

The Impact of Distance Education on University Retention Rates

Emily Breit

Associate Professor of Finance, Fort Hays State University
USA

Samuel Schreyer

Associate Professor of Economics, Fort Hays State University
USA

Abstract

This study uses new measures of distance education to assess the impact on retention rates at 4-year public and private non-profit universities in the U.S. We present evidence that the percent of undergraduates enrolled exclusively in distance education courses reduces a university's freshmen retention rate, particularly for institutions with a relatively low median SAT score. We find no clear evidence of lower retention rates when undergraduates are enrolled in a combination of on-campus and distance education courses. These findings suggest increased enrollment through distance education can come at the expense of lower retention.

Keywords: distance education, online, retention, university, higher education

1. Introduction

This paper is motivated by a simple question: to what extent does a university's offering of distance education impact its ability to retain undergraduate students? The growth of online education in recent decades has resulted in considerable attention to this topic. According to data from the National Center for Education Statistics (NCES), about 8% of all undergraduates at 4-year institutions were enrolled in at least one distance education course in 1999; this figure increased to 30.3% by 2015.¹ Not all universities have incorporated distance education with equal vigor. Almost 71% of undergraduate enrollment at for-profit institutions in fall 2015 constituted at least one distance education course. In contrast, the figures are 23% and 27% for non-profit and public institutions, respectively.

Distance education is expected to continue to grow in the public and non-profit sectors, as it allows a university to reach a cohort of students that may otherwise have forgone a university education. Another reason distance education may grow is that it may help lower tuition costs. Indeed, there is empirical support that online education places downward pressure on tuition costs (Deming et al., 2015). Average state appropriations in the U.S. declined from \$8,616 to \$7,116 per full-time equivalent (FTE) student from 1991 to 2016 (State Higher Education Finance, 2017), while expenditures per student have been increasing

¹ Changes to federal legislation in recent decades may have facilitated the growth of distance education. The Higher Education Act of 1992 prevented students from receiving Title IV financial aid if their institution provided more than 50 percent of their courses online. In 1998 the Distance Education Demonstration Program was created which granted waivers to this rule for certain institutions. Finally, the Higher Education Reconciliation Act (HERA) went into effect in 2006 which eliminated the 50-percent rule entirely.

(Baum, Kurose, and McPherson 2013). Given these dynamics, it is not surprising that a survey of chief academic officers across the U.S. recently found that nearly 63% of respondents agreed with the statement that online education is critical to the long-term strategy of their institution (Allen and Seaman, 2015).

Greater enrollment in distance education may increase revenue for a university, but it may also impact a university's ability to retain students. Retention ratings are critical to college administrators as the majority of state legislatures have moved away from enrollment-based funding in favor of performance-based funding, with a university's retention rate often being a key metric (NCSL, 2015). Additionally, retention rates are used by ratings agencies, such as U.S. News & World Report, in determining a university's ranking.

Research on undergraduate retention dates back to the 1930s (Berger and Lyon, 2005). However, the studies which use the institution as the observational unit is considerably smaller, with virtually all of these studies utilizing data from the NCES' Integrated Postsecondary Education Data System (IPEDS). This is a comprehensive dataset for U.S. institutions of higher learning that contains measures of retention, university characteristics, and student body characteristics.

In this paper we study the impact of distance education, a newly collected variable, on a university's ability to retain its undergraduate students.² The data used in this study is obtained from IPEDS and covers all higher education institutions that participate in the federal student financial aid programs. We focus on 4-year public and non-profit institutions as it is these institutions that enroll a large majority of undergraduates in the U.S. (about 92% in fall 2015) and where online education has yet to make a large headway. The analysis spans three academic years, beginning in 2012-13 when IPEDS first began reporting distance education data.

1. Data and Literature Review

This study examines the relationship between a university's retention rate and its offering of distance education to undergraduates. IPEDS reports institutional-level data gathered through a variety of surveys given to all institutions of higher education that participate in federal financial aid.³

Recent studies that have utilized IPEDS data to study retention rates, or the related topic of graduation rates, include Marsh (2014) which examines retention rates at four-year public universities for the 2007-08 academic year. This study finds that entering student characteristics, such as SAT scores, had the greatest impact on retention rates. Webber and Ehrenberg (2010) estimate whether various non-instructional types of expenditures by universities directly influence graduation and retention rates of undergraduate students in public and non-profit 4-year universities. The authors find expenditures on student services are positively related to graduation and retention rates, and this matters most for schools with lower entrance test scores and a larger number of Pell Grant dollars per undergraduate student.⁴ Chen (2012) merges data from IPEDS

² Distance education, as defined in the IPEDS database, is education that uses one or more technologies to deliver instruction to students who are separated from the instructor and to support regular and substantive interaction between the students and the instructor synchronously or asynchronously.

³ IPEDS stems from the Higher Education Act of 1965, and subsequent amendments, which requires that institutions participating in federal student aid programs report a wide variety of institutional data, including retention and graduation rates.

⁴ There are several studies preceding Webber and Ehrenberg which also look at how various types of institutional expenditures influence retention/graduation rates using institutional-level data. See, for example, Gansemer-Topf and Schuh (2006), Gansemer-Topf and Schuh (2003), Dolan and Schmidt (1994), and De Groot (1991).

with the *Beginning Postsecondary Students* survey, which tracks a nationally representative cohort of students who began college in fall 1995. Using an event history model, the study finds student integration, financial aid, and expenditures on student services are important predictors of attrition. Our study contributes to this literature by analyzing the impact of distance education on retention rate.

Beginning with 2012-13 academic year, IPEDS began reporting institutional-level data for distance education. A distance education course is defined in IPEDS as a course in which the instructor and students are physically separated and instructional content is delivered entirely via technologies such as the internet, CD-ROM, and audio conferencing. The IPEDS database contains two measures of distance education: (i) the percent of undergraduates enrolled *exclusively* in distance education coursework; and (ii) the percent of undergraduates enrolled in *some*, but not all, distance education coursework. It is worth emphasizing that students enrolled in “some” distance education courses are also enrolled in traditional, on-campus coursework and therefore most likely reside in the vicinity of the university. In contrast, students enrolled “exclusively” in distance education courses are most likely to be fully online students without a tie to the physical location of the university.

There are a number of studies that have investigated distance education from the perspective of student attrition in specific courses or programs (e.g., Carr, 2000; Diaz, 2002; Levy, 2007; Tello, 2007; and Patterson and McFadden, 2013). In a review of this literature, Lee and Choi (2011) indicate the majority of these studies find higher attrition in distance education courses and programs, and that students’ previous academic performance is among the most important determinants of attrition, although course design and institutional supports influenced students’ dropout decision. More recently, Huntington-Klein et al. (2017) use an endogenous switching model with data from Washington State community colleges and finds that the average student who takes an online class is less likely to continue in his/her field or earn a degree than if this student took the course on campus.⁵ To our knowledge, however, no previous study has examined the impact of distance education on university retention rates using institutional-level data.

2. Descriptive Statistics

Table 2 presents descriptive statistics for the main variables used in this study.⁶ Column (1) shows descriptive statistics for the full sample, while the remaining columns present information by various subsamples; columns (2) and (3) separate the full sample based on whether or not an institution has any undergraduates enrolled in *all* distance education coursework; columns (4) and (5) distinguish the full sample based on whether or not an institution has any undergraduates enrolled in *some*, but not all, distance education coursework; columns (6) through (11) separate the full sample into “low” or “high” categories according to where an institution’s SAT score, proportion of Pell Grant recipients, and net price falls relative to the sample medians; and columns (12) and (13) distinguish public universities from private, non-profit universities.

The question that motivates our study is how university retention rates are impacted by the prevalence of undergraduate distance education. The descriptive statistics in column (1) show that the mean retention rate in the full sample is 77.9%. However, columns (2) and (3) reveal that institutions without exclusive distance

⁵ A related line of literature examines how student performance in the classroom is impacted by face-to-face delivery versus blended or online delivery. See Asarta and Schmidt (2017) for a review of this literature.

⁶ A complete list of variable definitions is provided in Table 1.

education have, on average, a 7.9 percentage point higher retention rate than institutions with exclusive distance education. Similarly, columns (4) and (5) indicate that institutions that do not have a single undergraduate enrolled in *some* distance education have a mean retention rate that is 8.5 percentage points higher than institutions that do offer this type of coursework. Looking at the mean retention rate across the various subsamples in columns (6) to (13), the rates are lower for institutions with relatively low SAT scores, high Pell Grant recipients, a low net price, and public institutions.

Table 2 also presents descriptive statistics on the extent of distance education coursework at our sample of universities. Column (1) indicates that a university has an average of 4.1% of and 13.1% of its undergraduates enrolled in *all* and *some* distance education coursework. In addition, columns (6) through (13) indicate that both forms of distance education are more prevalent at universities where retention rates tend to be lower—i.e., institutions with relatively low SAT scores, high Pell Grant recipients, low net price, and public institutions.⁷

The relationship between retention rates and distance education is examined in greater detail in Table 3. Institutions are grouped together based upon the percentage of their undergraduates enrolled in distance education using the ranges shown in the left-most column. About 22.6% and 40.2% of the institutions in our full sample do not have a single undergraduate enrolled in *some* and *all* distance education coursework: these institutions also have the highest mean retention rates at 84.5% and 82.6%, respectively. Importantly,

Table 3 also shows that mean retention rates decrease as the percentage of undergraduate enrollment in *all* and *some* distance education coursework increases.

It is also worth noting that the prevalence of undergraduate enrollment in distance education is highly skewed, especially for *all* distance education. The greatest number of institutions, or about 47.6 percent of our sample, have 1-10% of their undergraduate body enrolled in *all* distance education coursework.

Table 1 - Variable Definitions

Variable	Description	Detailed Definition (IPEDS variable name)
RR	Retention rate (%)	Percent of the fall full-time cohort from the prior year that re-enrolled at the institution as either full- or part-time in the current year (IPEDS: ret_pcf).
ALL	% enrolled exclusively in distance education courses	Percent of undergraduate students who are enrolled only in courses that are considered distance education courses. (IPEDS: pcudeexc)
SOME	% enrolled in some distance education courses	Percent of undergraduate students who are enrolled in at least one course that is considered a distance education course, but are not enrolled exclusively in distance education courses. (IPEDS: pcudesom)
MEDIAN_SAT	Median SAT	The median SAT represents the score for first-time, degree/certificate-seeking undergraduates; it is calculated as the average of the 25th and 75th percentiles of math and

⁷ These findings loosely correspond to McPherson and Bacow (2015, p. 138) who write “[i]n general, use of online learning appears to be inversely proportional to prestige and selectivity”.

		verbal SAT scores. For institutions reporting the ACT scores instead, the ACT Composite score was converted into its SAT analog using the College Board concordance table.
PELL	% awarded Pell grants	Percentage of full-time, first-time degree/certificate-seeking undergraduate students who were awarded Pell grants. (IPEDS: pgrnt_p)
NET_PRICE	Net price (2014 dollars)	Average net price for full-time, first-time degree/certificate-seeking undergraduates paying the in-state or in-district tuition who were awarded grant or scholarship aid from federal, state or local governments, or the institution. Other sources of grant aid are excluded. Aid awarded anytime during the full aid year is included. (IPEDS: npist2)
SF_RATIO	Student-to-faculty ratio	Total FTE students not in graduate or professional programs divided by total FTE instructional staff not teaching in graduate or professional programs. (IPEDS: stufacr)
Expenditure Controls	Institutional spending	Instructional expenditures per FTE (INSTRUCTION_EXP); Research expenditures per FTE (RESEARCH_EXP); Academic expenditures per FTE (ACADEMIC_EXP); Student services expenditures per FTE (STSERVICE_EXP). All expenditures are in 2014 dollars.
Demographic Controls	Dummies for race, gender, and historically black colleges or universities	Percent of undergraduates who are Asian; Percent of undergraduates who are Black/African American; Percent of undergraduates who are Hispanic/Latino; Percent of undergraduates who are women (IPEDS: pcuenras, pcuenrbk, pcuenrhs, and pcuenrw). A dummy variable equal to 1 if an institution is a historically black college or university, and 0 otherwise (IPEDS: HBCU).
Enrollment Controls	Undergraduate enrollment	Total undergraduates enrolled for credit in the fall of the academic year (IPEDS: efug).
Degree Controls	% of bachelor's degrees awarded in education, engineering, health professions, and business	Percent of bachelor degrees awarded in education (CIP13); Percent of bachelor degrees awarded in engineering (CIP 14); Percent of bachelor degrees awarded in health professions and related programs (51); and Percent of bachelor degrees awarded in business, management, marketing, and related support services (CIP 52) (IPEDS: derived from ctotalt13, ctotalt14, ctotalt51, ctotalt52, and basdeg)
Carnegie Controls	Institution type by highest degree offered	Universities and colleges are categorized into bachelor, master, or doctorate-granting institutions according to the framework in Carnegie Classification 2010 Update: Basic Classification. (IPEDS: ccbasic)
Locale Controls	Degree of urbanization	Degree of urbanization categorized as City, Suburb, Town, or Rural (IPEDS: locale)

Table 2 - Descriptive Statistics 2012-2014

	Full Sample	Are any undergrads enrolled in "all" DE?		Are any undergrads enrolled in "some" DE?		SAT Score		Pell Grant Recipients		Net Price		Institutional Control	
	(1) Total	(2) No	(3) Yes	(4) No	(5) Yes	(6) Low SAT	(7) High SAT	(8) Low Pell	(9) High Pell	(10) Low Price	(11) High Price	(12) Public	(13) Private
Retention Rate % of first-time, full-time undergrads that re-enrolled following year	77.9 (10.8)	82.6 (11.0)	74.7 (9.4)	84.5 (11.8)	76.0 (9.7)	71.0 (8.0)	84.8 (8.6)	84.2 (8.9)	71.7 (8.7)	74.3 (9.5)	81.5 (10.8)	76.2 (9.7)	81.6 (11.9)
All Distance Education % of undergrads enrolled exclusively in DE courses	4.1 (7.5)	0 (0)	6.9 (8.7)	0.8 (3.8)	5.1 (8.1)	5.9 (8.8)	2.3 (5.4)	2.5 (6.0)	5.7 (8.5)	5.7 (8.0)	2.5 (6.7)	4.8 (7.4)	2.6 (7.6)
Some Distance Education % of undergrads enrolled in some, but not all, DE courses	13.1 (13.2)	4.7 (8.8)	18.8 (12.6)	0 (0)	17 (12.6)	17.3 (13.6)	9.0 (11.4)	8.2 (10.5)	18.0 (13.8)	18.5 (13.2)	7.8 (10.8)	17.1 (12.9)	5.0 (9.5)
Median SAT Median SAT score of first-time degree-seeking UGs	1072.6 (143.9)	1137 (169.2)	1029.3 (103.6)	1190.4 (177.8)	1038.3 (111.1)	962.9 (56.9)	1182.5 (122.0)	1162.3 (141.4)	984.8 (81.5)	1016.0 (102.7)	1129.4 (159.0)	1036.8 (108.2)	1147.1 (176.8)
Pell Grants % of full-time freshmen receiving Pell Grants	37.8 (16.5)	31.8 (17.1)	41.8 (14.8)	27.9 (16.5)	40.6 (15.3)	48.1 (13.9)	27.4 (11.7)	24.6 (7.8)	50.7 (11.9)	44.6 (15.4)	31.0 (14.6)	41.0 (15.7)	31.0 (16.0)
Enrollment Undergraduate enrollment, 1,000s	8.7 (8.6)	6.9 (8.3)	9.9 (8.5)	4.1 (5.5)	10.0 (8.9)	6.5 (6.0)	10.8 (10.1)	9.7 (9.5)	7.7 (7.4)	10.3 (8.6)	7.1 (8.2)	11.2 (9.1)	3.4 (3.4)
Net Price Average for first-time, full-time undergrads; 1,000s of, 2014 dollars	16.8 (7.1)	20.4 (7.4)	14.4 (5.8)	22.9 (7.0)	15.1 (6.2)	14.3 (5.7)	19.4 (7.5)	20.1 (7.3)	13.7 (5.3)	11.4 (2.5)	22.2 (6.1)	13.0 (3.4)	24.9 (6.0)
Instruction Instructional expenditures per FTE; 1,000s of 2014 dollars	11.6 (10.4)	15.3 (13.9)	9.1 (6.0)	18.6 (16.3)	9.5 (6.6)	7.7 (2.6)	15.4 (13.4)	15.0 (13.6)	8.2 (3.2)	8.6 (3.9)	14.6 (13.5)	9.0 (3.7)	16.9 (16.2)
Research Research expenditures per FTE; 1,000s of 2014 dollars	3.4 (9.2)	4.9 (12.9)	2.4 (5.3)	6.2 (16.5)	2.6 (5.3)	0.8 (1.5)	6.0 (12.4)	5.6 (12.5)	1.3 (2.5)	2.3 (4.6)	4.6 (12.1)	2.8 (4.8)	4.8 (14.6)
Academic Academic expenditures per FTE; 1,000s of 2014 dollars	3.1 (3.4)	4.2 (4.6)	2.4 (2.1)	5.1 (5.6)	2.5 (2.2)	2.0 (1.0)	4.2 (4.5)	4.1 (4.5)	2.1 (1.3)	2.3 (1.5)	3.9 (4.5)	2.4 (1.5)	4.6 (5.4)
Student Services Student services expenditures per FTE; 1,000s of 2014 dollars	2.9 (2.4)	4.1 (3.0)	2.1 (1.2)	5.5 (3.3)	2.1 (1.2)	2.3 (1.4)	3.5 (2.9)	3.5 (2.9)	2.3 (1.4)	1.9 (1.2)	3.9 (2.8)	1.8 (0.9)	5.1 (2.9)
Observations	2173	874	1299	490	1683	1086	1087	1076	1097	1086	1087	1468	705

Notes: table values represent the sample means (standard deviations); the observational units are 4-year public and private non-profit institutions. For the full sample, there are 748, 744, and 681 institutions in the years 2012, 2013, and 2014 respectively. Columns (2) and (3) distinguish institutions that do not have any undergraduates enrolled exclusively in distance education coursework from those that do. Similarly, columns (4) and (5) distinguish institutions that do not have any undergraduates enrolled in some, but not all, distance education coursework from those that do. Columns (6) and (7) categorize institutions according to whether their median SAT score was above or below the sample median of 1045. In similar fashion, columns (8) and (9) categorize institutions based on whether their percentage of Pell Grant recipients was above or below the sample median of 39%. Columns (10) and (11) separate the full sample into those institutions with a net price above or below the sample median net price of \$15,170. Lastly, columns (12) and (13) distinguish public institutions from private institutions.

Table 3 - Distribution of Institutions by Undergraduate Enrollment in Distance Education Courses (2012-2014)

Percent of undergrads enrolled	Some distance education		Retention rate	All distance education		Retention rate
	Number of institutions	Percent of institutions	Mean (se)	Number of institutions	Percent of institutions	Mean (se)
0%	490	22.6	84.5 (11.8)	874	40.2	82.6 (11.0)
1-10%	611	28.1	79.3 (9.9)	1,035	47.6	76.0 (9.1)
11-20%	493	22.7	75.6 (8.7)	180	8.3	70.7 (8.7)
21-100%	579	26.7	72.8 (9.1)	84	3.9	67.7 (7.8)
Total	2173	100.0	77.9 (10.8)	2173	100.0	77.9 (10.8)

Notes: the observational units are 4-year public and private non-profit institutions. The total number of observations is 2,173 which consists of 749, 744, and 681 institutions in the years 2012, 2013, and 2014 respectively.

3. The Regression Model

We begin the analysis by estimating equation (1) using panel data that spans three academic years 2012-13, 2013-14, and 2014-15 for nearly 750 undergraduate institutions in the U.S.

$$\ln(RR_{it}/(1-RR_{it})) = a_0 + a_1 ALL_{it} + a_2 SOME_{it} + \mathbf{X}'_{it}\boldsymbol{\Gamma} + \mathbf{Y}'_i\boldsymbol{\Omega} + \mathbf{Z}'_t\boldsymbol{\Phi} + u_{it} \quad (1)$$

RR_{it} is the retention rate of institution i at year t , which refers to the percentage of full-time first-time undergraduates who started college in year t and re-enrolled at that institution in year $t+1$. The log odds ratio of RR_{it} is taken to ensure predictions lie within the range of 0% and 100%. ALL_{it} is the percentage of an institution's undergraduates that are enrolled in *all* distance education courses. Similarly, $SOME_{it}$ is the percentage of an institution's undergraduates that are enrolled in *some*, but not all, distance education courses. Vector \mathbf{X}'_{it} consists of institutional controls that vary over time; vector \mathbf{Y}'_i consists of controls that only vary by institution; vector \mathbf{Z}'_t represents year dummies; and u_{it} is the random error term.

The covariates of particular interest to this study are *ALL* and *SOME*. We expect *ALL* and *SOME* to be negatively related with an institution's retention rate which is consistent with the general findings from previous studies of higher student attrition in distance education courses and programs (e.g., Lee and Choi, 2011). In addition, institutions with a larger proportion of undergraduates enrolled in distance education may have fewer extracurricular activities which, in a traditional on-campus environment, would serve to strengthen the bond between a student and institution. For similar reasons, the marginal effect of *ALL* and *SOME* on retention rate may be dependent on the amount of distance education already incorporated into an institution's undergraduate curriculum. For example,

Table 3 shows that 22.6% and 40.2% of institutions do not have a single undergraduate enrolled in *some* and *all* distance education coursework, respectively. A university may become more efficient at offering distance education as it gains experience with this mode of education (e.g., learning by doing and resource specialization). For this reason, the main results presented in this study measure *ALL* and *SOME* as indicator variables where the percentage of undergraduate enrollment in the type of distance education coursework is categorized into the following four ranges: (i) 0%, (ii) 1-10%, (iii) 11-20%, and (iv) 21% and higher.⁸

Vector \mathbf{X}'_{it} contains institutional control variables that vary over time. Previous studies on institutional-level retention and graduation rates guided our selection of control variables (see, for example, Marsh, 2014; Webber and Ehrenberg, 2010; Berger and Lyon, 2005; Ishitani and DesJardins, 2002; and Astin, 1997). These controls include *MEDIAN_SAT* which is the average of the 25th and 75th percentile SAT scores of first-time degree-seeking undergraduate students; *PELL* which is the percentage of

⁸ We experimented with measuring *ALL* and *SOME* as continuous variables, as well as alternative ranges. These alternative measures did not substantively change the results and conclusions in this study.

undergraduates receiving a Pell Grant, a federal grant that is based on financial need; and *NET_PRICE*, which is the average net price for full-time, first-time degree-seeking undergraduates (measured in 2014 dollars).⁹

Distance education may enable institutions to take advantage of economies of scale by way of increased class sizes and/or by employing relatively fewer full-time faculty members (e.g., Bowen, 2012; McPherson and Bacow, 2015). To address these possibilities, vector \mathbf{X}'_{it} includes the student-to-faculty ratio (*STUFACR*), measured as the total number of full-time equivalent (FTE) undergraduate divided by the total number of FTE faculty. Similarly, previous studies, such as Webber and Ehrenberg (2010), have found the level and type of institutional expenditures influence retention rate, thus we control separately for an institution's expenditures on academics (*ACADEMIC_EXP*), student services (*STSERVICE_EXP*), instruction (*INSTRUCTION_EXP*), and research (*RESEARCH_EXP*). The expenditure variables are all expressed in logged 2014 dollars.

There are additional control variables in vector \mathbf{X}'_{it} , but we omit their estimates in the following regression tables for brevity. We control for demographic factors including a dichotomous variable for institutions classified as a Historically Black College or University, and the percentages of a university's undergraduate body that is female, Black, Hispanic, and Asian. Following Webber and Ehrenberg (2010) we control for enrollment (both undergraduate and graduate) by specifying enrollment in quadratic form. The types of undergraduate degrees awarded by an institution may also influence an institution's retention rate because of differences across collegiate units, as well as differences in labor market opportunities (e.g., Des Jardins et al., 2002; Des Jardins, 1999). Thus, we control for the percentages of all undergraduate degrees awarded in the following fields: business, education, engineering, and health professions.

Vector \mathbf{Y}'_i consists of controls that vary only by institution. These include dichotomous variables to control for institutional type according to an institution's Carnegie classification (i.e., bachelors, masters, doctorate); public institutions and private non-profit institutions; and the degree of urbanization for the institution's location (i.e., city, suburb, town, rural). Vector \mathbf{Z}'_t represents dummies for the academic years 2012-13, 2013-14, and 2014-15, which help control for macroeconomic influences on retention rate.

Lastly, a note regarding the method of estimation. The distance education variables in equation (1) display little variability within an institution across the three academic years for which we have data, therefore we do not employ fixed effects with our panel data. Instead, we use OLS and cluster standard errors by institution to allow for error terms may be correlated across time for a given institution.

⁹ For universities reporting the ACT score instead, we converted the composite ACT score into its SAT analog using the College Board's concordance tables. Separately, the dollar amount received by a Pell Grant recipient has a minimum and a maximum; in 2015-16 these were \$600 and \$5,775. Cost of attendance is a prominent factor in determining the dollar amount a Pell Grant recipient receives.

4. Results

Table 4 presents estimates of the coefficients in equation (1) for our full sample. Column (1) limits the covariates to our measures of distance education and a limited set of controls. As discussed in the previous section, the proportion of undergraduates enrolled in *all* and *some* distance education at a given university are each categorized into the following four ranges: (i) 0%, (ii) 1-10%, (iii) 11-20%, and (iv) 21% and higher. The baseline category for both *all* and *some* distance education is 0%, which refers to institutions that do not offer distance education at the undergraduate level. The negative signs for the estimated coefficients on the *all* and *some* categories indicate that, on average, retention rates are lower for universities with undergraduate enrolled in distance education.

Column (2) in Table 4 adds to the model controls for student academic preparation, financial need, cost of attendance, the student-faculty ratio, and four types of university expenditures. As discussed in the previous section, these variables are common determinants in studies modelling retention rates (c.f., McPherson and Bacow, 2015; Bowen, 2012; Webber and Ehrenberg, 2010). The coefficient signs on the additional control variables are as expected: the positive sign on *MEDIAN_SAT* indicates retention is higher for institutions that admit students with higher SAT scores; the negative sign on *PELL*, albeit statistically insignificant in column (2), indicates that retention is lower for institutions that have a larger share of undergraduates receiving the need-based Pell Grant; the negative sign for *NET_PRICE* indicates that retention is lower at higher-cost institutions; and the positive signs for *INSTRUCTION_EXP* and *ACADEMIC_EXP* indicate that higher retention rates are associated with universities that spend more on instruction and academics on a per FTE basis. Finally, the positive sign for *SF_RATIO* is not expected as it suggests that retention is higher when there are more students per faculty member. Note, though, that *SF_RATIO* is not statistically significant in many of the following regressions. For this study, what is perhaps most notable about the estimates in column (2) is the lack of statistical significance for the categories measuring *some* distance education. Indeed, the lack of a statistical relationship between a university's retention rate and the proportion of its undergraduates enrolled in *some* distance education is a finding that holds for most of the remaining analysis in this paper.

Table 4 – Regression Estimates

	(1)	(2)	(3)
2.ALL (1-10%)	-0.220*** (0.0518)	-0.101*** (0.0298)	-0.0629** (0.0275)
3.ALL (11-20%)	-0.438*** (0.0642)	-0.200*** (0.0435)	-0.118*** (0.0425)
4.ALL (>20%)	-0.487*** (0.0817)	-0.238*** (0.0495)	-0.185*** (0.0569)
2.SOME (1-10%)	-0.485*** (0.0770)	-0.0209 (0.0384)	-0.00554 (0.0348)
3.SOME (11-20%)	-0.625*** (0.0802)	-0.0449 (0.0442)	-0.0306 (0.0400)
4.SOME (>20%)	-0.708*** (0.0838)	-0.0647 (0.0484)	-0.0313 (0.0445)
MEDIAN_SAT		0.00355*** (0.000200)	0.00338*** (0.000214)
PELL		-0.00200 (0.00137)	-0.00960*** (0.00164)
NET_PRICE		-0.00695** (0.00319)	-0.00641** (0.00295)
SF_RATIO		0.0166*** (0.00484)	0.000816 (0.00468)
INSTRUCTION_EXP		0.408*** (0.0572)	0.286*** (0.0533)
RESEARCH_EXP		0.000984 (0.00398)	-0.00778** (0.00385)
ACADEMIC_EXP		0.0531*** (0.0199)	0.0477*** (0.0151)
STSERVICE_EXP		-0.0383 (0.0295)	-0.0401 (0.0267)
Constant	2.468*** (0.0847)	-6.271*** (0.499)	-4.796*** (0.520)
Demographic Controls	No	No	Yes
Enrollment Controls	No	No	Yes
Degrees Awarded Controls	No	No	Yes
Carnegie Controls	Yes	Yes	Yes
Locale Controls	Yes	Yes	Yes
Year Controls	Yes	Yes	Yes
Observations	2,173	2,173	2,173
R-squared	0.417	0.797	0.828

Notes: the dependent variable is the log-odds ratio of retention rate. All equations are based off of the full sample of institutions, and are estimated using pooled OLS with standard errors are clustered at the institutional level. The omitted group for ALL is 0% (i.e., no undergraduates are enrolled exclusively in distance education). Similarly, the omitted group for SOME is 0%. Variables are defined in Table 1; *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

Column (3) in Table 4 presents estimates using our complete set of control variables, although a portion of the estimation output is not shown for brevity. The additional control variables pertain to enrollment (i.e., the number of undergraduates enrolled at an institution); student demographics (i.e., the shares of the undergraduate population that are female, Asian, Black, and Hispanic, respectively) and whether or not the institution is classified as a Historically Black College or University (HBCU); and the degrees awarded (i.e., the share of undergraduate degrees awarded in business, education, engineering, and health professions). As with the estimates shown in the previous columns, the categories measuring *all* distance education are statistically significant, negative in sign, and increasing in magnitude. Other factors held constant, a university's retention rate is adversely impacted by the proportion of undergraduates enrolled exclusively in distance education coursework.

Interpreting the coefficients for the distance education categories is hindered by the fact that the dependent variable in equation (1) is the log odds ratio of the retention rate. To facilitate interpretation, we convert these coefficients into marginal effects using the following equation:

$$\Delta rr = \frac{\rho'}{\rho'+1} - rr, \text{ where } \rho' = \frac{rr}{(1-rr)} \times \exp(\alpha) \quad (2)$$

Note that Δrr is the predicted change in the retention rate percentage; rr is the initial retention rate which we specify as the full-sample mean; and α is the estimated coefficient on a distance education category.

Table 5 – Regression Estimates and Marginal Effects

	Full Sample (1)	Low SAT (2)	High SAT (3)	Low Pell (4)	High Pell (5)	Low Price (6)	High Price (7)	Public (8)	Private (9)
ALL (1-10%)	- 0.0629** (0.0275)	-0.0592* (0.0331)	-0.0438 (0.0375)	-0.0134 (0.0353)	- 0.0655* (0.0387)	-0.0616 (0.0453)	-0.0327 (0.0311)	- 0.0756** (0.0347)	-0.0323 (0.0425)
	[-1.102]	[-1.036]	[-0.763]	[-0.231]	[-1.148]	[-1.078]	[-0.568]	[-1.328]	[-0.561]
ALL (11-20%)	- 0.118*** (0.0425)	- 0.138*** (0.0486)	-0.0782 (0.0712)	-0.146** (0.0606)	-0.105* (0.0568)	-0.112* (0.0578)	-0.0854 (0.0650)	-0.123** (0.0490)	-0.0219 (0.0626)
	[-2.097]	[-2.466]	[-1.375]	[-2.615]	[-1.860]	[-1.988]	[-1.505]	[-2.189]	[-0.379]
ALL (>20%)	- 0.185*** (0.0569)	- 0.226*** (0.0717)	-0.00659 (0.0869)	- 0.224*** (0.0655)	-0.129* (0.0732)	-0.142* (0.0793)	-0.177** (0.0865)	- 0.185*** (0.0696)	-0.175* (0.104)
	[-3.347]	[-4.133]	[-0.114]	[-4.094]	[-2.300]	[-2.540]	[-3.196]	[-3.347]	[-3.158]
Observations	2,173	1,086	1,087	1,076	1,097	1,086	1,087	1,468	705
R-squared	0.828	0.427	0.841	0.851	0.553	0.712	0.867	0.760	0.890

Notes: table values are the estimated coefficients (standard errors) [marginal effects]. Marginal effects represent the change in the retention rate percentage, and are determined using the sample mean retention rate of 77.9%. Each column represents

an estimate of equation (1) with the full set of control variables for a particular sample of institutions. All estimates use pooled OLS, with standard errors are clustered at the institutional level. Estimate coefficients for categories of *some* distance education are not shown because of a lack of statistical significance. The baseline group for *ALL* is 0% (i.e., institutions with no undergraduates enrolled in *all* distance education). *, **, and *** denote significance at the 10%, 5%, and 1% levels respectively.

Table 5 presents regression estimates of equation (1) for a variety of subsamples, and reports the corresponding marginal effects in brackets. For brevity, the table lists only the covariates corresponding to the categories for *all* distance education. Estimates for the categories of *some* distance education are not shown because of a lack of statistical significance. Column (1) shows the marginal effects using the full sample of institutions. For instance, the marginal effect for *ALL* (1-10%) is about -1.1%, meaning that the predicted retention rate is nearly 1.1 percentage points lower for an institution with 1-10% of its undergraduates enrolled exclusively in distance education courses relative to a university that does not offer undergraduate distance education. Not surprisingly, the marginal effect is greater for institutions with a larger share of undergraduates enrolled in *all* distance education. Indeed, relative to the baseline group, universities with more than 20% undergraduates enrolled in *all* distance education courses are predicted to have a lower retention rate by about 3.3 percentage points.

Columns (2) and (3) in Table 5 present the estimates for the *all* distance education categories when equation (1) is estimated separately for the “low” and “high” SAT subsamples. As discussed in earlier sections, these subsamples are based on whether an institution’s SAT score falls above or below the sample median. Only the coefficient estimates in column (2), corresponding to the low SAT group, are statistically significant. This finding suggests that undergraduates at low SAT institutions are not only less academically prepared as measured by their SAT score, but they are also less likely to possess skills necessary to succeed in an exclusively online environment.

In similar fashion, columns (4) and (5) in Table 5 present the estimates for institutions based on whether the proportion of undergraduates receiving Pell Grant is “low” or “high” (i.e., below or above the sample median). With one exception for the low Pell group, the coefficient estimates are negative and statistically significant for both subsamples. There does not appear to be a marked difference in how distance education impacts retention rates between these two subsamples. Columns (6) and (7) present estimates for institutions based on whether the net price of attendance is “low” or “high”. Again, there does not appear to be any clear difference in how distance education impacts retention rates between these two subsamples. Lastly, columns (8) and (9) present estimates for public versus private institutions. The categories for *ALL* are negative and statistically significant for public institutions.

There is one clear pattern that emerges across the subsamples: the estimated coefficients for *ALL*(>20%) are negative and statistically significant for nearly every subsample. This indicates that a university’s retention rate is nontrivially impacted when a sizeable proportion of its undergraduates are enrolled exclusively in distance education coursework.

5. Conclusion

This study examines the relationship between a university’s retention rate and its offering of distance education to undergraduates. We use data from IPEDS and focus on 4-year public and non-profit

institutions because this is where distance education is likely to grow in the future. The time period analyzed in our study spans three years, beginning with the 2012-13 academic year when distance education data first became available.

We find very little evidence that *some* distance education—defined as the percentage of undergraduates enrolled in some, but not all, distance education courses—impacts retention rates. In contrast, we do find evidence that the proportion of a university's undergraduate body enrolled exclusively in distance education courses has a negative impact on the university's retention rate. These findings lead us to question what the fundamental differences are between the “some” and “all” distance education variables. One key difference is that students enrolled in *some* distance education courses have, by definition, a physical presence at that institution. A related possibility is that students enrolled in some distance education are receiving a different educational experience than students who are enrolled exclusively in distance education courses.

How much does exclusive distance education reduce an institution's retention rate? Our complete model estimated over the full sample suggests that all other factors held constant, an institution that does not offer *all* distance education courses to its undergraduates will have a retention rate that is 1.1 percentage points higher than an institution that has 1-10% of its undergraduates enrolled in *all* distance education; nearly 2.1 percentage points higher than an institution with 11-20% of its undergraduates enrolled in *all* distance education; and about 3.3 percentage points higher than an institution with more than 20% of its undergraduates enrolled in *all* distance education.

The descriptive statistics in this paper indicate that exclusive distance education is more prevalent at institutions with relatively low SAT scores, high Pell Grant recipients, low net price, and public institutions. These same types of institutions also have relatively lower retention rates. However, the regression estimates in this study find no clear evidence that these subsamples matter in terms of how retention is impacted by distance education, with one exception. The negative relationship between retention rates and distance education is particularly acute at universities that admit undergraduates with relatively low SAT scores. This finding may be because these students are likely to possess skills necessary to succeed in an exclusively online environment. Another important finding is that, regardless of the type of institution, a university's retention rate is adversely impacted when a sizeable proportion of its undergraduates are enrolled exclusively in distance education coursework.

Public and non-profit universities will almost certainly increase their reliance on distance education as a source of revenue in coming years. The results in this study indicate there may be an important tradeoff: increased enrollment through distance education yet lower retention.

6. References

- Allen, I. E., and J. Seaman (2015). Grade level: Tracking online education in the United States. Babson Park, MA: Babson Survey Research Group. Retrieved October 1, 2018 from Education Advisory Board website <http://www.onlinelearningsurvey.com/reports/gradelevel.pdf>
- Alpert, W. T., K. A. Couch, and O.R. Harmon (2016). A randomized assessment of online learning. *American Economic Review*, 106(5): 378-82.

- Arano, K., S. Schreyer, and D. Toulaboe (2018). A quantile regression analysis of student performance in online vs. traditional principles of economics. *Perspectives on Economic Education Research*.
- Asarta, C. J. and J.R. Schmidt (2017). Comparing student performance in blended and traditional courses: does prior academic achievement matter? *Internet and Higher Education*, 32: 29-38.
- Astin, A. W. (1997). How 'good' is your institution's retention rate? *Research in Higher Education*, 38, 647-658.
- Baum, S., C. Kurose, and M. McPherson (2013). An overview of American Higher Education. *The Future of Children*, 23 (1): 17-39.
- Berger, J. B. and S. C. Lyons (2005). Past to present: A historical look at retention. In A. Seidman (Ed.), *College student retention: Formula for student success* (pp. 1-27). Westport, CT: American Council on Education and Praeger Publishers.
- Bowen, W.G. (2012). The 'cost disease' in higher education: is technology the answer? Lectures presented at the Tanner Lectures on Human Values at Stanford University, Stanford, CA.
- Carr, S. (2000). As distance education comes of age, the challenge is keeping the students. *Chronicle of Higher Education*, 46, 39-42.
- Chen, R. (2012). Institutional characteristics and college student dropout risks: a multilevel event history analysis. *Research in Higher Education*, 53(5), 487-505.
- Deming, D. J., C. Goldin, L.F. Katz, and N. Yuchtman (2015). Can online learning bend the higher education cost curve? *American Economic Review*, May, 105:5, pp. 496-501.
- DesJardins, S.L., D.A. Ahlburg, and B.P. McCall (1999). An event history model of student departure. *Economics of Education Review*, 18(3), 375-390.
- DesJardins, S.L., D.A. Ahlburg, and B.P. McCall, (2002). Simulating the longitudinal effects of changes in financial aid on student departure from college. *Journal of Human Resources*, 37(3), 653-679.
- DesJardins, S.L., D.O. Kim, and C.S. Rzonca, (2003). A nested analysis of factors affection bachelor's degree completion. *Journal of College Student Retention*, 4 (4), 407-435.
- Diaz, D. P. (2002). Online drop rate revisited. *The Technology Source*, May/June. Retrieved from <http://technologysource.org/issue/2002-05/>.
- Huntington-Klein, N., J. Cowan, and D. Goldhaber (2017). Selection into online community college courses and their effects on persistence. *Research in Higher Education*, 58(3), 244-269.
- Ishitani, T. and S. L. DesJardins (2002). A longitudinal investigation of dropout from college in the United

- States. *Journal of College Student Retention*, 4 (2), 173–201.
- Lee, Y. and J. Choi (2011). A review of online course dropout research: implications for practice and future research. *Educational Technology Research and Development*, 59(5), 593-618.
- Levy, Y. (2007). Comparing dropouts and persistence in e-learning courses. *Computers and Education*, 48(2), 185–204.
- Marsh, G. (2014). Institutional characteristics and student retention in public 4-year colleges and universities. *Journal of College Student Retention: Research, Theory & Practice*, 16(1), 127-151.
- McPherson, M. S. and L. S. Bacow (2015). Online higher education: Beyond the hype cycle. *The Journal of Economic Perspectives*, 29(4), 135–153.
- National Center for Educational Statistics. (2016).
https://nces.ed.gov/programs/digest/d16/tables/dt16_311.15.asp
- National Conference of State Legislatures. (2015). Retrieved August 18, 2018 from
<http://www.ncsl.org/research/education/higher-education-legislation-in-2015.aspx>.
- Patterson, B. and C. McFadden (2013). Attrition in online and campus degree programs. Retrieved from <http://www.westga.edu/~distance/ojdla/summer122/patterson112.pdf>.
- Porter, S. R. (2000). The robustness of the “graduation rate performance” indicator used in the U.S. News and World Report college ranking. *International Journal of Educational Advancement*, 1 (2), 1030.
- State Higher Education Executive Officers Association (2017). State higher education finance: FY2016, p. 20. Retrieved February 13, 2018 from http://www.sheeo.org/sites/default/files/project-files/SHEEO_SHEF_2016_Report.pdf.
- Tello, S. F. (2007). An analysis of student persistence in online education. *International Journal of Information and Communication Technology Education*, 3(3), 47–62.
- Titus, M. A. (2004). An examination of the influence of institutional context on student persistence at 4-year colleges and universities: A multilevel approach. *Research in Higher Education*, 45 (7), 673–699.
- Weber, D. A. and R. G. Ehrenberg (2010). Do expenditures other than instructional expenditures affect graduation and persistence rates in American higher education? *Economics of Education Review*, 29(6), 947-958.