on their distance from campus or the expected availability of specific in-person courses.

consistent with the prediction that the missing data result in bias that understates the negative impacts of taking a course online.

The remainder of Table 4 reports two additional results demonstrating the robustness of our main results. Panel F reports estimates where, for each student, $Distance_i$ is held constant over time at the distance measured using the student's first observed residence. The results suggest our main results are not driven by students' endogenous residence changes, at least endogenous changes after they matriculate at DeVry. Panel G reports estimates where we further restrict variation by replacing course, campus, and term fixed effects with fixed effects for course-by-campus, course-by-term, and campus-by-term. The results are quite similar to our main estimates.

Finally, we address the potential attrition bias in the estimated effects for future grades. The differential attrition is implicit in our results for "Enrolled next semester". We do not have a direct test for attrition bias in our main results. However, in Table 5 we examine future grades with an alternative outcome measure not subject to attrition. In Column 1, for example, the dependent variable is $= 1$ if the student achieved an A- or higher, on average, in the next semester. The outcome is $= 0$ for students who failed to achieve at least an A-, either because they enrolled in courses but their grades fell short or because they did not enroll at all. The other columns show results for outcomes defined by B- and C- grade thresholds. The estimated effect of taking a course is negative and significant at each of these points in the distribution of future grades.

III.c. Treatment Effect Heterogeneity

Next we examine whether the effects of taking a course online are heterogeneous by prior academic preparation, student major, and other characteristics. In the interest of space, estimates of heterogeneity are shown in Appendix Table 5. We find, first, that the negative effects of taking a course online are largest for students with relatively low prior GPAs, and the negative effects shrink as prior GPA rises. We estimate a specification where the treatment indicator $Online_{ict}$ is interacted with GPA in all prior courses.[19] For students with the average prior GPA (about 3.03 points), taking a course online reduces current grade by 0.36 points. The coefficient on the interaction between online and prior GPA is positive and significant (0.54).

Differences in effects by major are less strong than differences by prior GPA. The negative effects of online courses are somewhat larger in health-related majors than in business- or computer-related majors. Business- and technology-related majors comprise the majority of students, and thus the estimated effects for these students are quite similar to the overall average effects.

We further explore potential heterogeneity by comparing treatment effects in different types of courses. For students taking required courses (about half of the sample), the effects on student grades are somewhat larger and the effects on persistence are somewhat smaller, though the general pattern matches the overall results shown in Table 3. We also estimate the effects of taking a course online separately in introductory/intermediate courses and advanced courses. At both course levels, the negative effect on current course grade holds. Other effects may differ by level: effects on future course grades appear larger for students taking advanced courses, but effects on persistence appear larger for introductory or intermediate courses.

III.d. Effects on the Variance of Student Outcomes

In addition to reducing average achievement, online courses also increase the variance of student academic performance, as measured by course grades both in the target course and during the next semester. Table 6 reports our estimate of the difference between the standard deviation for online classes and the standard

[19] This specification has two endogenous treatment variables: $Online_{ict}$ and the interaction of $Online_{ict}$ and prior GPA. We instrument for these two variables with two instruments: our main instrument $\left(Offered_{b(i)ct} * Distance_{it}\right)$ and the interaction of our main instrument and prior GPA. Additionally, we center GPA at the sample mean to aid interpretation.

deviation for in-person classes, as estimated using the instrumental variables strategy described in Section II. When students take a course online, instead of in a conventional in-person class, the between-student standard deviation of course grades is about 40 percent larger. The between-student standard deviation of grades next semester also grows about 30 percent. This increase in variance is large but perhaps not surprising given our discussion in the introduction of differences between online and in-person practices and technologies, and the results on heterogeneity by prior GPA.[20]

IV. Conclusion

This study is the first of which we are aware to provide evidence on the mean effects of at-scale online courses and on the distributional consequences of online courses at non-selective 4-year colleges. In addition, this study uses an instrumental variables strategy for addressing selection that arguably relies on weaker identifying assumptions than do prior studies estimating the effects of online course-taking in broad-access settings. Finally, our setting provides a clean counterfactual in which the only difference between online and in-person courses is the medium of instructional delivery. All other aspects—professor assignment, class size, syllabus, textbooks—are identical across online and in-person courses.

Our analyses provide evidence that students in online courses perform substantially worse than students in traditional in-person courses, and these findings are robust to a number of potential threats. We also find that the variance of student outcomes increases, driven, at least in part, by differentially larger negative effects of online course-taking for students with lower prior GPA. The results are in line with prior studies of online education in showing that in-person

[20] As a robustness check, we also estimate the variance effects of online course-taking using a random-effects maximum-likelihood approach. Appendix Table 6 shows these results. The results are very similar to those using the instrumental variables approach.

courses yield better mean outcomes than online courses (Figlio, Rush, and Yin 2013, Couch, Harmon, and Alpert 2014, Joyce et al. 2014, Xu and Jaggers 2013, Hard, Friedman, and Hill 2014, and Streich 2014b). Our results also suggest one reason why, as other studies have found, for-profit students may have poorer labor market outcomes (Turner 2012, Lang and Weinstein 2013, Cellini and Chaudhary 2014, Darolia et al. 2015, Deming et al. forthcoming).

While we find that online courses lead to poorer student outcomes, we cannot provide a full welfare analysis. Most notably, the existence of an online course option might have enabled a large group of students to take college courses who otherwise would not have done so. Our estimates are local average treatment effects—the effect for students whose decision to take a course online, instead of in-person, is determined entirely by the (lack of) availability of an in-person class at their home campus, and the distance they live from that home campus. These effects are, for example, not necessarily applicable to students who live so far away from a campus that they would not plausibly ever take an in-person course. Thus, we cannot estimate the extent of this expansion in college course access in our setting. Our estimates are, nevertheless, a critical input to a more complete welfare analysis. Overall, the results—lower student performance and greater student variation—while not necessarily surprising, provide evidence that online courses are not yet as effective as in-person courses. This current state, however, is not a necessary end state.

References

Allen, I. Elaine, and Jeff Seaman. 2013. *Changing Course: Ten Years of Tracking Online Education in the United States.* Newburyport, MA: Sloan Consortium.

Bettinger, Eric P. 2004. “How Financial Aid Affects Persistence.” In Caroline Hoxby (Ed.), *College Choices: The Economics of Where to Go, When to Go, and How to Pay for It*. University of Chicago Press.

Bettinger, Eric P., and Rachel B. Baker. 2013. "The Effects of Student Coaching: An Evaluation of a Randomized Experiment in Student Advising." *Educational Evaluation and Policy Analysis, 36* (1):3-19.

Bettinger, Eric P., Susanna Loeb, and Eric S. Taylor. 2015. “Remote but Influential: Peer Effects and Reflection on Online Higher Education Classrooms.” Center for Education Policy Research, Stanford University.

Bowen, William G., Matthew M. Chingos, Kelly A. Lack, and Thomas I. Nygren. 2014. "Interactive Learning Online at Public Universities: Evidence from a Six-Campus Randomized Trial." *Journal of Policy Analysis and Management, 33* (1):94-111.

Brown, Byron W. and Carl E. Liedholm. 2002. "Can Web Courses Replace the Classroom in Principles of Microeconomics?" *American Economic Review, Papers & Proceedings, 92* (2):444-448.

Card, David. 1995. “Using Geographic Variation in College Proximity to Estimate the Returns to Schooling.” In Louis N. Christofides, Robert Swidinsky, E. Kenneth Grant (Eds.) *Aspects of Labour Market Behavior: Essays in Honor of John Vanderkamp.* Toronto: University of Toronto Press.

Card, David. 2001. "Estimating the return to schooling: Problems on some persistent econometric problems." *Econometrica, 69* (5):1127-1160.

Cellini, Stephanie Riegg. 2009. “Crowded Colleges and College Crowd-Out: The Impact of Public Subsidies on the Two-Year College Market.” *American Economic Journal: Economic Policy, 1* (2):1-30.

Cellini, Stephanie Riegg. 2012. “For-Profit Higher Education: An Assessment of Costs and Benefits.” *National Tax Journal, 65* (1):153-180.

Cellini, Stephanie Riegg and Latika Chaudhary. 2014. “The Labor Market Returns to a For-Profit College Education.” *Economics of Education Review, 43* :125-140.

Cellini, Stephanie Riegg and Nicholas Turner. 2016. “Gainfully Employed? Assessing the Employment and Earnings of For-Profit College Students Using Administrative Data.” National Bureau of Economic Research Working Paper 22287.

Chevalier, Arnaud, Peter Dolton, and Melanie Lührmann. 2016. "Making It Count: Evidence from a Field Study on Assessment Rules, Study Incentives and Student Performance.” Institute for the Study of Labor (IZA) Working Paper 8582.

Coates, Dennis, Brad R. Humphreys, John Kane, and Michelle A. Vachris. 2004. “’No significant distance’ between face-to-face and online instruction: Evidence from principles of economics.” *Economics of Education Review, 23* (5), 533–546.

Couch, Kenneth A., William T. Alpert, and Oskar R. Harmon. 2014. “Online, Blended and Classroom Teaching of Economics Principles: A Randomized Experiment.” University of Connecticut Working Paper.

Darolia, Rajeev, Cory Koedel, Paco Martorell, Katie Wilson and Francisco Perez-Arce. 2015. “Do Employers Prefer Workers Who Attend For-Profit Colleges? Evidence from a Field Experiment.” *Journal of Policy Analysis and Management, 34* (4):881-903.

Deming, David J., Claudia Goldin, and Lawrence F. Katz. 2012. "The For-Profit Postsecondary School Sector: Nimble Critters or Agile Predators*?" Journal of Economic Perspectives, 26* (1):139-164.

Deming, David J., Claudia Goldin, Lawrence F. Katz, and Noam Yuchtman. 2015. Can Online Learning Bend the Higher Education Cost Curve? *American Economic Review, Papers & Proceedings, 105* (5):496-501.

Deming, David J., Noam Yuchtman, Amira Abulafi, Claudia Goldin, and Lawrence F. Katz. Forthcoming. "The Value of Postsecondary Credentials in the Labor Market: An Experimental Study." *American Economic Review*.

Figlio, David, Mark Rush, and Lu Yin. 2013. "Is It Live or Is It Internet? Experimental Estimates of the Effects of Online Instruction on Student Learning." *Journal of Labor Economics, 31* (4):763-784.

Hart, Cassandra, Elizabeth Friedmann, and Michael Hill. 2014. "Online Course-Taking and Student Outcomes in California Community Colleges." Working Paper.

Inside Higher Ed. 2015, May 7. "Vanishing Profit, and Campuses."

Jaggars, Shanna Smith and Thomas Bailey. 2010. "Effectiveness of Fully Online Courses for College Students: Response to a Department of Education Meta-Analysis." Community College Research Center, Teachers College, Columbia University.

Jensen, Michael C., and William H. Meckling. 1976. "Theory of the Firm: Managerial Behavior, Agency Costs and Ownership Structure." *Journal of Financial Economics, 3* (4):305-360.

Joyce, Theodore J., Sean Crockett, David A. Jaeger, Onur Altindag, and Stephen D. O'Connell. 2014. "Does Classroom Time Matter? A Randomized Field Experiment of Hybrid and Traditional Lecture Formats in Economics." National Bureau of Economic Research Working Paper 20006.

Lang, Kevin and Russell Weinstein. 2013. “The Wage Effects of Not-for-Profit and For-Profit Certifications: Better Data, Somewhat Different Results.” *Labor Economics, 24* :230-243.

McPherson, Michael S. and Lawrence S. Bacow. 2015. "Online Higher Education: Beyond the Hype Cycle." *Journal of Economic Perspectives, 29* (4):135-54.

Streich, Francie E. 2014a. “Estimating the Impact of Online Education on Labor-Market Outcomes.” University of Michigan Working Paper.

Streich, Francie E. 2014b. “Online and Hybrid Instruction and Student Success in College: Evidence from Community Colleges in Two States.” University of Michigan Working Paper.

Turner, Nicholas. 2012. “Do Students Profit from For-Profit Education? Estimating the Returns to Postsecondary Education with Tax Data.” Manuscript, U.S. Treasury.

U.S. Department of Education, Office of Planning, Evaluation, and Policy Development. 2009. *Evaluation of evidence-based practices in online learning: A meta-analysis and review of online learning studies.* Washington, DC: U.S. Department of Education.

Xu, Di, and Shanna Smith Jaggars. 2013. "The Impact of Online Learning on Students’ Course Outcomes: Evidence from a Large Community and Technical College System." *Economics of Education Review, 37* :46-57.

Xu, Di, and Shanna Smith Jaggars. 2014. "Performance Gaps Between Online and Face-to-Face Courses: Differences Across Types of Students and Academic Subject Areas." *The Journal of Higher Education, 85* (5):633-659.

TABLE 1—STUDENT CHARACTERISTICS AND OUTCOMES

	All	Online	In-person
	(1)	(2)	(3)
Took course online	0.591	1	0
Observations			
Student-by-course-by-term	2,323,023	1,373,521	949,502
Students	230,484	184,799	118,041
Courses	750	559	653
Course sections	168,223	63,443	104,780
Student characteristics			
Female	0.467	0.545	0.354
Age	31.107	32.986	28.390
Prior GPA	3.027	3.057	2.983
	(0.866)	(0.873)	(0.853)
Student outcomes			
Grade in the course (0-4)	2.821	2.798	2.856
	(1.329)	(1.357)	(1.285)
GPA next semester	2.784	2.822	2.732
	(1.159)	(1.179)	(1.128)
Enrolled next semester	0.882	0.874	0.893
Credits attempted	9.764	9.126	10.652
next semester	(4.657)	(4.555)	(4.651)
Enrolled semester one year later	0.686	0.681	0.693
Credits attempted	7.737	6.899	8.906
semester one year later	(5.642)	(5.392)	(5.774)

Note: Authors' calculations. Means (standard deviations) for DeVry University undergraduate course enrollments from May 2009 to November 2013.

TABLE 2—FIRST STAGE

	Second-stage dependent variable		
	Course grade	GPA next semester	Enrolled next semester or one year later
	(1)	(2)	(3)
Excluded instrument			
Offered * distance	0.014***	0.014***	0.013***
	(0.001)	(0.001)	(0.001)
F-statistic	100.07	99.1	96.84
Additional controls			
Offered in-person at home campus	-0.220***	-0.223***	-0.210***
	(0.015)	(0.016)	(0.015)
Distance to home campus (10s of miles)	-0.001	-0.001	-0.001
	(0.001)	(0.001)	(0.001)
Female	0.040***	0.041***	0.040***
	(0.001)	(0.001)	(0.001)
Age	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)
Prior GPA, online courses	0.017***	0.018***	0.013***
	(0.001)	(0.001)	(0.001)
Prior GPA, in-person courses	0.006***	0.004*	0.003*
	(0.002)	(0.001)	(0.001)
Observations	2,323,023	2,106,090	2,360,645

Note. Each column reports the first stage coefficients from a separate two-stage least squares regression. The dependent variable in the first stage is an indicator = 1 if the student took the course online. Different second stage outcome variables have different samples, and thus the different first stage results. All specifications also include separate fixed effects for course, term, home campus, and major. When a prior GPA variable is missing we set the value to the sample mean and include an indicator = 1 in all cases where the variable is missing. The estimation sample is limited to students who have address information. Standard errors allow for clustering within campuses.
* indicates $p < 0.05$, ** 0.01, ***$p<.001$

TABLE 3—EFFECT OF TAKING A COURSE ONLINE, INSTEAD OF IN-PERSON, ON STUDENT ACHIEVEMENT AND PERSISTENCE (LOCAL AVERAGE TREATMENT EFFECT)

	Dependent variable			
	Course grade (A=4...F=0)	GPA next semester	Enrolled next semester	Enrolled one year later
	(1)	(2)	(3)	(4)
Took course online	-0.440***	-0.151*	-0.093***	-0.105***
	(0.049)	(0.058)	(0.014)	(0.017)
F-statistic for excluded instrument	100.07	99.10	96.84	96.84
Sample mean (st. dev.) for dep. var.	2.821	2.784	0.882	0.686
	(1.329)	(1.159)		
Observations	2,323,023	2,106,090	2,360,645	2,360,645

Note: The first row of each column reports the estimated local average treatment effect from a separate two-stage least squares regression. The second row reports the *F*-statistic for the excluded instruments from the first stage, and the third row shows the sample mean (standard deviation) of the dependent variable. Dependent variables are described in the column headers. The specification includes one endogenous treatment variable, an indicator = 1 if the student took the course online. The excluded instrument is the interaction between (a) an indicator variable = 1 if the course was offered in-person at the student's home campus (defined as the nearest campus) and (b) the distance in miles from the student's home address to her home campus. All specifications include the main effects of (a) and (b). All specifications also include controls for prior GPA, gender, age, and separate fixed effects for course, term, home campus, and major. Standard errors allow for clustering within campuses.
* indicates p < 0.05, ** 0.01, *** 0.001

TABLE 4—ROBUSTNESS CHECKS

	Dependent variable			
	Course grade (A=4...F=0)	GPA next semester	Enrolled next semester	Enrolled one year later
	(1)	(2)	(3)	(4)
A. Only students within 20 miles of campus				
Took course online	-0.264	0.055	-0.064*	-0.046
	(0.166)	(0.164)	(0.031)	(0.058)
Observations	1,614,646	1,481,373	1,641,395	1,641,395
B. Only students within 50 miles of campus				
Took course online	-0.293***	-0.074	-0.054**	-0.047
	(0.069)	(0.070)	(0.017)	(0.027)
Observations	1,918,781	1,754,160	1,946,970	1,946,970
C. Omit all prior GPA controls				
Took course online	-0.382***	-0.120	-0.089***	-0.097***
	(0.066)	(0.077)	(0.015)	(0.020)
Observations	2,323,023	2,106,090	2,360,645	2,360,645
D. Only new students				
Took course online	-0.372***	-0.234*	-0.144***	-0.207***
	(0.103)	(0.112)	(0.025)	(0.032)
Observations	270,917	239,181	289,151	289,151
E. Only students who have previously taken an online course				
Took course online	-0.590***	-0.276***	-0.087***	-0.123***
	(0.050)	(0.066)	(0.021)	(0.023)
Observations	1,446,162	1,299,656	1,421,103	1,421,103
F. Constant distance and home campus for a student				
Took course online	-0.434***	-0.139*	-0.091***	-0.100***
	(0.049)	(0.057)	(0.013)	(0.018)
Observations	2,323,023	2,106,090	2,360,645	2,360,645
G. Alternative set of fixed effects				
Took course online	-0.430***	-0.138*	-0.072***	-0.069***
	(0.046)	(0.061)	(0.012)	(0.018)
Observations	2,323,023	2,106,090	2,360,645	2,360,645

Note. Each column, within panels, reports estimates from a separate two-stage least squares regression. The estimation procedure is described in the note for Table 3, however, each panel makes one change to that procedure. In Panels A, B, D, and E the estimation sample is restricted as described in the panel labels. In Panel C the prior GPA controls are omitted. In Panel F each student's distance measure and home campus are held constant over time at the distance and campus first observed in our data. In Panel G the fixed effects for campus, course, and term are replaced with campus-by-course, course-by-term, and campus-by-term fixed effects. Standard errors allow for clustering within campuses.
* indicates $p < 0.05$, ** 0.01, *** 0.001

TABLE 5—ALTERNATIVE MEASURES OF FUTURE GRADES

	Dependent variable: GPA next semester is…		
	A- or higher	B- or higher	C- or higher
	(1)	(2)	(3)
Took course online	-0.045***	-0.062**	-0.077***
	(0.011)	(0.019)	(0.021)
Observations	2,323,023	2,323,023	2,323,023

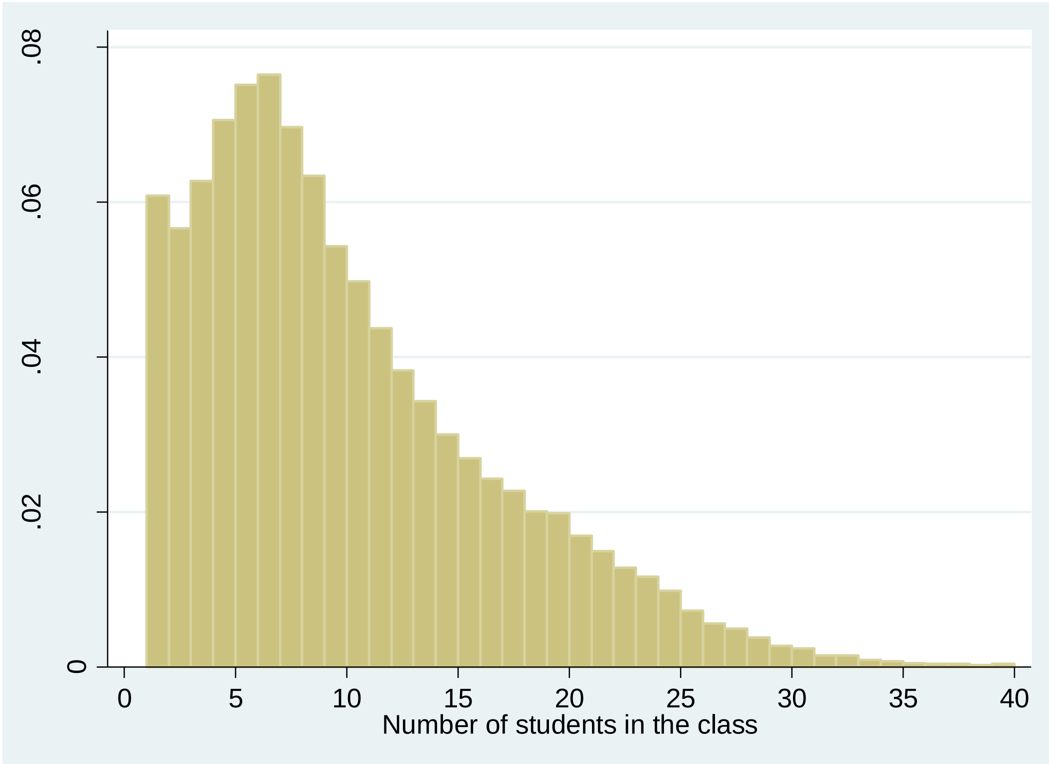
Note: Each column reports estimates from a separate two-stage least squares regression. . The estimation procedure is as described in the note for Table 3; only the dependent variables are different. In Column 1 the dependent variable is an indicator = 1 if the student achieved a grade of A or higher, on average, in the next semester, and = 0 in all other cases including students who did not enroll in courses. Dependent variables for Columns 2 and 3 are constructed similarly. Standard errors allow for clustering within campuses.
* indicates $p < 0.05$, ** 0.01, *** 0.001

TABLE 6—EFFECT OF TAKING A COURSE ONLINE, INSTEAD OF IN-PERSON, ON THE STANDARD DEVIATION OF STUDENT ACHIEVEMENT

	Sample st. dev.	Online − in-person diff. in st. dev.	P-value test diff. = 0	Observations
	(1)	(2)	(3)	(4)
Course grade	1.329	0.567	0.000	2,323,023
GPA next semester	1.163	0.342	0.007	2,106,090

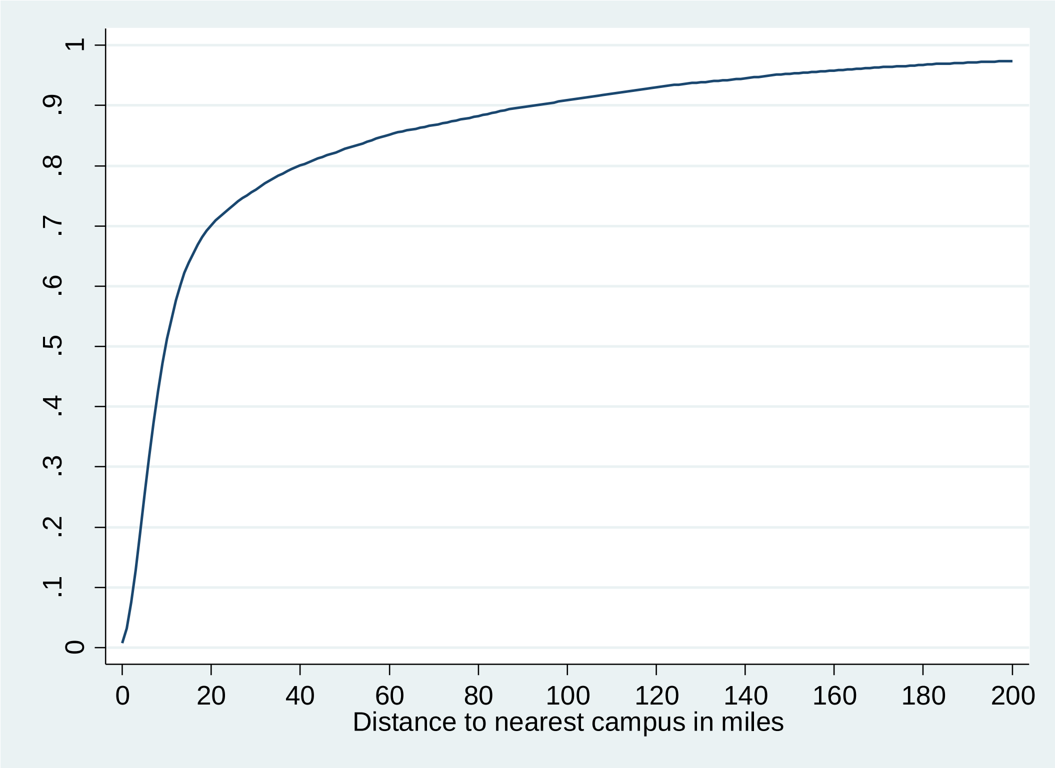
Note: Column 2 reports the estimated effect of taking a course online on the standard deviation of student course grade and GPA next semester. The dependent variable scale is 0-4, A = 4 ... F = 0. Column 1 reports the sample standard deviation of the dependent variable. The estimates in Column 2 are instrumental variables estimates obtained in two steps. Step 1 obtain the fitted residuals from the two-stage least squares regression used to estimate the mean effects in Table 3. Step 2 repeat the identical two-stage least squares regression except that the dependent variable is replaced with the squared residuals from Step 1. In both steps the specification includes one endogenous treatment variable, an indicator = 1 if the student took the course online. The excluded instrument is the interaction between (a) an indicator variable = 1 if the course was offered in-person at the student's home campus (defined as the nearest campus) and (b) the distance in miles from the student's home address to her home campus. All specifications include the main effects of (a) and (b). All specifications also include controls for prior GPA, gender, age, and separate fixed effects for course, term, home campus, and major. The test in column 3 allows for error clustering within campuses.

Appendix A: Additional figures and tables



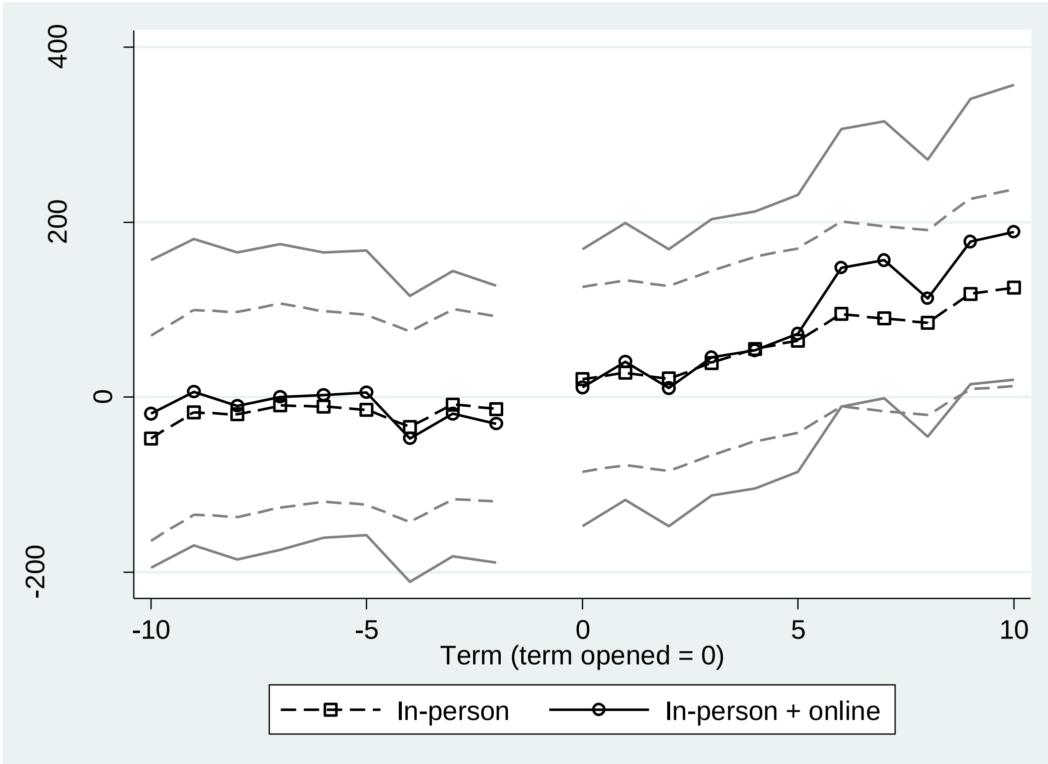
Appendix Figure 1—Histogram of class size for in-person classes

Note: Based on 112,628 in-person classes in the study data.



Appendix Figure 2—CDF of the distance in miles between student residence and nearest campus

Note: Based on 2,323,023 student-by-course observations. The x-axis is truncated at 200 miles to improve readability of the figure. 200 miles is the 97th percentile of the data.



Appendix Figure 3—Enrollment trends before and after a campus opening

Note: The square markers represent point estimates from a regression with campus-by-term observations. The dependent variable is the count of students enrolled in in-person courses. Students are assigned to the campus closest to the student's residence, even if that campus is not yet open. Thus the dependent variable counts enrollments for the assigned campus, not necessarily the campus where the student actually took the course. This enrollment count is regressed on a set of indicators for term relative to opening (leaving out the term just prior to opening, and existing campuses), campus FE, and calendar term FE (to control for university-wide trends). The circle markers repeat this process but for total enrollment: online plus in-person. They gray lines are 95% CIs.

APPENDIX TABLE 1—STUDENT CHARACTERISTICS, INSTITUTION TYPES, AND DEVRY UNIVERSITY (PERCENTAGES)

	Public 4-year	Private non-profit 4-year	Public 2-year	Private for-profit	DeVry University
	(1)	(2)	(3)	(4)	(5)
Age					
23 or younger	69.6	71.2	49.1	31.6	28.1
24-39	24.8	19.0	36.4	50.7	52.5
40+	5.6	9.9	14.4	17.7	19.4
Female	53.9	56.6	55.7	64.1	44.5
Race/Ethnicity					
African-American	12.8	13.4	16.4	25.6	25.5
Asian	6.9	6.9	5.0	2.9	4.9
Hispanic or Latino	13.8	10.1	18.6	18.5	18.7
Other, More than one	4.4	4.5	4.3	4.6	4.8
White	62.2	65.1	55.8	48.5	46.1

Note: Columns 1-4 come from the 2011-12 National Postsecondary Student Aid Study (NPSAS:12) as reported by NCES QuickStats. Column 5 is from DeVry administrative data. Race/ethnicity is imputed for 16.9 percent of students assuming missing at random.

APPENDIX TABLE 2—COVARIATE BALANCE

	Dependent variable			
	Female	Age	Prior GPA, online courses	Prior GPA, in-person courses
	(1)	(2)	(3)	(4)
Took course online	0.018	-0.941	0.004	0.170
	(0.019)	(0.519)	(0.082)	(0.194)
Observations	2,323,023	2,323,023	1,446,162	1,144,734

Note: Each column reports the coefficient from a separate two-stage least squares regression. Dependent variables are described in the column headers. The estimation procedure is identical to that described in the note for Table 3 except that when "Female" is the outcome variable it is removed from the right hand side controls. The same is true for "Age". When a prior GPA variable is the outcome, all prior GPA variables are removed from the right hand side. Standard errors allow for clustering within campuses.
* indicates $p < 0.05$, ** 0.01, *** 0.001

APPENDIX TABLE 3—OLS ESTIMATES OF THE 'EFFECT' OF TAKING A COURSE ONLINE, INSTEAD OF IN-PERSON, ON STUDENT ACHIEVEMENT AND PERSISTENCE

	Dependent variable			
	Course grade (A=4...F=0)	GPA next semester	Enrolled next semester	Enrolled one year later
	(1)	(2)	(3)	(4)
Took course online	-0.381***	-0.120***	-0.053***	-0.051***
	(0.012)	(0.007)	(0.002)	(0.003)
Observations	2,323,023	2,106,090	2,360,645	2,360,645

Note: Each column reports the OLS coefficient from a separate regression. Dependent variables are described in the column headers. The specification includes one 'treatment' variable, an indicator = 1 if the student took the course online. All specifications also include controls for distance to home campus, availability of an in-person section, prior GPA, gender, age, and separate fixed effects for course, term, home campus, and major. Standard errors allow for clustering within campuses.
* indicates $p < 0.05$, ** 0.01, *** 0.001

APPENDIX TABLE 4—EFFECT OF TAKING A COURSE ONLINE, INSTEAD OF IN-PERSON, ON STUDENT ACHIEVEMENT AND PERSISTENCE (LOCAL AVERAGE TREATMENT EFFECT) ADDITIONAL OUTCOMES

	Dependent variable		
	Course grade A- or higher	Course grade B- or higher	Course grade C- or higher
	(1)	(2)	(3)
Took course online	-0.121***	-0.133***	-0.101***
	(0.013)	(0.014)	(0.014)
Sample mean dep. var.	0.410	0.696	0.831
Observations	2,323,023	2,323,023	2,323,023

	Dependent variable			
	Passed course	Withdrew from course	Credits next semester	Credits one year later
	(4)	(5)	(6)	(7)
Took course online	-0.085***	0.066***	-0.623**	-1.254***
	(0.014)	(0.009)	(0.194)	(0.257)
Sample mean (st. dev.) for dep. var.	0.884	0.092	9.764	7.737
	(1.329)		(4.657)	(5.642)
Observations	2,323,023	2,601,742	1,980,377	1,520,954

Note. Each column, within panels, reports estimates from a separate two-stage least squares regression. The estimation procedure is described in the note for Table 3, only the dependent variables are different. Column 6 (7) is conditional on enrolling next semester (one year later). Standard errors allow for clustering within campuses.
* indicates p < 0.05, ** 0.01, *** 0.001

APPENDIX TABLE 5—TREATMENT EFFECT HETEROGENEITY

	Dependent variable			
	Course grade (A=4...F=0)	GPA next semester	Enrolled next semester	Enrolled one year later
	(1)	(2)	(3)	(4)
A. By prior achievement				
Took course online	-0.363***	-0.102	-0.061***	-0.065***
	(0.045)	(0.053)	(0.012)	(0.015)
Online * prior GPA	0.538***	0.495***	0.143***	0.167***
	(0.021)	(0.026)	(0.004)	(0.005)
Observations	2,323,023	2,106,090	2,360,645	2,360,645
B. Business-related majors				
Took course online	-0.446***	-0.145**	-0.068***	-0.092***
	(0.045)	(0.052)	(0.014)	(0.016)
Observations	1,007,534	900,335	1,006,664	1,006,664
C. Technology-related majors				
Took course online	-0.314***	-0.083	-0.094***	-0.094***
	(0.068)	(0.087)	(0.018)	(0.022)
Observations	880,811	797,795	877,484	877,484
D. Health-related majors				
Took course online	-0.875***	-0.267	-0.184***	-0.095
	(0.170)	(0.173)	(0.047)	(0.068)
Observations	371,605	348,615	407,444	407,444
E. Courses required for the student's major				
Took course online	-0.523***	-0.262***	-0.053***	-0.083***
	(0.072)	(0.070)	(0.014)	(0.021)
Observations	1,072,736	987,179	1,070,763	1,070,763
F. Introductory and intermediate courses (below 300 level)				
Took course online	-0.391***	-0.076	-0.094***	-0.077***
	(0.066)	(0.075)	(0.018)	(0.023)
Observations	1,460,463	1,325,177	1,511,154	1,511,154
G. Advanced courses (300 level or higher)				
Took course online	-0.470***	-0.190***	-0.029*	-0.041
	(0.053)	(0.056)	(0.013)	(0.021)
Observations	862,560	780,913	849,491	849,491

Note. Each column, within panels, reports estimates from a separate two-stage least squares regression. The note for Table 3 describes the estimation procedure; however, each panel makes one change to the procedure. Panel A adds an endogenous variable: the interaction between taking the course online and prior GPA (measured in all courses). The main effect of taking a course online and the new interaction are instrumented for with two instruments: the main offered*distance instrument, and the interaction between the main instrument and prior GPA. In Panels B-D the estimation sample is restricted to students in each category of majors. In Panels E-G the estimation sample is restricted by type of course as described in the panel labels. Standard errors allow for clustering within campuses.
* indicates $p < 0.05$, ** 0.01, *** 0.001

APPENDIX TABLE 6—EFFECT OF TAKING A COURSE ONLINE, INSTEAD OF IN-PERSON, ON THE STANDARD DEVIATION OF STUDENT ACHIEVEMENT RANDOM EFFECTS ESTIMATION APPROACH

	Sample st. dev.	Online − in-person diff. in st. dev.	P-value test diff. = 0	Observations
	(1)	(2)	(3)	(4)
Course grade	1.329	0.563	0.000	1,165,902
GPA next semester	1.163	0.267	0.000	1,116,350

Note: Column 2 reports the estimated effect of taking a course online on the standard deviation of student course grade and GPA next semester. The dependent variable scale is 0-4, A = 4 ... F = 0. Column 1 reports the sample standard deviation of the dependent variable. The estimates in Column 2 are random effects estimates from separate linear mixed models. The specification includes a random student effect which is allowed to be different for online and in-person classes. The fixed portion of the model includes controls for prior GPA, gender, age, and separate fixed effects for course, term, home campus, and major. The test in column 3 is a likelihood-ratio test where the constrained model requires the online and in-person variance parameters to be equal. The estimation sample is limited to students who have taken both online and in-person classes.

Appendix B: Characterizing the Bias from Missing Data on "Never Takers"

Observations for "never takers" are missing in our data. The "never takers" in this setting are students who are unwilling take course c online in term t. These "never takers" would prefer to take course c during term t, but will only do so if there is an in-person class at their home campus b. Thus the "never takers" will be observed in our data only when $Offered_{b(i)ct} = 1$, and missing from the data when $Offered_{b(i)ct} = 0$. By contrast, "always takers" and "compliers" are never missing from our data because courses are always offered online.

To simplify the notation, let (i) Y be the student outcome variable of interest, (ii) T be the treatment indicator called $Online_{ict}$ in Equation 1, (iii) Z be the instrument $\left(Offered_{b(i)ct} \times Distance_i\right)$, and (iv) W be a vector of all the remaining right hand side covariates included in our 2SLS first- and second-stages, including the main effects for $Offered_{b(i)ct}$ and $Distance_i$. Further let (v) m be an indicator $= 1$ if the student is a "never taker" as defined above, and (vi) $\rho = \mathrm{E}[m = 1]$, the probability of being a "never taker".

We can write the true effect of interest, δ in Equation 1, as a ratio of conditional covariances

$$\delta = \frac{cov(Y, Z|W)}{cov(T, Z|W)} = \frac{\mathrm{E}[YZ|W] - \mathrm{E}[Y|W]\mathrm{E}[Z|W]}{\mathrm{E}[TZ|W] - \mathrm{E}[T|W]\mathrm{E}[Z|W]} \tag{B1}$$

From here on we drop the $|W$ notation to simplify, but $|W$ should be thought of as implicit in all (cov)ariances and expectations below.

We can also write δ as a function of weighted sums of expectations. In particular, the numerator in B1 can be written

$$\begin{aligned} &\{\rho\mathrm{E}[YZ|m = 1] + (1 - \rho)\mathrm{E}[YZ|m = 0]\} - \\ &\quad \{\rho\mathrm{E}[Y|m = 1] + (1 - \rho)\mathrm{E}[Y|m = 0]\} \\ &\quad \times \{\rho\mathrm{E}[Z|m = 1] + (1 - \rho)\mathrm{E}[Z|m = 0]\}. \end{aligned} \tag{B2}$$

To simplify B2 first recall that the missing observations, $m = 1$, are missing because $Offered_{b(i)ct} = 0$. Thus when $m = 1$ it will always be the case that $Z = (0 \times Distance_i) = 0$. With this fact and a little algebra we can simplify B2 to

$$(1-\rho)(\mathrm{E}[YZ|m=0]-\{\rho\mathrm{E}[Y|m=1]+(1-\rho)\mathrm{E}[Y|m=0]\}\mathrm{E}[Z|m=0]). \tag{B3}$$

The denominator in B1 can be simplified the same way by replacing Y with T.

Thus we can write the true effect of interest, δ, as

$$\delta=\frac{\mathrm{E}[YZ|m=0]-\{\rho\mathrm{E}[Y|m=1]+(1-\rho)\mathrm{E}[Y|m=0]\}\mathrm{E}[Z|m=0]}{\mathrm{E}[TZ|m=0]-\{\rho\mathrm{E}[T|m=1]+(1-\rho)\mathrm{E}[T|m=0]\}\mathrm{E}[Z|m=0]} \tag{B4}$$

Contrast B4 with our empirical estimate $\hat{\delta}$ which is

$$\hat{\delta}=\frac{\mathrm{E}[YZ|m=0]-\mathrm{E}[Y|m=0]\mathrm{E}[Z|m=0]}{\mathrm{E}[TZ|m=0]-\mathrm{E}[T|m=0]\mathrm{E}[Z|m=0]} \tag{B5}$$

Subtracting the true numerator in B4 from our estimate of the numerator in B5 leaves

$$\rho\{\mathrm{E}[Y|m=1]-\mathrm{E}[Y|m=0]\}\mathrm{E}[Z|m=0]. \tag{B6}$$

Notice, first, that the missing data bias in the numerator will be proportional to ρ, the share of "never takers". Second, that the numerator's bias will be positive if $\mathrm{E}[Y|m=1]>\mathrm{E}[Y|m=0]$ and negative if the inequality is reversed. Assuming taking a class online, instead of in-person, has a negative effect on student outcomes, then positive bias would mean our estimates, $\hat{\delta}$, understate the true negative effects of online classes.

Similarly, subtracting the true denominator in B4 from our estimate of the denominator in B5 leaves

$$-\rho\mathrm{E}[T|m=0]\mathrm{E}[Z|m=0]. \tag{B7}$$

Expression B7 parallels B6, but further simplifies by noting that $\mathrm{E}[T|m=1]=0$. The missing observations, $m=1$, are all "never takers" where $Online_{ict}=0$ in all cases by definition. Again, first, notice that the denominator bias is proportional to ρ. Second, the denominator bias will always negative; that is, the estimated denominator is too small relative to the truth. This negative bias would mean our estimates, $\hat{\delta}$, are too large in absolute value. Put differently, the denominator bias makes the first-stage too small leading us to scale-up the reduced-form too much.

To summarize, first, the missing data bias is proportional to ρ. Second, our estimates will overstate the negative effects of taking a class online, instead of in-person, if $\mathrm{E}[Y|m = 1] < \mathrm{E}[Y|m = 0]$. If the inequality goes the other direction $\mathrm{E}[Y|m = 1] > \mathrm{E}[Y|m = 0]$ then the direction of the bias is ambiguous; bias in the numerator will understate the effects, but bias in the denominator will overstate the effects. The empirical tests presented in section III.b suggest that potential bias created by "never takers" likely leads to a small underestimation of the effects on online courses.