Making the connection: broadband access and online course enrollment at public open admissions institutions*

Benjamin T. Skinner†

1 January 2019

Abstract

Postsecondary students increasingly enroll in online courses, which have the potential to further democratize higher education by expanding access for historically underserved populations. While a number of studies have investigated student outcomes in online courses, past data limitations have hindered robust examination of a potential mechanism underlying the decision to enroll in an online course: access to high speed broadband. With data from the National Broadband Map and IPEDS, I fit a number of Bayesian regression models to investigate the relationship between various measures of broadband access—download speed, upload speed, and the number of providers—and the number of students who take online courses at public colleges and universities with open admissions policies. Results show that increases in broadband speed at the lower end of the speed spectrum are positively associated with the number of students who take some of their courses online, but that the marginal gain diminishes as speeds increase. This finding suggests that there may be a minimum threshold of necessary broadband access, beyond which increases in speed become a less important factor in the take up of online coursework. Open admissions colleges seeking to improve access for local students through increased online course offerings should consider broadband access in the area, particularly if the targeted populations live in communities with low average broadband speeds.

*Published in Research in Higher Education at: https://link.springer.com/article/10.1007/s11162-018-9539-6

†Email: btskinner@virginia.edu. This research was supported in part by the Bonsal Applied Education Research Award. I thank two anonymous reviewers as well as Angela Boatman, Josh Clinton, Stephen DesJardins, Will Doyle, Carolyn Heinrich and participants at the 2015 ASHE annual conference for their helpful comments and suggestions. This article is better for their help. All findings and conclusions remain my own.

1 Introduction

The number of students who enroll in online courses has increased dramatically since the early 2000s (Radford and Weko 2011; Snyder, Brey, and Dillow 2016). Though the popular press has often focused more on massive open online courses, or MOOCs, (e.g., Pappano 2012), relatively few college students take only distance education credits. Many instead split their time between face-to-face courses on campus and those offered online through their home institution (Snyder, Brey, and Dillow 2016; Snyder and Dillow 2015). As more people decide to pursue postsecondary education, it is likely that a significant portion of the higher education sector's expansion will occur online.

In the midst of this growth, those studying the outcomes of students who enroll in online course sections have found mixed results. Some studies that compare online students to those who enroll in face-to-face sections have found little to no average difference in end-of-course grades for online students (Bowen et al. 2012; Figlio, Rush, and Yin 2010; Joyce et al. 2014) and even some evidence for a positive association between first-year online course enrollment and persistence to earning an associate's degree (Ortagus 2018). Other studies, however, have shown the opposite: that students who enroll in online courses, particularly those who attend broad access institutions such as community colleges, may perform worse in terms of course persistence, grades, subsequent enrollment, and eventual degree attainment (Bettinger et al. 2017; Hart, Friedmann, and Hill 2018; Huntington-Klein, Cowan, and Goldhaber 2017; Xu and Jaggars 2011; Xu and Jaggars 2013).

Largely unexamined in this literature on outcomes, however, is a key mechanism through which most online students access their work: high speed broadband. Students who do not have access to quality broadband cannot take online courses. Existing disparities in broadband access (Federal Communications Commission 2016a; Grubesic 2008b; Grubesic and Murray 2002; Prieger and Hu 2008) are likely to lead to disparities in online education, which, as more courses become digital, could exacerbate inequities currently seen across the higher education sector as a whole. This concern is especially salient for public colleges and universities with open admissions policies that, due to strong equity missions, enroll greater numbers of first-generation and non-traditional students (Cox 2006).

In this paper I investigate the relationship between access to high speed broadband and the number of students at public universities and community colleges with open admissions policies who take some of their courses online. To do this I rely primarily on enrollment data from the Integrated Postsecondary Education Data System (Education 2015) and census block-level measures of broadband access from the National Broadband Map (National Broadband Map) from 2012 to 2014. I specifically operationalize broadband access using three key measures: download speed, upload speed, and the number of providers. In order to approximate the download/upload speeds and number of providers experienced by the average student enrolled at each institution in my sample, I assign each school broadband measures that are the population/inverse distance-weighted averages of those recorded in surrounding census block groups. Due to the nested nature of the data, with schools located in states that have different postsecondary policy contexts, I fit multilevel Bayesian regression models in addition to single level pooled models in order to better account for potential state-level differences in access to online learning.

Among institutions with open admissions policies, I find a generally positive association between download speed and the number of students who take some of their courses online. The marginal gain decreases, however, as download speeds increase. Whereas moving from tier 4 ($\approx$ 1.5–3 megabytes/sec) to tier 5 ($\approx$ 3–6 megabytes/sec) download speeds is associated with a 14% average increase in the percentage of students who take some courses online, a similar one tier increase from tier 7 to 8 (from $\approx$ 10–25 megabytes/sec to $\approx$ 25–50 megabytes/sec) is only associated with an average 2% increase. This finding suggests that after a certain threshold of download speed is met, subsequent speed increases do not make online coursework more attractive on the margins. I do not find strong evidence of similar associations for upload speed or the number of internet service providers in the area.

On the whole, this paper offers the first empirical evidence of a generally positive relationship between download speed and enrollment in online courses at public colleges and universities with open admissions policies across the country. If, as part of their equity mission, these institutions seek to increase digital offerings, the results from this study suggest that broadband access among the populations they wish to serve should be an important consideration.

2 Literature review

2.1 Rise of online education

Over time, a number of technologies have been used to make education at a distance possible. Correspondence courses that took advantage of the postal system in the 19th century mark a more formal start to higher education at a distance (Johnson 2003). Moore and Kearsley (2011) note five generations of distance education: correspondence, broadcast radio and television, open universities (as seen in the United Kingdom and Australia, among other countries), teleconferencing, and the Internet/Web. With each new generation, improvements in technology meant potential improvements in the delivery of education. While radio and television broadcasts were faster and potentially more inclusive than direct mail, teleconferencing and the Internet once again "allowed a student to answer back," Moore and Kearsley 2011, p. 36.

The United States Armed Forces itself took advantage of distance education during the Second World War. Under the theory that better educated soldiers made better all-around soldiers, the federal government built a robust system of correspondence courses that troops could (and did) take during their downtime (Loss 2012). Though these courses gave way to temporary place-based universities that soldiers attended before returning home at the end of the war, the idea that higher education was fit for many more than society's elite found its first application in these correspondence courses.

In the postwar period, a number of other countries similarly decided to make higher education available to new populations of students. The Open University of Great Britain, founded in the late 1960s, was new university model that used distance education technologies to enroll students from all over the country (Moore and Kearsley 2011). Not a part of existing universities, it instead combined distance learning technologies, distributed brick-and-mortar locations around the country, and a "radical open admissions policy" (The OU story) to enroll a large number of students. Still operating today, the Open University has seen countries around the world borrow its name and model for their own similarly structured institutions (Keegan 1996).

Aside from experiments among some state university systems such as the Articulated Instructional Media Project (AIM) at the University of Wisconsin (Moore and Kearsley 2011), the United States has not attempted to develop an open national university along the same lines as Britain's

Open University. This is in spite of the fact that the number of college-going students in the United States has steadily risen since the end of the Second World War. From the early 1960s to 2010, the number of postsecondary enrollments increased from fewer than 5 million students to over 20 million, a sizable proportion of which came from non-traditional, historically underserved populations (Snyder and Dillow 2015). Enrollments at large public universities greatly increased to meet this new mass demand for higher education, as did the number of two-year institutions (Cohen and Brawer 2003; Crookston and Hooks 2012; Thelin 2011).

One popular solution to the problem of rising enrollments of non-traditional students coupled with declining funding (Carlson et al. 2015; Tandberg 2010) has been to expand online education offerings through existing colleges and universities[1] (Bowen 2013; Deming et al. 2015; Goldrick-Rab 2010). Online learning has been touted as a way to educate students in a flexible and cost-effective manner (Moloney and Oakley 2010). Such a move is also in line with the equity mission found at open access public institutions (Cox 2006).

While some researchers have held up massive open online courses, or MOOCs, as models for a new type of higher education (Bowen 2013; Selingo 2013), the newest digital revolution in higher education has not been limited to these unique courses. Public institutions in many states, especially those at the two-year level, have also steadily increased the number of online credit hours they offer in the past few years (Allen and Seaman 2011; Allen et al. 2016; Radford and Weko 2011; Southern Region Educational Board 2013; Southern Region Educational Board 2015). In the fall of 2012, 27% of all college students took at least one online course, with 13% completing all coursework through online classes (Snyder and Dillow 2015). Two years later in the fall of 2014, those same numbers had increased to 28% and 14%, respectively (Snyder, Brey, and Dillow 2016). Unlike MOOC students, many of whom take courses anonymously and for no credit, these data show that formally enrolled students often incorporate online courses into their degree pathways, taking them alongside more traditional face-to-face courses.

From correspondence classes meant to replace traditional higher education programs to online course sections as simply another option in the menu of higher education choices, distance education has remained one way to open higher education to a greater number of people than possible at traditional campuses. But as the technology of distance education has improved—from mail, radio

[1] cf. the Open University and its founding as an autonomous institution.

and television to the Internet—so too have the requirements for participation increased. Where an address, paper, and pen had once sufficed, a computer and steady Internet connection are now required to take most distance education courses. Students who do not have these technological resources may find themselves effectively shut out. The open access rhetoric of online education, therefore, need first acknowledge the digital divide between those with access to the Internet and those without.

2.2 The digital divide: an overview

Scholars and policy-makers have long noted the divide between those who have access to communications technology and those who do not (Brown, Barram, and Irving 1995; Irving et al. 1999; McConnaughey and Lader 1998). Of recent concern is the digital divide between those with access to broadband and those without (Federal Communications Commission 2016a). Because a large majority of persons in the United States have access to some form of broadband, researchers have transitioned from questions that ask if persons have access to questions that ask what kind of access (Brown et al. 2010). This represents an important shift as research has shown that local infrastructure, regardless of the relative affluence of the population, can have a major impact on the availability of service in a particular area (Grubesic and Murray 2002). Even within a socioeconomically homogeneous local area, topological features can cause the quality of broadband connections to vary substantially across households and neighborhoods (Oyana 2011).

Local variations notwithstanding, less affluent rural areas have generally had poorer access to broadband than wealthier and more urban areas (Brown, Barram, and Irving 1995; Copps 2009; Federal Communications Commission 2015). One reason for this disparity lies in prohibitive "last-mile" infrastructure costs that communications firms, local governments, and residents are reluctant to cover (Grubesic and Murray 2004). In rural, suburban, and urban areas alike, those of lower socioeconomic status usually have the least access to broadband (Horrigan 2010). Even the near ubiquity of cellular and satellite technology does not close the gap since wireless services still cannot compete with wired services in terms of speed and reliability (Brown et al. 2010) or coverage (Grubesic 2012b).

Due to growing concern over these issues, Congress charged the Federal Communications Commission in 2009 with instituting the National Broadband Plan. This plan represents a concerted national attempt to close the digital divide by making sure that "every American 'has access to broadband capability' " Federal Communications Commission 2009, p. XI, a goal that is seen as worthwhile due its economic, civic, and educational benefits (Copps 2009; Czernich et al. 2011). Since the inception of the plan, a number of government reports have been issued detailing the status of broadband in the country (latest: Federal Communications Commission 2016b). Data collected from ISPs about broadband penetration has also been opened to the public in the form of the National Broadband Map (National Broadband Map). While these data come with their own limitations mostly owing to the size of the raw data (Grubesic 2012a; Grubesic 2012b), they nonetheless represent a vast improvement over previously available data on broadband (Grubesic 2008a; Grubesic 2008b; Grubesic 2008c) and make possible new quantitative research on the effect that being on the wrong side of the digital divide may have on educational access and outcomes.

2.3 The digital divide in higher education

Regarding the effect of the digital divide on students, the National Broadband Plan's founding document notes that

> [t]oday, millions of students are unprepared for college because they lack access to the best books, the best teachers and the best courses. Broadband-enabled online learning has the power to provide high-quality educational opportunities to these students—opportunities to which their peers at the best public and private schools have long had access. (Federal Communications Commission 2009, p. 5)

In concurrence, a number of studies have highlighted the negative effects of the digital divide for communities and students in postsecondary institutions. From the early days of the Internet, scholars have noted that white students were more likely than their African-American peers to have computers in their homes and to have used the Internet (Hoffman and Novak 1998). Even in later years as more students gained access to the Internet, differential usage across gender and racial/ethnic groups in terms of communication and academic usage suggest a continued divide (Cotten and Jelenewicz 2006; Jones et al. 2009).

Two recent studies in particular reveal how infrastructural differences in broadband access among higher education students remain. Using GIS spatial data to examine the availability of broadband in the region around Southern Illinois University Carbondale, Oyana (2011) found that broadband quality was not uniform in the region and was generally inferior in poorer and more rural areas. Presenting data on the quality of signal as function of land topography, he showed a correlation between signal quality and median household income. Based on his analysis, Oyana concluded that without improvements to broadband infrastructures, much of the southern Illinois area under study could not support the large data requirements of the types of libraries and research labs that area students would need in order to have the same educational advantages as their peers who live in areas with better broadband access.

Hurst (2010) highlighted similar concerns for knowledge production in low broadband areas with a survey that asked students at Walters State Community College in eastern Tennessee to describe their home broadband access and how it related to and affected their coursework. He found that while 65% of his survey respondents ($N = 740$) said they felt having broadband access was very important for completing their schoolwork, 20% reported having no Internet access or only dial-up at home and 30% reported feeling dissatisfied with their broadband quality. He also found a statistically significant relationship between having faster internet speeds at home and the propensity for taking an online course.

Both of these studies provide suggestive evidence of a gap in broadband access that may make it more difficult for rural and low-income students to fully participate in courses that require online coursework. The authors' respective findings are echoed by others who suggest that some areas of the country lack access to the quality broadband, computers, and human capital required to successfully integrate online coursework into higher education (Cejda 2007). These studies, nonetheless, have limitations. Singly, each considers only a small local area; together, they focus on the mid-southern region of the country. The generalizability of Hurst's findings are further limited by the fact that his analyses do not utilize metrics of broadband connectivity but instead rely exclusively on survey data.

This paper adds to the literature on online higher education by rigorously investigating the connection between broadband access and the take up of online courses across the country. This study focuses specifically on open admissions colleges and universities due to the convergence of their equity-based missions and the potential for online course offerings to further democratize access to higher education. With its national scope and novel data, this study further bridges the gap between research on the digital divide and online education. Though it may seem obvious that increased download speeds should be connected to increased online-based distance enrollment, the uncertainty surrounding the demands of online coursework may mean that students do not readily make the connection.

3 Theoretical framework

Under the human capital model of college enrollment, the enrollment decision requires a weighing of the potential gains of increased education against its costs (Becker 2009; Turner 2004). For many students, gains are realized as improved job prospects or increased wages (Eagan et al. 2014; Fishman 2015). College costs generally include direct costs such as tuition, books, fees, room and board as well as indirect opportunity costs like forgone wages (Manski and Wise 1983).

A key selling point of online courses is their potential to lower costs for both the institution and the student (Deming et al. 2015). For the institution, online students do not require physical plant space or the costs associated with using the space: desks, lights, chalk/whiteboards, paper handouts, electricity, support staff to maintain the space, etc. For this reason, the marginal cost of the course in terms of an additional student is close to zero. For courses in which grading may be largely automated (e.g., those that rely on multiple choice tests rather than written assessments), the cost of an additional student in terms of instruction is also effectively zero. If most online courses are asynchronous, meaning that students are not required to view course materials at specific class meeting times but rather can do so when they choose, then schools can save money by reducing or eliminating off-hour course sections in which the student-instructor ratio may be comparatively lower and less cost-effective.

Students may also see online courses as personal cost reducers. Because online courses do not require travel to campus, a student can save on travel-related expenses such as gas or public transportation costs. Asynchronous courses do not demand a set block of time in which to work on them. Students in these types of online courses can self-pace and work at times with lower opportunity costs. The flexibility given by the online course may allow for a greater range of options in other areas—work, family—that themselves may not be as time flexible (Jaggars 2014).

Online courses, however, may increase rather than decrease some educational costs. These additional costs may be difficult to quantify due to informational asymmetry and uncertainty on the part of the student. When deciding whether to enroll in an online course, an important consideration is whether one has access to the quality of broadband required to successfully complete the course. Slow speeds, poor connections, and high connection prices may all reduce the likelihood of success by increasing the direct and indirect costs of enrollment. More money spent on improved broadband access is less money to spend on books or other supplies; more time spent completing and submitting online assignments due to low quality broadband is less time to complete other assignments, work, or spend with family.

At the time of enrollment, students may have difficulty connecting perceptions of their broadband connections as experienced through other devices and services (e.g., cellular phone) to the demands of online coursework. Students who are older or who face socioeconomic barriers to their enrollment may have had little experience with broadband upon which to base their enrollment decision. It could be the case that broadband access is moot for students with constrained choice sets that effectively limit their options to online work. For these reasons, it is unclear whether broadband access should be correlated with the decision to enroll in online courses in the aggregate.

It is outside of the scope of this paper to fully investigate whether students perceive online courses as increasing or decreasing the costs associated with their enrollment decision. Instead, I focus on a single aspect—the association between broadband access and enrollment in online courses. The limited prior literature suggests that students with better access to quality broadband connections are more likely to attempt and succeed in online coursework than their peers who lack access (Hurst 2010). If individuals respond this way, it may also be true that institutions in areas with better broadband are more likely to offer course sections in an online format and encourage their faculty to move some coursework online into a hybrid setting. In either of these situations, I would expect measures of broadband access to be positively correlated with the number of students attempting online courses. If uncertainty surrounding online courses or broadband connectivity is high then students may perceive online courses to be a less sure bet and avoid them when possible. Or if schools have other prerogatives regarding online education that outweigh considerations of broadband access (e.g., cost reductions), broadband access among students they hope to enroll may be moot. In either of these latter scenarios, I would not expect a strong connection between broadband measures and online enrollment.

To investigate the relationship between broadband and online enrollment at open admissions institutions, I make two key assumptions. First, I assume that the broadband measures I compute and assign to each institution in my sample (a process discussed in more detail in section 5.2) also apply to students who take online courses at the school. To support this assumption of student proximity, I only use the number of students who take some of their course online—that is, those who split their coursework between face-to-face and online courses—as the dependent variable in the models I estimate. While it is logical to think that students who need to be on campus to attend their face-to-face courses live nearby, prior research and national survey data also support this assumption in that they show that most students attend schools close to home (Mattern and Wyatt 2009) and that students who take some courses online are more likely to live nearby than students who enroll entirely in online courses (U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics 2012).

Table 1: Distance from student's home (in miles) to NPSAS institution by percentage of courses taken completely online and institution level

	< 21	21–50	51–100	101–500	> 500
Total	59.9	13.6	7.2	11.8	7.6
% of classes taken completely online					
All classes					
Public two-year	64.6	15.4	7.4	8.5	4.1
OA public four-year	35.4	18.1	7.4	14.1	25.0
Some classes					
Public two-year	78.1	14.9	3.4	2.6	1.1
OA public four-year	66.8	13.8	4.2	9.9	5.2

Notes. OA: Open admissions. Data come from the U.S. Department of Education, National Center for Education Statistics, 2011-12 National Postsecondary Student Aid Study (NPSAS:12). All numbers are percentages, which add to 100% within each row. Computation by NCES PowerStats.

Table 1 shows that in the years just before the study period, approximately three quarters of all postsecondary students attended college within 50 miles of their permanent address. Among students who attended public two-year colleges and took their courses entirely or in some part online, the majority lived within 20 miles, with a comparatively greater share of sometimes online students living that close (78.1%) than entirely online students (64.6%). Among students who attended a public four-year institution with an open admissions policy and took all or some of their courses online, the difference in proximity between the groups is larger. While 35.4% of entirely online students lived within 20 miles of the school, 66.8% of sometimes online student did. Unlike at two-year colleges, however, a substantial percentage entirely online students at the four-year level, 25%, had homes more than 500 miles distant; only 5.2% of sometimes online students lived this far away. Supported by these descriptive survey data, I argue my assumption of student proximity to institutions in the sample is best held for students who take some, but not all of their courses online.

As a second assumption, I assume that the outcome of interest involves online coursework rather than other methods of distance education (e.g., television, radio, CDs, mail-based correspondence). While the IPEDS variable I use to get the number of students who took some distance education courses allows for many types of distance learning technology, historical trends and the current technological landscape (Allen and Seaman 2011; Allen et al. 2016; Moore and Kearsley 2011) mean that online work makes up the vast majority, if not functionally all, of the distance coursework during the years under study.

4 Estimation strategy

I estimate the association between three measures of broadband access—download speed, upload speed, and the number of providers—and the number of students enrolling in some online courses using both Bayesian single and multilevel regression models. The Bayesian paradigm, as opposed to a frequentist paradigm, is warranted in this instance for two primary reasons.

First, the data I have represent the entire population of public open admissions postsecondary institutions that have students who enroll in online courses. Furthermore, these data are situated in their specific historical context, a time in which online-based online courses were growing in popularity and the technological means by which they were made possible were also developing and improving (Allen et al. 2016; Federal Communications Commission 2016a; Snyder, Brey, and Dillow 2016). A repeated sampling framework, on the other hand, is predicated on the belief that the analysis data represent a random sample of data from a population. The key assumption is that under repeated samples, which would realize different analysis data, fitted models would produce parameters that would converge to the true and fixed values. Were I to gather these institutional and broadband data again, however, I would not expect differences since the data represent the entire population of interest (Western and Jackman 1994). A Bayesian approach, which instead views the data as fixed and the parameters as random, is more applicable (Jackman 2009).[2]

A second reason to employ a Bayesian framework is the nested nature of the data. Because schools are nested in states, multilevel models that can account for correlation between units in groups may be preferred to single-level linear models (Gelman et al. 2014). The nature of the Bayesian framework is such that it is straightforward to move from single-level to multilevel models in terms of estimation and interpretation.

In the analysis, I first estimate a number of single-level Bayesian linear regression models that take the form

$$log(y_i) \sim N(\alpha + \beta Broadband_i + X\gamma, \sigma_y^2), \tag{1}$$

where y_i is the number of students who enroll in some online courses; α is a constant term; β is the parameter of interest for $Broadband_i$, the institution's assigned measure of broadband; and X is a matrix of covariate data values with γ as its corresponding vector of parameters. Because the number of students who take some online courses is right skewed, I fit the log transformation of these values. This procedure normalizes the outcome, which allows me to use a normal likelihood function with a variance of σ_y^2. Using the natural log of the outcome has the added benefit of making β represent the percentage change in the number of students who take some online courses for each unit increase in the broadband measure of interest (Greene 2012).

A key feature of Bayesian estimation is that the researcher must place prior distributions on all non-fixed parameters. One way of interpreting Bayesian analysis is that it is an update of prior beliefs using information provided by new data. More formally, a Bayesian posterior distribution, $p(\theta \mid X)$, is proportional to the joint probability of the unknown parameters and data, $p(\theta, X)$, which can be decomposed into the prior, $p(\theta)$, multipled by the likelihood, $p(X \mid \theta)$, or $p(\theta)p(X \mid \theta)$.[3]

[2]Theoretical justifications notwithstanding, it is important to note that when weakly-informative priors are used in applied analyses, otherwise comparable Bayesian and frequentist models generally produce substantively similar estimates. Prior distributions are discussed later in the section.

[3]A Bayesian posterior distribution is properly written as

$$p(\theta \mid X) = \frac{p(\theta)p(X \mid \theta)}{p(X)},$$

Broadly, priors may be weak or strong. For example, the regression coefficients α, β, and γ in equation (1) might be assigned normal distributions with means at zero and variances large enough to allow for a wide range of possible posterior values, positive or negative. A stronger prior might assign a non-zero mean and smaller variance to one or more parameters, signalling more strongly held beliefs. Priors may also come from prior research or be elicited from those with domain expertise (Gill and Walker 2005).

A Bayesian posterior distribution, therefore, is therefore a compromise between the prior and likelihood distributions that weights each by the strength of its information. Strong priors in the face of a weakly-informative likelihood result in posterior beliefs that are not much changed. Weak priors estimated alongside a strong likelihood allow the data to "speak for themselves" (Gelman et al. 2014), giving results that are generally similar to those returned in a comparable frequentist analysis. In the single-level models described by equation (1), I utilize weakly-informative priors meaning that results are driven by the data and may be interpreted much as they would be were they generated using a frequentist estimation.[4]

To account for the nested structure of the data, in which institutions are located in states, I also fit Bayesian multilevel models that allow model intercepts to vary with each state. One particular advantage of multilevel models over single-level models is that they allow information to flow between observations within groups. To account for state-level differences in the number of students at open admissions institutions who take some of their courses online—which may be warranted due unique higher education policy contexts across the states—I could conduct separate estimations for each state. Due to small numbers of observations in some states, however, this procedure would produce noisy estimates. The multilevel model allows for partial pooling of the estimates across states. Estimates for states with few observations can "borrow strength" (Jackman 2009) from other states based on their group-level characteristics. With this "bias/variance" trade-off (Carlin and Louis 2009; Gelman et al. 2014), I am be able to estimate the association between broadband access and the number of students enrolled in online courses that takes into account

which includes the probability of the data, $p(X)$. Because data are assumed fixed in many analyses, $p(X)$ may be treated as a normalizing constant and dropped, yielding, $p(\theta \mid X) \propto p(\theta)p(X \mid \theta)$.

[4]When coding the model, I technically use improper priors. Priors are improper when $\int p(\theta)d\theta \neq 1$, that is, the probabilities do not sum to one. All regression coefficients are drawn from a uniform distribution with support on $\theta \in (-\infty, \infty)$ and all variances from a positive uniform distribution: $\theta \in (0, \infty)$. Improper priors may combine with a likelihood function, however, to produce proper posterior distributions (Gelman et al. 2014).

differences across the states, even when some states have relatively few observations.

With the multilevel model, each state, represented by α_j, is allowed to have its own intercept. It takes the form

$$log(y_i) \sim N(\alpha_j + \beta Broadband_i + X\gamma, \sigma_y^2) \tag{2}$$
$$\alpha_j \sim N(\delta_s Region_s + Z\psi, \sigma_s^2),$$

in which X represent a vector of school and county level covariates. Each state intercept, α_j, is modeled using state-level covariates, Z, and a region-specific intercept, δ_s. As with the single-level models, all unknown parameters $(\beta, \gamma, \psi, \delta_s, \sigma_y^2, \sigma_s^2)$ are given diffuse priors.

I fit both single-level and multilevel models using each measure of broadband separately and together in a single equation. This produces four models—one each for download speed, upload speed, and the number of providers and a joint model with all three—for each model type for a total of eight sets of results. Each model includes measures of broadband at both the level and squared value as well as a number of covariates taken from a variety of data sources. These are discussed in more detail in the next section.[5]

5 Data

5.1 Institution data

Data on the number of students who enroll in online courses were taken from the Integrated Postsecondary Education Data System (Education 2015). Though the IPEDS survey has asked institutions if they are primarily distance learning schools for a number of years (a binary yes or no response), it has only asked institutions to break down the total number of students who attempt distance learning coursework since the fall of 2012.

IPEDS defines distance education as "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." (IPEDS

[5] In an alternative specification, I use the proportion of students as the outcome of interest. To properly model the proportion, which is bounded by [0,1], I use a beta likelihood function. The results for these models are qualitatively the same as those given by the log outcome/normal likelihood models, so I present the latter for ease of interpretation. More details about the beta likelihood specification as well results from the models are shown in Appendix A.

Glossary) The code book further defines the technologies that may fall under the heading of distance education as including "Internet; one-way and two-way transmissions through open broadcasts, closed circuit, cable, microwave, broadband lines, fiber optics, satellite or wireless communication devices; audio conferencing," (IPEDS Glossary). While not all of these technologies are strictly broadband-based, it is likely that most distance education students at public open access institutions in the study period experienced distance education through online portals (Allen and Seaman 2011; Allen et al. 2016; Moore and Kearsley 2011). For this reason, I refer to these students as online students.

Specifically, I use the IPEDS variable that gives the number of students who took some of their courses online as the primary outcome.[6] This number indicates the number of students who enrolled in both online and face-to-face courses. Because the data do not include the number of online courses that students took, the dosage of online course-taking within and between schools is unobserved. Sometimes-online students need only take one of each type of course to be labeled as such. This means that whereas some students represented in the data could have had balanced course loads, others were predominantly online students who took one face-to-face course or mostly in traditional courses with a single online course enrollment. Though it may be the case that broadband speeds are positively correlated with the number of online courses attempted, these data do not support this particular analysis.

Using the theoretical framework as a guide, I include a number of school-level covariates in all model specifications. Based on the mixture of potential costs and benefits to online learning, it is possible that student populations will be differentially affected. To account for potential heterogeneity in response to online coursework, covariate models include, at the institutional level, proportions of students who receive Pell grants, students of color, women, students who are 25 years old or older, and part-time attenders. Only undergraduate enrollments are considered. Models also include indicators for status as a two-year college and whether the institution provides on-campus housing. Though the weighted broadband measures apply to students living in housing provided by the college, including the indicator controls for potential differences in how residential and non-residental students associate broadband with online course enrollment. As further control measures, all models include the log transformation of the total number of students enrolled and indicators

[6]IPEDS variable: EFDESOM.

for the survey year of the observation.

Table 2: Descriptive statistics of the institution sample

	Mean/(SD)
Total enrollment	7725
	(7819)
Some online enrollment	1361
	(1536)
Two year institution	0.88
	(0.33)
Has on-campus housing	0.25
	(0.44)
Non-white enrollment	0.42
	(0.24)
Women enrollment	0.42
	(0.07)
Pell grant recipients	0.42
	(0.15)
Part-time enrollment	0.57
	(0.15)
Aged 25 years and older	0.37
	(0.11)
2013	0.4
	(0.49)
2014	0.3
	0.46
N (2012)	750
N (2013)	1003
N (2014)	741

Notes. Total enrollment and some online enrollment represent the average number of students rounded to nearest student. Other rows are proportions. Standard deviations are shown in parentheses. Schools included in the sample are public, open admissions postsecondary institutions that report at least one student who took some distance education courses.

Table 2 shows means and standard deviations of the numbers of students taking courses for credit, those taking some online courses, and other institutional covariates. Schools in two states, Alaska and Hawaii, are dropped from the final analysis data set due to their unique contexts—Alaska as a large, but sparsely populated state, and Hawaii an island group—that may bias the weights used to construct the broadband measures that I assign to each institution. Because data for all branches of Indiana's public two-year institution, Ivy Tech, were aggregated under a single identification code, these institutions were also dropped from the data set as broadband measures could not be accurately assigned to them. The final estimation data represent 1,017 unique public open admissions institutions observed across three years.

5.2 Broadband data

Broadband data were collected from the National Broadband Map website.[7] Gathered at the behest of the National Telecommunications and Information Administration in partnership with the Federal Communications Commission (FCC), these data were collected from internet service providers (ISPs) within each state by appointed grantee agencies. Each service provider gave information about upload and download rates at the census block level as well as information on the number and types of community anchor institutions (typically libraries, K-12 schools, college, etc.) in the area (National Broadband Map). These data were corroborated against other sources of broadband information and released to the public. Data released in June 2012, 2013, and 2014 were used since they were collected around the same time period as the IPEDS data and represent the best estimates of broadband connectivity surrounding study institutions during the time period under study.[8]

Internet service providers reported measures of broadband speed in ordered categories that range from 1 (greater than 200 kilobytes/sec and less than 768 kilobytes/sec) to 11 (greater than 1 gigabyte/sec). Past FCC guidelines suggested that the minimum required speed to watch university lectures was 4 megabytes/sec, which would fall under category 5.[9] Acknowledging the increasing "speeds required to use high-quality video, data, voice, and other broadband applications," (Federal Communications Commission 2015, p. 3) as well as the demands placed on broadband connections by multiple users within the average household, the FCC recently updated these benchmarks, arguing that "having 'advanced telecommunications capability' requires access to actual download speeds of at least 25 Mbps and actual upload speeds of at least 3 Mbps," (Federal Communications Commission 2015, p. 3). These new benchmark speeds fall under category 8.

A primary question for this analysis: how does one assign a broadband measure to each institution? Were the unit of analysis the student and home address known, I could simply assign each student the broadband measures of their home census block group. But because the unit of analysis

[7] www.broadbandmap.gov

[8] Though subsequent years of student enrollment data have since been released, the National Broadband Map stopped being updated in June 2014. I have not incorporated the newest enrollment data for this reason.

[9] www.fcc.gov/guides/broadband-speed-guide

is the school, I am unable to see where each student lives and must assign broadband measures to the institution that take into account those experienced by the average student who is enrolled there.

One solution would be to assign each school the measures of its census block, census tract, or county. Yet postsecondary students, especially those who are part-time, older, or attend non-residential schools, may not live in the same county as the one in which their institution is located. To assign each institution the average download speed in its county, for example, may neglect the broadband experience of students who live in a very large county or in an adjacent county with very different download speeds. Aggregating the measures up to the county level would also mean losing the fine grain differences in broadband access within each county.

To better approximate the most likely distribution of students around institutions, I instead assign each school broadband measures using a weighting process that takes into account both its distance to surrounding census block group-level measures and the population sizes of those census block groups. The parts of the combined inverse-distance/population weight are discussed in turn below.

Inverse-distance weights were constructed by first computing the Great Circle distance, d_{sc}, from each institution, s, to all census block group centroids, c, in the lower 48 states, producing an $N \times K$ matrix where $N = \#$ schools and $K = \#$ census block groups. For the year 2013, a $1,004 \times 217,290$ matrix of distances was the result. So that download speeds of nearby census block groups would count more than those of distant census block groups, these values were inverted and scaled to create inverse distances, id_{sc}:

$$id_{sc} = \frac{1}{(d_{sc})^r}. \tag{3}$$

Inverse-distance weights were created by dividing each inverse distance over the row sums of the inverse distances (each row representing the distance between a single institution and all census block group centroids):

$$idw_{sc} = \frac{id_{sc}}{\sum_{c=1}^{C} id_{sc}}. \tag{4}$$

Were the data used in this study at the student level, the inverse distance weight would be sufficient to approximate the broadband measures experienced by each student (if I did not wish to simply assign them the values of their census block group). But because the data are aggregated to the institution level, I do not know exactly where the students live in relation to the school. Some may live very close whereas others commute from further away. It is also unlikely that students are evenly distributed around the school. It is more likely that students come from nearby population centers: towns, suburbs, and neighborhoods. The inverse distance-weighted average assigned to the school, therefore, is likely to be different from what would be assigned to the students were their addresses known.

To mitigate this potential bias, I employed a second weight that adjusted each school's inverse distance-weighted broadband measure average back toward values recorded in more highly populated areas. Accounting for population in spatial-weighting schemes has been shown to improve upon estimates in which only inverse distance weights are used (Hanigan, Hall, and Dear 2006). Using census block group population estimates taken from the 2010 Census, I constructed a second $N \times K$ matrix (again, $N = \#$ schools and $K = \#$ census block groups) in which each column represents a census block group's population, pop_c, repeated in each row. The population weight, pw_c, is simply

$$pw_c = \frac{pop_c}{\sum_{c=1}^{C} pop_c}, \tag{5}$$

or the population in the matrix cell divided by the row sums.[10] In applying the second weight, I make the assumption that the likelihood a student lives in a particular census block group around the institution is proportional to that block group's population size. Lest major metropolitan centers unduly skew the average too far away from the institution (e.g., census block groups in Charlotte, North Carolina, affecting averages of schools in the eastern part of the state), I use a quadratic decay ($r = 2$) in the inverse distance weight formula (Shepard 1968). At this rate of decay, the influence of a census block group's broadband values on one institutional average quickly diminishes with distance, even when population sizes are taken into account. Thus the second population weight serves as a slight correction to the inverse distance weight, which dominates in

[10]Because the census block group population values are constant, it is not strictly necessary build an entire $N \times K$ matrix, which is simply a vector of values of length K repeated N times, or compute the row sums and weights N times. For purposes of computation, however, it is easier to build a full matrix that can be easily combined with the other weighting matrix when computing the final average.

the final computation of each broadband measure average at an institution.

I used these combined weights, $w_{sc} = pw_c \times idw_{sc}$, to create weighted average broadband measures for each school, $wbroadband_s$,

$$wbroadband_s = \sum_{c=1}^{C} \frac{w_{sc} \cdot broadband_c}{\sum_{c=1}^{C} w_{sc}}, \tag{6}$$

in which $broadband_c$ is the average broadband measure—download speed, upload speed, or number of providers—in the census block group.[11] Each weighted average was computed using census block group broadband values that covered the fall term in which the online course enrollment numbers were reported. Schools that appear more than once in the data set (the majority), therefore, have distinct broadband measure averages for each year.

Figure 1 offers a stylized visualization of this weighting process. Using Nashville State Community College (NSCC) as an example, dotted lines connect its location in Davidson County, Tennessee, shown by the diamond, to a few census block group centroids in surrounding counties, indicated by black dots. (For clarity, the black dots represent a small fraction of the census block group centers that were actually used to compute weighted broadband measures for NSCC.) Among these census block groups, which have broadband measures given by the National Broadband Map, the average download speed assigned to NSCC, for example, would be most influenced by the value recorded for the centroid in Davidson County, which is black dot just the right of NSCC in the center of the map. Because the Davidson County centroid also happens to represent the largest population among the subgroup of census block groups shown, NSCC's average download value would be even further weighted towards its value.

Figure 2 shows the variability in both weighted download and upload speeds assigned to institutions in the data set. Due to the asymmetric design of most residential broadband networks (Federal Communications Commission 2016a), the distribution of download speeds is generally greater than that of upload speeds. Each plot indicates the old and new thresholds for definition as broadband. While almost all institutions in the analysis sample had download and upload speeds

[11] Combined weighted broadband equation:

$$wbroadband_s = \sum_{c=1}^{C} \frac{\left(\frac{pop_c}{\sum_{c=1}^{C} pop_c}\right)\left(\frac{id_{sc}}{\sum_{c=1}^{C} id_{sc}}\right) \cdot broadband_c}{\sum_{c=1}^{C}\left(\frac{pop_c}{\sum_{c=1}^{C} pop_c}\right)\left(\frac{id_{sc}}{\sum_{c=1}^{C} id_{sc}}\right)}$$

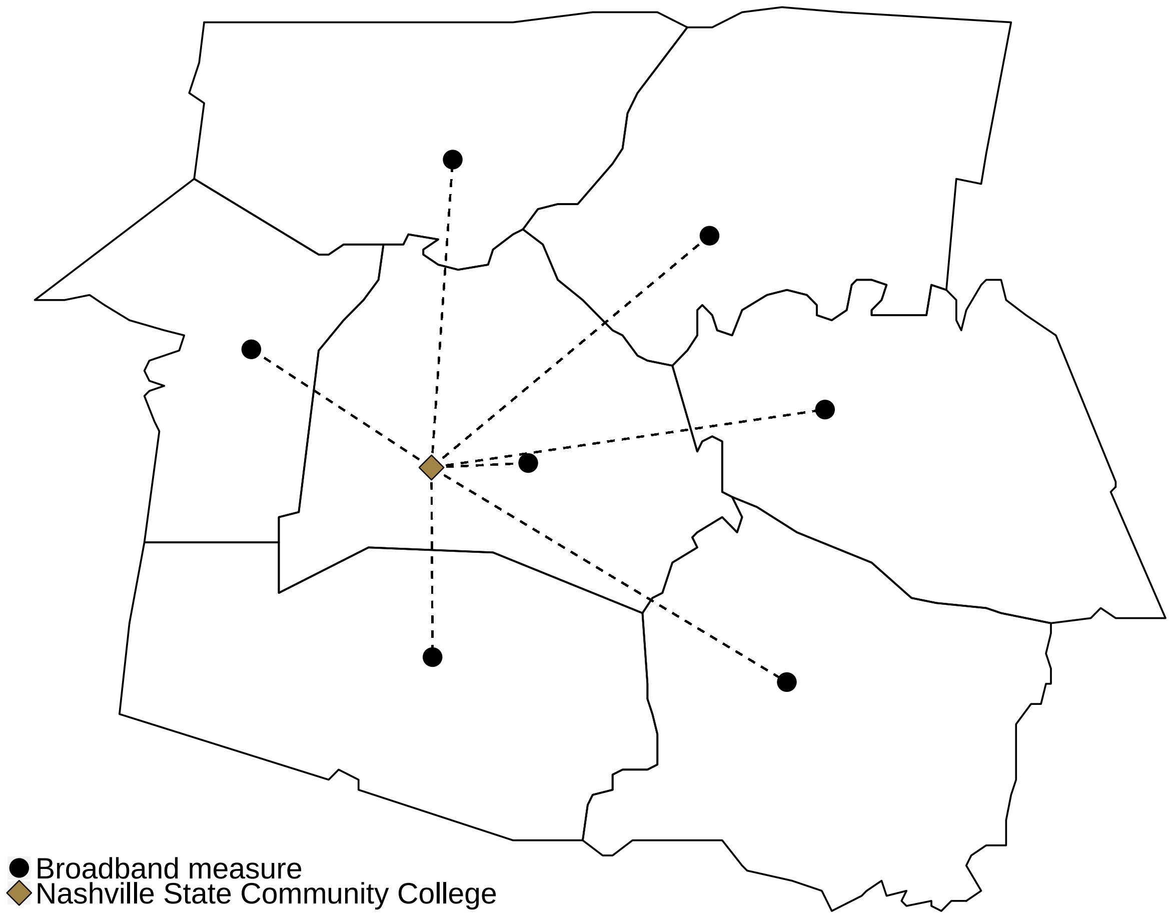

Figure 1: An illustrative example of the weighting scheme used to compute broadband measures—download speed, upload speed, number of providers—at an institution. For clarity, only one census block group per county is shown. Raw broadband measures from the National Broadband Map, which are recorded at the census block level, are first aggregated to the census block group level (represented by the dots). The distances between each institution (indicated by the diamond) and all census block group centers are then computed (represented by the dotted lines). The inverse of these distances along with each census block group's population are then used as weights in equation (6), which computes a weighted average of surrounding broadband measures for each institution, with values from nearby and more populated census block groups having comparatively more influence. In this example, weighted broadband measures assigned to Nashville State Community College would be more influenced by those of the census block in Davidson County (nearest black dot to the right) than those of outlying areas both due to its proximity and larger population.

in the surrounding area that met the benchmark for broadband during the years covered by this study (tier 5, 4 megabytes/sec, for download; tier 3, 1 megabyte/sec, for upload), the majority did not meet the definition under the new thresholds defined in 2015 (tier 8, 25 megabytes/sec, for download; tier 5, 4 megabytes/sec, for upload). Across the sample, the average number of ISPs providing broadband service around each school went from 2.73 at the 25th percentile to 3.88 at the 75th percentile.

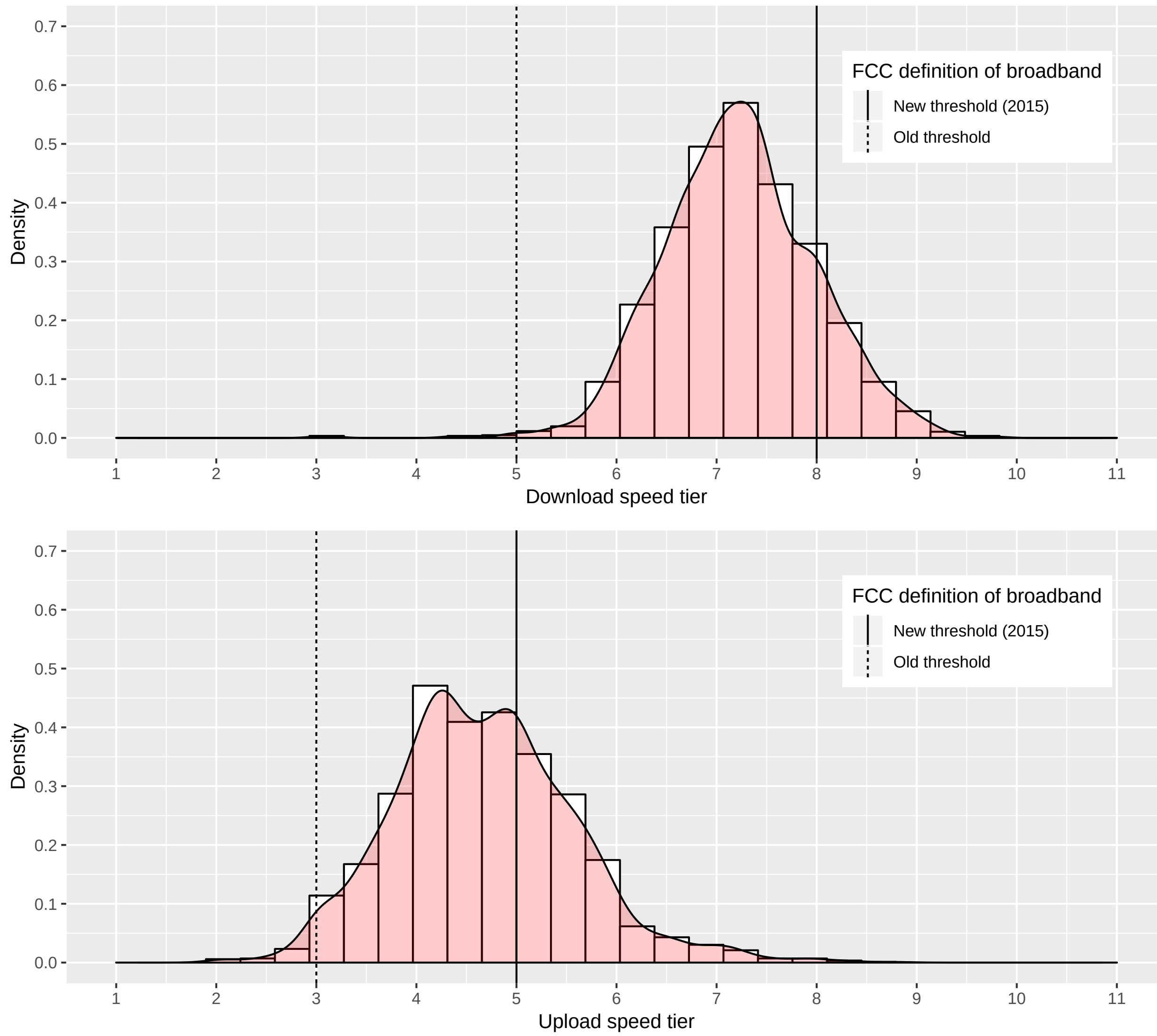

Figure 2: Histograms with overlaying density of distance- and population-weighted download (top) and upload (bottom) speeds assigned to institutions in the analytic sample. The National Broadband Map reports ordinal categories of speed from 1 (< 200 kilobytes/sec) to 11 (> 1 gigabytes/sec). Prior to 2015, the Federal Communications Commission (FCC) had set category 5 download speeds (≈ 4 megabytes/sec) and category 3 upload speeds (≈ 1 megabyte/sec) as the minimum thresholds for broadband designation. These points are indicated by the dashed vertical lines. As of 2015, the FCC upgraded the definition of broadband to category 8 for download speeds (at least 25 megabytes/sec) and category 5 for upload speeds. These are indicated by the solid vertical lines.

5.3 Geographic and demographic data

When estimating multilevel models, I use additional state level covariates to aid in fitting second level parameters. These include measures of statewide unemployment rates, which were taken from data provided by the Bureau of Labor Statistics (Bureau of Labor Statistics 2012; Bureau of Labor Statistics 2013; Bureau of Labor Statistics 2014). To account for potential differences that funding structures could have on the availability of online courses, measures of state appropriations per full-time equivalent student in each year were gathered from a report produced by the State Higher Education Executive Officers Association (SHEEO) (Carlson et al. 2015). I also include a measure of the proportion of two-year open admissions public institutions within each state that I computed using data from IPEDS. These variables were intended to account for potential differences across states in the number of students who might be likely to attend open admissions institutions or attempt some online courses.

Because the likelihood of students enrolling in online courses (or an institution offering more online course sections) might be correlated with the average distance between the student and the postsecondary institution (Xu and Jaggars 2013), I incorporate a measure of the average distance a person would have to travel to get to the nearest open admissions institution. To compute this measure, I first found the distance from each census block group centroid to the nearest public open admissions postsecondary school. I then averaged these distances to the state level, using the relative population in each block group as the weight. Though a rough measure, it does give an indication of the spread of institutions around each state in terms of its population centers and is variable across the states. I include the log transformation of this measure alongside other state-level measures.

Finally, to account for potential differences between students living in rural and urban areas (Cejda 2007), I include both a self-constructed measure of population density for each school and an array of indicators for degree of urbanicity/rurality. I computed the first measure by summing the averages of census tract density (tract population divided by land area) to the county level and assigning each institution the value of its county. Indicator variables for degree of urbanicity come from the United States Department of Agriculture, which assigns all counties one of nine rural-urban continuum codes (United States Department of Agriculture Economic Research Service 2013). As with the population density, I assign each school the value of its county. Unlike the measures above, I incorporate both of these measures into the vector of first-level covariates in all models.

6 Results

To generate my results, I utilized a computationally-based Markov chain Monte Carlo algorithm that fit each Bayesian model a repeated number of times. While simple Bayesian models may be solved analytically, non-trivial equations are often too complex or have no closed-form solution (Gelman et al. 2014). To solve these problems, a computer program uses an iterative process to propose, compare, and either accept or reject parameter values. With enough iterations, parameters produced by the process will come from the true posterior distribution (Brooks et al. 2011). Though no tests exist that can determine whether enough samples have been drawn so that the true posterior is effectively summarized by their distribution, there are a number of best practices that support such a conclusion.

First, the algorithm is generally run multiple times to generate multiple chains of results. If these independent chains converge, that is, give distributions of results that are the same, this supports a conclusion that the draws summarize the posterior density distribution. For this analysis, I ran four independent chains with different starting values that appear to converge based on visual inspection of density plots and Rubin-Gelman statistics, which compare within-chain variance to between-chain variance, close to 1 (no significant difference) (Gelman et al. 2014).

Second, each chain should have a large number of draws. Because each chain starts with different values that are unlikely to come from the posterior, a large number of iterations is needed so that chains have time to reach the posterior. To prevent the initial and likely improbable starting values from biasing the results, it is common practice to discard some number of initial draws (Gelman et al. 2014). For all models, each of the four chains was run for 2,000 iterations and the first 1,000 of these values discarded. Combining these chains means that results for each model are a function of 4,000 draws.[12]

[12]All models were estimated using (CmdStan, version 2.18.0), the command line version of Stan's No-U-Turn Sampler (NUTS), a variant of the Hamiltonian Monte Carlo sampler that may more efficiently explore the parameter space. To reduce the amount of lagged auto-correlation between successive draws in the chains, improving convergence and the number of effective samples, all models were estimated using centered data and a QR reparameterization (Lunn et al. 2013; Stan Development Team 2018).

6.1 Single-level models

Table 3 shows the results for the single-level models. The Bayesian point estimates represent the mean of the posterior distribution with the accompanying numbers in the square brackets showing the 95% credible interval.[13] Models 1-3 use each broadband measure—download speed, upload speed, and the number of providers—in turn along with its quadratic. Model 4 uses all broadband measures and their quadratics in the same equation. All models include indicators for two-year institutions, availability of on-campus housing, year, and USDA urban/rural continuum code (not reported); the natural log of the institution's total of student enrollment as well as its proportions of students of color, women, Pell grant recipients, part-time attendees, and students 25 years and older; and the county-level measure population density, logged. In all models, the dependent variable is the natural log of the number of students who took some online courses. The parameter posterior distributions therefore represent the percent change in the number of sometimes online students for a one unit change in the covariate.

I first consider model (1), which uses download speed as the measure of broadband access. Because all right-hand-side variables were grand-mean centered, the intercept, α, may be interpreted as the expected log number of students who take some online courses for average institution.[14] At $\alpha = 6.726$, this translates to about 834 students (95% credible range: [814, 853]). Though this number is lower than that shown in Table 2 (1361 students), it accounts for different enrollment sizes across institutions as well as changes over time that skew the unconditional average. Indeed, the mean values for the year indicators, $\beta_{2013} = 0.103$ and $\beta_{2014} = 0.15$, show positive growth in the number of students taking online courses on the order of 11 to 16% over that seen in 2012. This aligns with general trends described elsewhere in the literature (Allen et al. 2016; Snyder, Brey, and Dillow 2016).

All else equal, two-year institutions are likely to have around 6% more ($\beta_{two-year} = 0.058$) of their students take some online courses than four-year institutions. Though all schools in this analysis utilize open admissions practices, this finding may reflect the two-year sector's particular focus on

[13] All reported 95% credible intervals represent the middle 95% of the posterior distribution, meaning that the lower and upper bound values are the 2.5% and 97.5% quantile values, respectively.

[14] Throughout the rest of the chapter, I will generally use the Bayesian point estimates when referring to the mean of parameter's posterior distribution. They should be understood, therefore, in their proper context as useful summaries of full probability distributions.

Table 3: Single level Bayesian regressions of log number of students who enrolled in some distance education courses on broadband measures.

	(1)	(2)	(3)	(4)
Download speed	0.333			0.41
	[-0.027,0.701]			[-0.007,0.82]
Download speed^2	-0.025			-0.027
	[-0.05,0]			[-0.056,0.002]
Upload speed		-0.077		-0.214
		[-0.272,0.114]		[-0.44,0.021]
Upload speed^2		0.003		0.016
		[-0.016,0.022]		[-0.007,0.039]
# Providers			0.007	0.029
			[-0.102,0.112]	[-0.084,0.144]
# Providers^2			-0.003	-0.004
			[-0.015,0.01]	[-0.017,0.009]
Two year institution	0.058	0.067	0.065	0.058
	[-0.016,0.131]	[-0.007,0.139]	[-0.009,0.14]	[-0.015,0.132]
Has on-campus housing	-0.049	-0.067	-0.048	-0.063
	[-0.12,0.024]	[-0.136,0.004]	[-0.12,0.024]	[-0.133,0.006]
log(Total enrollment)	1.154	1.156	1.155	1.157
	[1.117,1.191]	[1.119,1.193]	[1.119,1.189]	[1.12,1.195]
Prop. non-white	-0.633	-0.63	-0.635	-0.611
	[-0.755,-0.509]	[-0.746,-0.509]	[-0.75,-0.519]	[-0.73,-0.492]
Prop. women	-2.219	-2.201	-2.237	-2.203
	[-2.653,-1.799]	[-2.624,-1.784]	[-2.656,-1.82]	[-2.626,-1.765]
Prop. Pell grant	0.614	0.568	0.608	0.578
	[0.395,0.832]	[0.353,0.786]	[0.4,0.818]	[0.353,0.801]
Prop. part-time	-0.476	-0.495	-0.469	-0.488
	[-0.697,-0.247]	[-0.717,-0.273]	[-0.695,-0.241]	[-0.708,-0.266]
Prop. 25 years and older	0.397	0.405	0.408	0.413
	[0.155,0.638]	[0.158,0.646]	[0.156,0.662]	[0.171,0.665]
log(Pop. density)	-0.091	-0.084	-0.087	-0.087
	[-0.118,-0.064]	[-0.111,-0.058]	[-0.114,-0.06]	[-0.114,-0.06]
2013	0.103	0.11	0.1	0.105
	[0.045,0.159]	[0.054,0.165]	[0.045,0.158]	[0.047,0.164]
2014	0.15	0.164	0.142	0.159
	[0.088,0.211]	[0.102,0.225]	[0.08,0.204]	[0.095,0.222]
(Intercept)	6.726	6.726	6.726	6.726
	[6.703,6.749]	[6.702,6.75]	[6.701,6.751]	[6.702,6.749]
Unique institutions	1017	1017	1017	1017
N	2494	2494	2494	2494

Notes. Bayesian point estimates represent posterior mean values. Values in the square brackets are 95% credible intervals. Covariates not reported include indicators for USDA urban/rural community codes. Parameter distributions in each model are the combination of four independent MCMC chains of 1000 draws each (with 1000 initial draws discarded as burn-in) for a total of 4000 draws. All models were estimated using the Stan NUTS sampler with a normal likelihood sampling statement. The outcome measure in all models is the log number of students at each institution who enrolled in some distance education courses.

access (Cox 2006). Institutions that provide on-campus housing have 5% fewer students taking some classes online, which is unsurprising since online courses intended for distance education are less likely to appeal to students who have chosen to live on campus. Concerning characteristics of the student body, the results are mixed. On one hand, institutions with greater proportions of Pell grant recipients and students over 25 years old show increases in the percentage of students taking some online courses. These findings are in line with arguments of online courses as means to increase access (Bowen 2013; Selingo 2013). On the other hand, greater proportions of women, students of color, and those attending part time are associated with lower percentages of students taking some online courses, which runs counter to the same arguments. One potential explanation may be that the high correlation between some of these student body characteristics, many of which fall under the "non-traditional student" designation (Snyder, Brey, and Dillow 2016), produces counter-intuitive marginal results.

Two other institution-specific characteristics are of interest. First, schools located in areas with higher population density have fewer students who take some online courses. For a 10% increase in the population density surrounding an institution, 0.9% fewer students choose to enroll in some online courses ($\beta_{log(population density)} = -0.091$). Though the relationship is weak (keeping in mind that the degree of urbanicity is also accounted for in the model by the USDA urban continuum codes), it does support arguments that online courses may be particularly appropriate in less populated areas that are less likely to have strong transportation networks (Copps 2009). Second, the model shows that as the total number of enrolled students increases, proportionally greater numbers of them take online courses, with every 1% increase in the former associated with a 1.154% increase in the latter ($\beta_{log(enrollment)} = 1.154$). The positive elasticity of this relationship provides evidence that recent growth in enrollments may be due in part to increases in the number of students taking online courses, which, once prepared, benefit from economies of scale (Deming et al. 2015).

Turning to the parameter of interest in model (1), $\beta_{download}$, Table 3 shows a positive association between tiers of download speed and the number of students enrolling in some online courses that quickly diminishes as download speeds become faster. For a single-tier increase in speed at the lowest speed levels, tier 1 to tier 2, the marginal increase in students taking some online courses is 23.4%.[15] Going from tier 5 to tier 6, however, the increase is only 8.7%. At higher levels, the average marginal influence of download speed becomes negative: from tier 7 to tier 8 (just meeting the new threshold for broadband), the result is -1.2%. Wide credible intervals around the main parameter indicate a large variance in the posterior distribution. Though Bayesian analyses do not

[15]The margins in this section are computed using the full posterior distribution of the main and quadratic parameters. They are therefore slightly different than what is computed using only the Bayesian point estimates presented in the results tables.

utilize null hypothesis significance tests in the way that frequentist analyses do, it nonetheless may be useful to consider the 95% credible intervals as boundaries of interest. Because the lower bound of the credible interval on $\beta_{download}$ is negative, it suggests some uncertainty about the parameter's marginal influence on the outcome.[16] Yet since fully 85% of the sample draws are greater than zero at the marginal change from tier 6 to tier 7, $Pr(\frac{\Delta y}{\Delta x} > 0 = 0.854)$, a more properly Bayesian interpretation is that the model indicates that there is a greater than 85% probability that the marginal association between download speed and online course enrollment at this point in the download speed spectrum is positive. At lower download speed levels, the probability of a positive association is higher; at higher levels, tier 8 and above, the association is effectively zero and potentially negative.

Results for models (2) and (3), which operationalize broadband access using upload speed and the number of providers, respectively, are reported in the next two columns of Table 3. The posterior distributions of the covariate parameters in each remain much as they were in the first model. But unlike download speed, neither upload speed ($\beta_{upload} = -0.077$, $\beta_{upload^2} = 0.003$), nor provider count ($\beta_{provider} = 0.007$, $\beta_{provider^2} = -0.003$) are strongly predictive of online course enrollment numbers. Across each margin, neither parameter shows strongly positive or negative association with the number of students who take some classes online.

In the fully specified model, (4), which includes all broadband measures simulaneously, download speed once again is the most strongly predictive of the outcome with a diminishing relevance as speeds increase. As in model (3), the number of providers is comparatively uninformative. Though the parameters for upload speed remain less informative than those for download speed, their magnitudes have become larger, which is a counter-intuitive result. Why should improved upload speeds be associated with reductions in online course take-up? One possibility lies in the asymmetric nature of most residential broadband connections. Downloaded and uploaded data travel the same line, so ISPs must decide how to balance the load. Because most internet services are structured to send audio, video, text, and other files to the end-user who in turn needs only send small files of instructions regarding what to download, most residential broadband connections are asymmetric, meaning they allow download rates to be much higher than upload rates (Federal Communications

[16] If the mean value presented in the table represented a frequentist point estimate, one could not reject the null that $\beta = 0$ under a two-tailed test of significance at conventional levels of significance ($\alpha = 0.05$). This is not to say, however, that it would not be jointly significant with its quadratic term.

Commission 2014; Federal Communications Commission 2016a). If increased upload rates come at the expense of download rates and download rates are more salient for online students, then it is reasonable that the main parameter for upload speed skew negative.

Another reason for this counter-intuitive finding may lie in the technologies used to serve broadband. Though some people have access to broadband served through fiber optic lines, which use glass fibers and light to send digital signals, many users with wired connections have their broadband delivered through copper wires that also transmit their telephone or cable television signals (National Broadband Map). Broadband data transmitted over these lines more quickly degrades with distance, making it more difficult for ISPs to separate signal from noise as well as separate the data requests of users who share the line (reduce "crosstalk") (Grubesic and Murray 2002). Holding download speeds and the number of providers constant as in model (4), it may be that increases in upload speed allow for more traffic on wired lines that degrades the average signal quality for users along the line. The negative sign, therefore, could reflect a decrease in the proportion of students who attempt some online courses due to a less robust broadband connection, not necessarily faster upload speeds.

Regardless of the reason for the negative sign on the main upload parameter, however, upload speeds should be taken into account when computing the marginal change in the outcome for increases in download speed since increases in download speeds usually come with concommitant increases in upload speeds. In other words, holding upload speeds constant while increasing download speeds does not reflect the positive correlation between the two found in the analysis data set. Thus when computing the margins for download speed in equation (4), I include changes at the margins of upload speed. Figure 3 shows these changes across the full range of download speed tiers. Because upload speeds tend to be less than download speeds, the value I use for the change in upload speed at a specific level change in download speed are lower.[17] For example, at tier 2 download speed, the average upload speed is 1.37. These are shown on the x-axis of the plot. The black line shows the average marginal value and the shaded area the 95% credible interval.

With the inclusion of all broadband measures in the same model, the highest average marginal value at the bottom of the speed tiers is 19.6% and becomes zero at tier 8. Along this range, the

[17] These were computed from the data by taking the fractional value of the minimum upload speed over the minimum download speed, the same fractional value of the relative maximum speeds, and building a sequence of 11 values (1 to 11 on the NBM scale) that range between them.

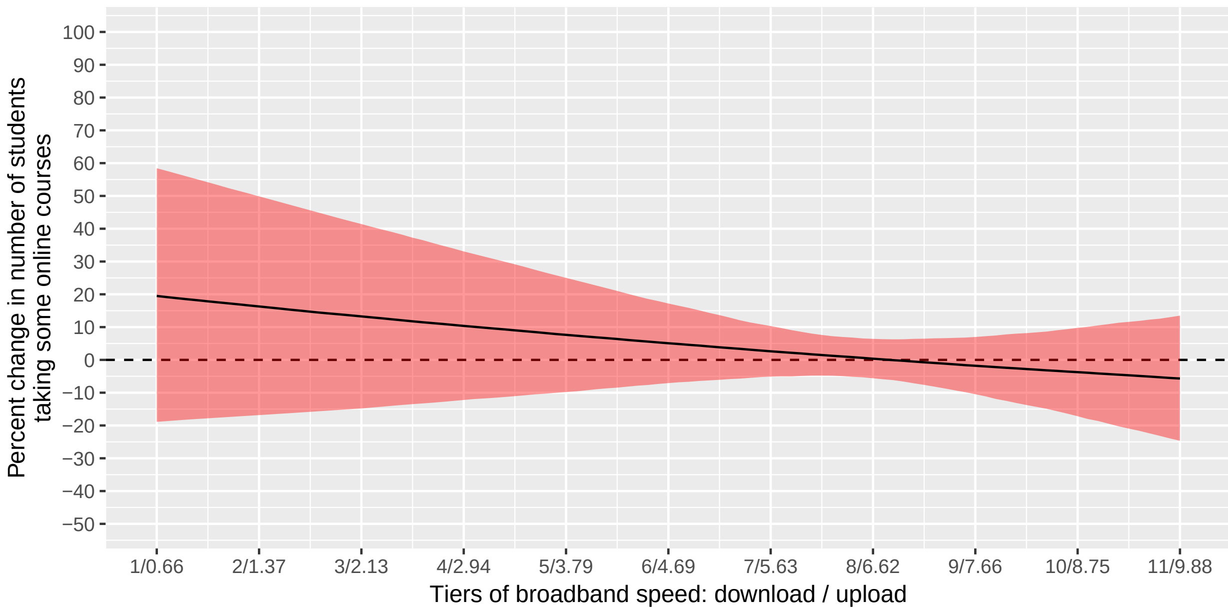

Figure 3: Computed marginal association between increasing broadband speed tiers and the number of students who take some courses online. Samples were taken from equation (4) shown in Table 3. Because increases in download speed in the sampel are accompanied by increases in upload speed, the margins were computed using scaled values of upload speed determined from the data. The black line represents the average marginal value; the shaded area represents the 95% credible intervals of the estimated marginal value.

probability of a positive association between download speed and some online enrollment shifts from 85.2% to 55.3%. Immediately above tier 8 download speeds, the average marginal association becomes negative with substantial proportions of values above and below zero, indicating little connection either way. These results provide some evidence that for open admissions institutions in areas with download speeds below the current threshold of broadband, increases in download speed are more likely than not to be connected to increases in the number students taking some of their courses online. For similar institutions in areas with already fast download speeds, there may be little gain or even decreasing numbers of these types of student.

6.2 Multilevel models

To better account for the nested nature of the data and differences in higher education policy contexts across states, I fit multilevel models that allow each state to have its own intercept (Gelman et al. 2014). As with the single-level linear regression models, the dependent variable is the log transformation of the number of students who attempt online courses. In addition to the covariates used in the single-level equations, I include a number of second-level covariates to help predict each state's unique intercept. These include the state average unemployment rate, statewide average appropriations per full time equivalent student, the proportion of open admissions public two-year institutions within the state, and the population-weighted average distance to the nearest public open admissions institution. In interest of space, Bayesian point estimates for second-level parameters and state-specific intercepts are not reported.

Table B.4 shows the posterior means for these models. Again, I fit four separate equations in which each broadband parameter was included on its own and alongside the others. Broadly, the results from the varying intercept multilevel models are similar to those found in the single level models. In models (1), (2), and (3), all of the broadband parameters of interest measures have 95% credible intervals that include zero. In the fully specified model (4), however, the Bayesian point estimate for download speed is once again positive ($\beta_{download} = 0.539$) and with a 95% credible interval that does not cover zero [0.137, 0.938].

As with the single level model using all measures of broadband, I again compute the margins for download speed including commensurate changes in upload speed. Figure 4 shows these changes across the full range of download speed tiers. While the trend is similar, the marginal change estimated using the varying intercept multilevel model is generally more positive. At the lowest broadband speeds, a single tier increase in download speed is associated with an average 33.2% increase in the number of students taking some courses online. At the new FCC threshold for broadband speed, tier 8, the average marginal increase is 2.3%. The probability of a positive change on the margin between tier 1 and 8 ranges from 95.9% to 77.5%. Only at average download speeds near tier 9 (approximately 8.7) does that probability become less than half. Accounting for state-level differences, results from the varying intercept models provide stronger evidence of a potentially small, but positive association between increases in download speed and the number of students at open admissions institutions that take some of their courses onlines, particularly at lower speed tiers.[18]

[18]While the focus of this paper has been on public institutions with open admissions policies, I also ran all models using the full census of colleges and universities during this period who reported having any students take some courses online as part of a sensitivity analysis. All models were the same, but with the inclusion of indicators for being private-non-profit, private-for-profit, or having an open admissions policy. Findings were qualitatively similar to those discussed in the paper. Tables showing these results are included in Appendix B.

Table 4: Varying intercept Bayesian regressions of log number of students who enrolled in some distance education courses on broadband measures

	(1)	(2)	(3)	(4)
Download speed	0.327			0.539
	[-0.027,0.674]			[0.137,0.938]
Download speed^2	-0.025			-0.039
	[-0.048,0]			[-0.067,-0.011]
Upload speed		-0.104		-0.211
		[-0.302,0.092]		[-0.438,0.027]
Upload speed^2		0.008		0.021
		[-0.012,0.027]		[-0.002,0.044]
# Providers			-0.045	-0.059
			[-0.156,0.066]	[-0.173,0.056]
# Providers^2			0.003	0.004
			[-0.01,0.016]	[-0.009,0.018]
Two year institution	0.088	0.099	0.099	0.092
	[0.003,0.172]	[0.012,0.183]	[0.014,0.183]	[0.006,0.178]
Has on-campus housing	-0.017	-0.026	-0.024	-0.025
	[-0.09,0.054]	[-0.1,0.05]	[-0.099,0.049]	[-0.099,0.049]
log(Total enrollment)	1.132	1.132	1.133	1.135
	[1.097,1.168]	[1.095,1.17]	[1.095,1.171]	[1.098,1.172]
Prop. non-white	-0.782	-0.777	-0.77	-0.785
	[-0.94,-0.624]	[-0.933,-0.625]	[-0.924,-0.615]	[-0.943,-0.626]
Prop. women	-2.105	-2.114	-2.117	-2.13
	[-2.502,-1.703]	[-2.517,-1.713]	[-2.538,-1.702]	[-2.546,-1.714]
Prop. Pell grant	0.467	0.442	0.434	0.449
	[0.225,0.716]	[0.188,0.689]	[0.189,0.675]	[0.208,0.695]
Prop. part-time	-0.849	-0.84	-0.857	-0.851
	[-1.111,-0.586]	[-1.097,-0.59]	[-1.121,-0.596]	[-1.111,-0.596]
Prop. 25 years and older	0.263	0.258	0.274	0.272
	[0.002,0.527]	[0.003,0.527]	[0.006,0.541]	[0.006,0.533]
log(Pop. density)	-0.065	-0.058	-0.059	-0.059
	[-0.095,-0.035]	[-0.088,-0.029]	[-0.088,-0.03]	[-0.091,-0.029]
2013	0.099	0.1	0.096	0.105
	[0.044,0.153]	[0.046,0.155]	[0.044,0.148]	[0.049,0.16]
2014	0.159	0.16	0.15	0.166
	[0.098,0.217]	[0.1,0.22]	[0.09,0.208]	[0.106,0.226]
Unique institutions	1017	1017	1017	1017
N	2494	2494	2494	2494

Notes. Bayesian point estimates represent posterior mean values. Values in the square brackets are 95% credible intervals. Intercepts (not reported) were allowed to vary at the state level. First level covariates not reported are indicators for USDA urban/rural community codes. Second level covariates include state unemployment rate, statewide average appropriations per FTE student, the proportion of public open admissions institutions in the state that are two-year institutions, and a population-weighted measure of the average distance to the nearest open admissions institution in the state. Parameter distributions in each model are the combination of four independent MCMC chains of 1000 draws each (with 1000 initial draws discarded as burn-in) for a total of 4000 draws. All models were estimated using the Stan NUTS sampler with a normal likelihood sampling statement. The outcome measure in all models is the log number of students at each institution who enrolled in some distance education courses.

7 Conclusion

This paper offers some evidence of a positive association between broadband access via download speeds and student enrollment in online courses at public colleges and universities with open ad-

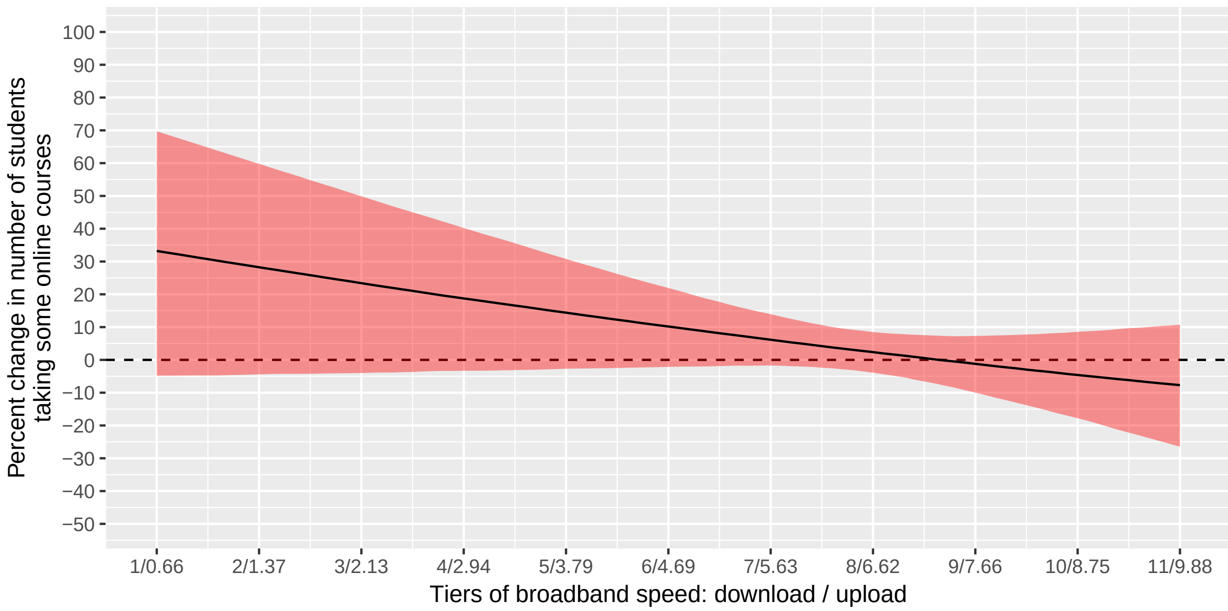

Figure 4: Computed marginal association between increasing broadband speed tiers and the number of students who take some courses online. Samples were taken from equation (4) shown in Table B.4. Because increases in download speed in the sampel are accompanied by increases in upload speed, the margins were computed using scaled values of upload speed determined from the data. The black line represents the average marginal value; the shaded area represents the 95% credible intervals of the estimated marginal value.

missions across the United States. In preferred model specifications, I find that average increases in download speeds in the area surrounding an institution correspond to increases in the number students who take some online courses may be as high as 33.2% at the lowest speeds and a more moderate 2.3% at the new threshold for broadband speed. At the highest speed levels, the average change on the margin may be negative, though credible intervals in this range are sufficiently spread across zero to suggest little change either way. While the range of the potential marginal values is wide—meaning that particular values should be interpreted with caution—the probability of a positive marginal association between download speed and sometimes online students is better than 50% between the lowest speed tiers and that which current defines broadband according to the FCC.

I do not find similar independent marginal associations for upload speed or the number of providers in the area. While I account for upload speed when computing the margins for increased download speed, theory would suggest that increases in the number of ISPs should foster competition that lowers prices and improves speeds, thus supporting a positive correlation between average provider counts and online course enrollment. That I find no such evidence may be due, in part, to the relatively small variation in this measure in the data set. It may also be the case that despite nominal access to the services offered by ISPs, geographic restrictions and costs may prevent some users from actually accessing those services (Grubesic and Murray 2002; Oyana 2011). The lack of clear association between the number of providers and the take-up of online courses may reflect this difference between the nominal number of ISPs, which I can observe in the data, and actual number of ISPs reasonably available to users, which I do not.

As a limitation, I note that the dependent variable I utilize does not measure course completion, only the number of students who enrolled in some online courses. Student persistence within the course is an important outcome to measure since students may make the rational decision to withdraw rather than persist and earn low grade or fail the course (Xu and Jaggars 2011; Xu and Jaggars 2013). Because the course enrollment decision typically comes in the months prior to the start of the course, it may be that some students only realize broadband access barriers after enrolling. As I note in section 3, students may not have a clear idea about the quality of their area broadband or, if they do, how it corresponds to the demands of an online course. Thus while my results speak to the possible association between broadband access and a student's decision of whether to enroll in online courses (or for institutions to offer them), they do not extend to student performance in online courses.

A final limitation of this paper is that I cannot differentiate between student demand for online courses and institutional supply. Increases in broadband speeds may induce students to demand more online course options. Conversely, postsecondary institutions situated in areas with better broadband access may choose to offer more online course options. Under this second scenario, students only respond to increases in broadband quality insofar as their schools do by offering more online courses (or comparatively fewer face-to-face courses). Findings, therefore, should not be interpreted causally, but instead as regression-adjusted measures of the association between broadband speed and online course enrollment. Future research that can link student-level data with similar measures of broadband speed would provide important new insights into how students respond to broadband access in terms of online coursework.

These potential limitations notwithstanding, this paper offers support for the conclusion that access to quality broadband, particularly in the form of faster download speeds for students who experience the lowest speed levels (Rosenboom and Blagg 2018), may be an important component of the choice to enroll in some online courses. Schools, however, may choose to offer online courses regardless of area broadband speeds based on their institutional needs and goals (Allen and Seaman 2011; Allen et al. 2016; Moore and Kearsley 2011). If students are effectively limited in their choice set to these courses, they may be forced to register for them regardless of their broadband connectivity. State-level higher education policies may also support online education to varying degrees based on financial and political rather than technological considerations (Johnstone 2006; Kinser 1999). Though these results speak only to enrollments and not completions, it is logical to think that improved broadband would also be associated with improved student outcomes in online courses.

A scenario of infrastructure-irrelevant policy-making, therefore, has important ramifications for the future of online higher education policy. As recent trajectories suggest, distance education delivered through online technology will almost certainly play an important part in the future of higher education (Allen et al. 2016; Bowen 2013; Selingo 2013; Snyder, Brey, and Dillow 2016). Insofar as it has the potential to be less expensive on the margins and reach more non-traditional students due to its flexibility, online education will be appealing to higher education institutions, especially those strong equity missions such as those with open admissions policies that serve large numbers of non-traditional students. But if these populations are also most at risk for lacking access to quality broadband (Cejda 2007; Hurst 2010), then the move to online coursework may have negative effects on their educational outcomes.

Should it be true that students respond positively to broadband speeds, public open admissions institutions that desire to increase the number of online learning opportunities may first wish to consider speeds in their service area. They may also find partners with those in other sectors who wish to improve broadband infrastructure for economic reasons. If, conditional on equitable access, students at these institutions realize equal or better educational outcomes, improving broadband infrastructure around public postsecondary institutions may pay triple dividends in the form of improved student outcomes, lower costs, and improved infrastructure that can be shared with all members of the community.

References

Allen, I Elaine and Jeff Seaman (2011). Going the Distance. Tech. rep. Babson Survey Research Group.

Allen, I Elaine et al. (2016). Online Report Card: Tracking Online Education in the United States. Tech. rep. Babson Survey Research Group.

Becker, Gary S (2009). Human capital: A Theoretical and Empirical Analysis, with Special Reference to Education. 3rd. University of Chicago Press.

Bettinger, Eric P. et al. (2017). "Virtual Classrooms: How Online College Courses Affect Student Success." In: American Economic Review 107.9, pp. 2855–75.

Bowen, William G (2013). Higher Education in the Digital Age. Princeton, NJ: Princeton University Press.

Bowen, William G et al. (2012). Interactive learning online at public universities: Evidence from randomized trials. Tech. rep. ITHAKA.

Brooks, Steve et al. (2011). Handbook of Markov Chain Monte Carlo. Handbooks of Modern Statistical Methods. Boca Raton, FL: Chapman and Hall/CRC.

Brown, Ronald H., David J. Barram, and Larry Irving (1995). Falling Through the Net: A Survey of the 'Have Nots' in Rural and Urban America. Tech. rep. Washington, D.C.: National Telecommunications and Information Administration, U.S. Department of Commerce.

Brown, Thomas et al. (2010). The Broadband Availability Gap. OBI Technical Paper 1. Federal Communications Commission.

Bureau of Labor Statistics (2012). Unemployment Rates for States, 2012 Annual Averages. online table.

Bureau of Labor Statistics (2013). Unemployment Rates for States, 2013 Annual Averages. online table.

Bureau of Labor Statistics (2014). Unemployment Rates for States, 2014 Annual Averages. online table.

Carlin, Bradley P. and Thomas A. Louis (2009). Bayesian Methods for Data Analysis. 3rd. CRC Press.

Carlson, Andrew et al. (2015). State Higher Education Finance FY 2014. Tech. rep. Boulder, CO: State Higher Education Executive Officers Association.

Cejda, Brent D (2007). "Distance Education in Rural Community Colleges." In: Community College Journal of Research and Practice 31.4, pp. 291–303.

Cohen, Arthur M. and Florence B Brawer (2003). The American Community College. 4th. John Wiley & Sons.

Copps, Michael J. (2009). Bringing Broadband to Rural America: Report on Rural Broadband Strategy. Federal Communications Commission.

Cotten, Shelia R and Shameeka M Jelenewicz (2006). "A disappearing digital divide among college students? Peeling away the layers of the digital divide." In: Social Science Computer Review 24.4, pp. 497–506.

Cox, Rebecca D. (2006). "Virtual Access." In: Defending the Community College Equity Agenda. Ed. by Thomas Bailey and Vanessa Smith Morest. Baltimore, MD: Johns Hopkins University Press. Chap. 5, pp. 110–131.

Crookston, Andrew and Gregory Hooks (2012). "Community Colleges, Budget Cuts, and Jobs: The Impact of Community Colleges on Employment Growth in Rural U.S. Counties, 1976-2004." In: Sociology of Education 85.4, pp. 350–372.

Czernich, Nina et al. (2011). "Broadband Infrastructure and Economic Growth." In: The Economic Journal 121.552, pp. 505–532.

Deming, David J et al. (2015). “Can Online Learning Bend the Higher Education Cost Curve?” In: American Economic Review 105.5, pp. 496–501.
Eagan, Kevin et al. (2014). The American Freshman: National Norms Fall 2014. Tech. rep. Los Angeles, CA: Higher Education Research Institute at UCLA.
Education, National Center for Education Statistics Department of. IPEDS Glossary. (Visited on 2015).
Education, National Center for Education Statistics Department of (2015). Integrated Postsecondary Education Data System (IPEDS). Data file.
Federal Communications Commission (2009). Connecting America: The National Broadband Plan. Federal Communications Commission.
Federal Communications Commission (2014). Measuring Broadband America Fixed Broadband Report: A Report on Consumer Fixed Broadband Performance in the U.S. Tech. rep. Washington, D.C.: Federal Communications Commission; Office of Engineering et al.
Federal Communications Commission (2015). 2015 Broadband Progress Report. FCC 14-113. Washington, D.C.: Federal Communications Commission.
Federal Communications Commission (2016a). 2016 Broadband Progress Report. FCC 16-6. Washington, D.C.: Federal Communications Commission.
Federal Communications Commission (2016b). Measuring Broadband America Fixed Broadband Report: A Report on Consumer Fixed Broadband Performance in the United States. Tech. rep. Washington, D.C.: Federal Communications Commission; Office of Engineering et al.
Figlio, David N., Mark Rush, and Lu Yin (2010). Is it Live or is it Internet? Experimental Estimates of the Effects of Online Instruction on Student Learning. Working Paper 16089. National Bureau of Economic Research.
Fishman, Rachel (2015). Deciding to go to college. 2015 College Decisions Survey: Part 1. Washington, D.C.: New America Education Policy.
Gelman, Andrew et al. (2014). Bayesian Data Analysis. 3rd. CRC Press.
Gill, Jeff and Lee D Walker (2005). “Elicited priors for Bayesian model specifications in political science research.” In: The Journal of Politics 67.3, pp. 841–872.
Goldrick-Rab, Sara (2010). “Challenges and Opportunities for Improving Community College Student Success.” In: Review of Educational Research 80.3, pp. 437–469.
Greene, William H. (2012). Econometric Analysis. 7th. Boston, MA: Prentice Hall.
Grubesic, Tony H. (2008a). “Spatial data constraints: Implications for measuring broadband.” In: Telecommunications Policy 32.7, pp. 490 –502.
Grubesic, Tony H. (2008b). “The spatial distribution of broadband providers in the United States: 1999–2004.” In: Telecommunications Policy 32.3–4, pp. 212 –233.
Grubesic, Tony H. (2008c). “Zip codes and spatial analysis: Problems and prospects.” In: Socio-Economic Planning Sciences 42.2, pp. 129 –149.
Grubesic, Tony H. (2012a). “The U.S. National Broadband Map: Data limitations and implications.” In: Telecommunications Policy 36.2, pp. 113 –126.
Grubesic, Tony H. (2012b). “The wireless abyss: Deconstructing the U.S. National Broadband Map.” In: Government Information Quarterly 29.4, pp. 532 –542.
Grubesic, Tony H. and Alan T. Murray (2002). “Constructing the divide: Spatial disparities in broadband access.” In: Papers in Regional Science 81.2, pp. 197–221.
Grubesic, Tony H. and Alan T. Murray (2004). “Waiting for Broadband: Local Competition and the Spatial Distribution of Advanced Telecommunication Services in the United States.” In: Growth and Change 35.2, pp. 139–165.

Hanigan, Ivan, Gillian Hall, and Keith BG Dear (2006). "A comparison of methods for calculating population exposure estimates of daily weather for health research." In: International Journal of Health Geographics 5.1, p. 38.

Hart, Cassandra M. D., Elizabeth Friedmann, and Michael Hill (2018). "Online Course-taking and Student Outcomes in California Community Colleges." In: Education Finance and Policy 13.1, pp. 42–71.

Hoffman, Donna L and Thomas P Novak (1998). "Bridging the Racial Divide on the Internet." In: Science.

Horrigan, John B. (2010). Broadband Adoption and Use in America. OBI Working Paper Series 1. FCC.

Huntington-Klein, Nick, James Cowan, and Dan Goldhaber (2017). "Selection into Online Community College Courses and Their Effects on Persistence." In: Research in Higher Education 58.3, pp. 244–269.

Hurst, Mark A. (2010). "Residential Broadband Access for Students at Walters State Community College." dissertation. East Tennessee State University.

Irving, Larry et al. (1999). Falling Through the Net: Defining the Digital Divide. Tech. rep. Washington, D.C.: National Telecommunications & Information Administration, U.S.Department of Commerce.

Jackman, Simon (2009). Bayesian Analysis for the Social Sciences. Wiley Series in Probability and Statistics. United Kingdom: John Wiley & Sons.

Jaggars, Shanna Smith (2014). "Choosing between online and face-to-face courses: Community college student voices." In: American Journal of Distance Education 28.1, pp. 27–38.

Johnson, Judith L. (2003). Distance Education: The Complete Guide to Design, Delivery, and Improvement. Teachers College Press.

Johnstone, Douglas (2006). The Rise of Western Governors University, 1996-2006. Western Governors University.

Jones, Steve et al. (2009). "U.S. college students' Internet use: Race, gender and digital divides." In: Journal of Computer-Mediated Communication 14.2, pp. 244–264.

Joyce, Theodore J et al. (2014). Does classroom time matter? A randomized field experiment of hybrid and traditional lecture formats in economics. Working Paper w20006. Cambridge, MA: National Bureau of Economic Research.

Keegan, Desmond (1996). Foundations of Distance Education. Third. New York, NY: Routledge.

Kinser, Kevin (1999). "The Origin, Development, and Implications of the Western Governors University." Doctoral dissertation. Columbia University.

Loss, Christopher P. (2012). Between Citizens and the State: The Politics of American Higher Education in the 20th Century. Princeton University Press.

Lunn, David et al. (2013). The BUGS book: A practical introduction to Bayesian analysis. CRC press.

Manski, Charles F. and David A. Wise (1983). College Choice in America. Cambridge, MA: Harvard University Press.

Mattern, Krista and Jeff N. Wyatt (2009). "Student Choice of College: How Far Do Students Go for an Education?" In: Journal of College Admission 203, pp. 18–29.

McConnaughey, James W. and Wendy Lader (1998). Falling through the Net: New Data on the Digital Divide. Tech. rep. Washington, D.C.: National Telecommunications & Information Administration, U.S.Department of Commerce.

Moloney, Jacqueline F. and Burks Oakley (2010). "Scaling Online Education: Increasing Access to Higher Education." In: Journal of Asynchronous Learning Networks 14.1, pp. 55–70.

Moore, Michael G. and Greg Kearsley (2011). Distance education: A Systems View of Online Learning. Cengage Learning.

National Telecommunications and Information Administration. National Broadband Map. (Visited on 2011).

Ortagus, Justin C. (2018). "National Evidence of the Impact of First-Year Online Enrollment on Postsecondary Students' Long-Term Academic Outcomes." In: Research in Higher Education.

Oyana, Tonny J. (2011). "Exploring geographic disparities in broadband access and use in rural southern Illinois: Who's being left behind?" In: Government Information Quarterly 28.2, pp. 252 –261.

Pappano, Laura (2012). "The Year of the MOOC." In: The New York Times.

Prieger, James E and Wei-Min Hu (2008). "The broadband digital divide and the nexus of race, competition, and quality." In: Information economics and Policy 20.2, pp. 150–167.

Radford, Alexandria Walton and Thomas Weko (2011). Learning at a Distance: Undergraduate Enrollment in Distance Education Courses and Degree Programs. NCES 2012-154. Washington, D.C.: U.S. Department of Education, National Center for Education Statistics, Institute of Education Sciences.

Rosenboom, Victoria and Kristin Blagg (2018). Disconnected from Higher Education. Tech. rep. Urban Institute.

Selingo, Jeffrey J. (2013). College (Un)Bound. New York: Houghton Mifflin Harcourt Publishing Company.

Shepard, Donald (1968). "A two-dimensional interpolation function for irregularly-spaced data." In: Proceedings of the 1968 23rd ACM national conference. ACM, pp. 517–524.

Snyder, Thomas D., Cristobal de Brey, and Sally A. Dillow (2016). Digest of Education Statistics: 2015. NCES 2016-014. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Snyder, Thomas D. and Sally A. Dillow (2015). Digest of Education Statistics: 2013. NCES 2015-011. Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Southern Region Educational Board (2013). State Data Exchange: 2011-2012 Indicators Report. Southern Regional Education Board.

Southern Region Educational Board (2015). State Data Exchange: 2013-2014 Indicators Report. Southern Regional Education Board.

Stan Development Team (2018). Stan Modeling Language Users Guide and Reference Manual. Version 2.18.0.

Tandberg, David A. (2010). "Politics, Interest Groups and State Funding of Public Higher Education." In: Research in Higher Education 51.5, pp. 416–450.

The Open University. The OU story. (Visited on 2015).

Thelin, John R. (2011). A History of American Higher Education. Second. Baltimore, MD: The Johns Hopkins University Press.

Turner, Sarah E. (2004). "Going to College and Finishing College: Explaining Different Educational Outcomes." In: College Choices: The Economics of Where to Go, When to Go, and How to Pay for It. Ed. by Caroline M. Hoxby. Chicago: University of Chicago Press, pp. 13–62.

United States Department of Agriculture Economic Research Service (2013). 2013 Rural-Urban Continuum Codes. online data set.

U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics (2012). National Postsecondary Student Aid Study, 2012 (NPSAS:12). online data set.

Western, Bruce and Simon Jackman (1994). "Bayesian Inference for Comparative Research." In: The American Political Science Review 88.2, pp. 412–423.

Xu, Di and Shanna Smith Jaggars (2011). "The Effectiveness of Distance Education Across Virginia's Community Colleges: Evidence From Introductory College-Level Math and English Courses." In: Educational Evaluation and Policy Analysis 33.3, pp. 360–377.

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." In: Economics of Education Review 37, pp. 46–57.

A Alternative specifications

In an alternative specification, I fit equations using the proportion of students who took some online courses as the outcome. To model this outcome, I use a beta likelihood function with dispersion parameter that accurately accounts for the [0,1] bounds of the dependent variable (Gelman et al. 2014):

$$\begin{aligned}
\frac{online}{total} &\sim beta(a, b) \\
a &= \mu \times \phi \\
b &= (1 - \mu) \times \phi \\
\mu &= \frac{exp(X\beta)}{1 + exp(X\beta)}
\end{aligned}$$

As with the primarily analysis models, I fit single level as well as multilevel models in which the intercepts and parameters on broadband measures are allowed to vary at the state level. Results from these models show the marginal effect of broadband on the logged-odds of a percentage increase in the number of students who took some online courses. The results from these models are qualitatively the same as those reported in the main text of the chapter.

Table A.1: Single level Bayesian beta regressions of percentage of students who enrolled in some distance education courses on broadband measures.

	(1)	(2)	(3)	(4)
Download speed	0.236			0.227
	[-0.119,0.598]			[-0.168,0.63]
Download speed^2	-0.017			-0.012
	[-0.043,0.007]			[-0.04,0.016]
Upload speed		-0.069		-0.202
		[-0.259,0.124]		[-0.416,0.014]
Upload speed^2		0.001		0.012
		[-0.018,0.02]		[-0.01,0.033]
# Providers			0.019	0.055
			[-0.084,0.124]	[-0.058,0.171]
# Providers^2			-0.004	-0.006
			[-0.017,0.008]	[-0.019,0.008]
Two year institution	0.086	0.095	0.091	0.085
	[0.014,0.16]	[0.023,0.168]	[0.017,0.164]	[0.01,0.163]
Has on-campus housing	-0.038	-0.054	-0.035	-0.054
	[-0.107,0.032]	[-0.126,0.015]	[-0.104,0.035]	[-0.127,0.017]
log(Total enrollment)	0.143	0.146	0.145	0.146
	[0.107,0.18]	[0.107,0.185]	[0.107,0.181]	[0.109,0.183]
Prop. non-white	-0.581	-0.579	-0.584	-0.542
	[-0.702,-0.459]	[-0.7,-0.46]	[-0.708,-0.465]	[-0.663,-0.418]
Prop. women	-2.377	-2.32	-2.395	-2.329
	[-2.833,-1.924]	[-2.768,-1.881]	[-2.851,-1.953]	[-2.785,-1.894]
Prop. Pell grant	0.659	0.619	0.656	0.602
	[0.454,0.869]	[0.409,0.829]	[0.448,0.865]	[0.386,0.825]
Prop. part-time	-0.492	-0.51	-0.484	-0.498
	[-0.72,-0.26]	[-0.74,-0.282]	[-0.706,-0.259]	[-0.724,-0.268]
Prop. 25 years and older	0.452	0.468	0.473	0.474
	[0.202,0.698]	[0.223,0.715]	[0.214,0.726]	[0.228,0.713]
log(Pop. density)	-0.076	-0.069	-0.073	-0.07
	[-0.102,-0.049]	[-0.096,-0.042]	[-0.1,-0.044]	[-0.096,-0.043]
2013	0.111	0.123	0.111	0.114
	[0.056,0.168]	[0.065,0.181]	[0.054,0.169]	[0.055,0.172]
2014	0.149	0.172	0.145	0.161
	[0.087,0.212]	[0.11,0.234]	[0.082,0.206]	[0.099,0.223]
(Intercept)	-1.433	-1.434	-1.434	-1.434
	[-1.458,-1.408]	[-1.46,-1.41]	[-1.459,-1.409]	[-1.459,-1.409]
Unique institutions	1017	1017	1017	1017
N	2494	2494	2494	2494

Notes. Bayesian point estimates represent posterior mean values. Values in the square brackets are 95% credible intervals. Covariates not reported include indicators for USDA urban/rural community codes. Parameter distributions in each model are the combination of four independent MCMC chains of 1000 draws each (with 1000 initial draws discarded as burn-in) for a total of 4000 draws. All models were estimated using the Stan NUTS sampler with a reparameterized beta likelihood sampling statement. The outcome measure in all models is the percentage of students at each institution who enrolled in some distance education courses.

Table A.2: Varying intercept Bayesian beta regressions of percentage of students who enrolled in some distance education courses on broadband measures

	(1)	(2)	(3)	(4)
Download speed	0.277			0.361
	[-0.046,0.606]			[-0.037,0.758]
Download speed^2	-0.021			-0.026
	[-0.044,0.001]			[-0.054,0.002]
Upload speed		-0.029		-0.109
		[-0.232,0.174]		[-0.34,0.114]
Upload speed^2		-0.001		0.009
		[-0.021,0.019]		[-0.014,0.032]
# Providers			-0.031	-0.044
			[-0.14,0.079]	[-0.156,0.07]
# Providers^2			0	0.002
			[-0.013,0.013]	[-0.012,0.015]
Two year institution	0.096	0.105	0.107	0.101
	[0.014,0.184]	[0.019,0.193]	[0.021,0.192]	[0.016,0.187]
Has on-campus housing	-0.006	-0.012	-0.01	-0.014
	[-0.079,0.064]	[-0.088,0.061]	[-0.084,0.063]	[-0.091,0.06]
log(Total enrollment)	0.114	0.115	0.116	0.118
	[0.077,0.152]	[0.078,0.152]	[0.078,0.154]	[0.08,0.156]
Prop. non-white	-0.709	-0.697	-0.689	-0.697
	[-0.864,-0.555]	[-0.848,-0.537]	[-0.846,-0.528]	[-0.856,-0.539]
Prop. women	-2.312	-2.303	-2.321	-2.312
	[-2.757,-1.886]	[-2.734,-1.868]	[-2.753,-1.897]	[-2.731,-1.891]
Prop. Pell grant	0.44	0.418	0.409	0.418
	[0.194,0.691]	[0.161,0.653]	[0.163,0.66]	[0.174,0.657]
Prop. part-time	-0.908	-0.897	-0.908	-0.913
	[-1.171,-0.643]	[-1.147,-0.643]	[-1.172,-0.654]	[-1.176,-0.656]
Prop. 25 years and older	0.385	0.382	0.404	0.403
	[0.114,0.654]	[0.106,0.642]	[0.146,0.666]	[0.145,0.66]
log(Pop. density)	-0.052	-0.046	-0.047	-0.046
	[-0.083,-0.023]	[-0.077,-0.016]	[-0.077,-0.017]	[-0.077,-0.015]
2013	0.107	0.11	0.105	0.116
	[0.052,0.16]	[0.057,0.164]	[0.051,0.159]	[0.062,0.171]
2014	0.161	0.168	0.153	0.174
	[0.103,0.217]	[0.107,0.228]	[0.096,0.211]	[0.114,0.236]
Unique institutions	1017	1017	1017	1017
N	2494	2494	2494	2494

Notes. Bayesian point estimates represent posterior mean values. Values in the square brackets are 95% credible intervals. Intercepts (not reported) were allowed to vary at the state level. First level covariates not reported are indicators for USDA urban/rural community codes. Second level covariates include state unemployment rate, statewide average appropriations per FTE student, the proportion of public open admissions institutions in the state that are two-year institutions, and a population-weighted measure of the average distance to the nearest open admissions institution in the state. Parameter distributions in each model are the combination of four independent MCMC chains of 1000 draws each (with 1000 initial draws discarded as burn-in) for a total of 4000 draws. All models were estimated using the Stan NUTS sampler with a reparameterized beta likelihood sampling statement. The outcome measure in all models is the percentage of students at each institution who enrolled in some distance education courses.

B Sensitivity analyses

Table B.1: Descriptive statistics of the institution sample using all institutions

	Mean/(SD)
Total enrollment	6395
	(8633)
Some online enrollment	1050
	(1679)
Two year institution	0.4
	(0.49)
Has on-campus housing	0.51
	(0.5)
Private, nonprofit	0.22
	(0.41)
Private, for-profit	0.17
	(0.38)
Open admissions policy	0.54
	(0.5)
Non-white enrollment	0.44
	(0.24)
Women enrollment	0.41
	(0.12)
Pell grant recipients	0.45
	(0.18)
Part-time enrollment	0.37
	0.25
Aged 25 years and older	0.36
	0.21
2013	0.44
	0.5
2014	0.27
	0.45
N (2012)	1816
N (2013)	2801
N (2014)	1732

Notes. Total enrollment and some online enrollment represent the average number of students rounded to nearest student. Other rows are proportions. Standard deviations are shown in parentheses. Schools included in the sample are public, open admissions postsecondary institutions that report at least one student who took some distance education courses.

Table B.2: Descriptive statistics of broadband measures using all institutions

	2012	2013	2014
Download tier	6.98	7.3	7.45
	(0.77)	(0.74)	(0.79)
Upload tier	4.54	4.84	5.03
	(0.83)	(0.93)	(0.96)
Number of providers	3.41	3.73	3.67
	(1.08)	(1.27)	(1.21)

Notes. Values are the average of broadband measures assigned across all schools in the sample in a given year. Each school is given a value that is the population-distance-weighted average of surrounding measures (at the census block level). Download and upload speeds are reported in ordered categorical tiers from 2 to 11. Broadband data come from the National Broadband Map. Standard deviations are shown in parentheses.

Table B.3: Single level Bayesian regressions of log number of students who enrolled in some distance education courses on broadband measures using all institutions

	(1)	(2)	(3)	(4)
Download speed	0.336			0.671
	[-0.066,0.729]			[0.22,1.137]
Download speed^2	-0.024			-0.045
	[-0.051,0.004]			[-0.078,-0.014]
Upload speed		-0.238		-0.41
		[-0.43,-0.05]		[-0.642,-0.189]
Upload speed^2		0.021		0.037
		[0.003,0.04]		[0.015,0.06]
# Providers			-0.009	0.019
			[-0.095,0.075]	[-0.072,0.11]
# Providers^2			0.002	-0.001
			[-0.007,0.011]	[-0.01,0.009]
Two year institution	-0.071	-0.062	-0.065	-0.068
	[-0.15,0.007]	[-0.14,0.018]	[-0.145,0.011]	[-0.146,0.012]
Has on-campus housing	0.106	0.096	0.104	0.098
	[0.027,0.185]	[0.014,0.18]	[0.025,0.184]	[0.016,0.181]
Private, nonprofit	-0.802	-0.796	-0.801	-0.8
	[-0.887,-0.716]	[-0.882,-0.711]	[-0.884,-0.716]	[-0.885,-0.717]
Private, for-profit	-0.655	-0.647	-0.658	-0.647
	[-0.775,-0.537]	[-0.767,-0.525]	[-0.78,-0.538]	[-0.767,-0.527]
Open admissions policy	-0.2	-0.203	-0.199	-0.201
	[-0.282,-0.118]	[-0.283,-0.123]	[-0.279,-0.119]	[-0.281,-0.121]
log(Total enrollment)	0.933	0.933	0.932	0.934
	[0.905,0.961]	[0.905,0.96]	[0.903,0.962]	[0.907,0.96]
Prop. non-white	-0.54	-0.544	-0.537	-0.531
	[-0.663,-0.418]	[-0.668,-0.423]	[-0.663,-0.41]	[-0.665,-0.401]
Prop. women	-0.363	-0.355	-0.364	-0.372
	[-0.593,-0.139]	[-0.577,-0.136]	[-0.591,-0.14]	[-0.606,-0.139]
Prop. Pell grant	1.366	1.339	1.359	1.335
	[1.172,1.553]	[1.143,1.536]	[1.167,1.559]	[1.138,1.534]
Prop. part-time	0.971	0.974	0.974	0.97
	[0.83,1.111]	[0.828,1.12]	[0.825,1.126]	[0.827,1.115]
Prop. 25 years and older	0.859	0.869	0.859	0.859
	[0.661,1.054]	[0.674,1.065]	[0.653,1.061]	[0.664,1.058]
log(Pop. density)	-0.022	-0.014	-0.02	-0.019
	[-0.047,0.002]	[-0.039,0.011]	[-0.044,0.005]	[-0.045,0.008]
2013	0.156	0.162	0.151	0.152
	[0.1,0.211]	[0.103,0.221]	[0.095,0.21]	[0.09,0.212]
2014	0.25	0.257	0.241	0.251
	[0.187,0.313]	[0.192,0.323]	[0.175,0.306]	[0.183,0.32]
(Intercept)	5.877	5.877	5.877	5.877
	[5.853,5.901]	[5.853,5.901]	[5.854,5.901]	[5.852,5.902]
Unique institutions	2975	2975	2975	2975
N	6349	6349	6349	6349

Notes. Bayesian point estimates represent posterior mean values. Values in the square brackets are 95% credible intervals. Covariates not reported include indicators for USDA urban/rural community codes. Parameter distributions in each model are the combination of four independent MCMC chains of 1000 draws each (with 1000 initial draws discarded as burn-in) for a total of 4000 draws. All models were estimated using the Stan NUTS sampler with a normal likelihood sampling statement. The outcome measure in all models is the log number of students at each institution who enrolled in some distance education courses.

Table B.4: Varying intercept Bayesian regressions of log number of students who enrolled in some distance education courses on broadband measures using all institutions

	(1)	(2)	(3)	(4)
Download speed	0.316			0.605
	[-0.071,0.696]			[0.152,1.065]
Download speed^2	-0.022			-0.041
	[-0.049,0.004]			[-0.073,-0.01]
Upload speed		-0.189		-0.325
		[-0.389,0.016]		[-0.567,-0.081]
Upload speed^2		0.018		0.031
		[-0.002,0.037]		[0.008,0.054]
# Providers			0.016	0.029
			[-0.079,0.109]	[-0.07,0.125]
# Providers^2			0.002	0.001
			[-0.007,0.012]	[-0.009,0.011]
Two year institution	-0.084	-0.077	-0.082	-0.084
	[-0.162,-0.006]	[-0.153,0]	[-0.162,-0.005]	[-0.165,-0.007]
Has on-campus housing	0.104	0.1	0.1	0.1
	[0.018,0.185]	[0.018,0.184]	[0.019,0.183]	[0.018,0.183]
Private, nonprofit	-0.806	-0.803	-0.806	-0.809
	[-0.896,-0.716]	[-0.89,-0.715]	[-0.894,-0.723]	[-0.897,-0.72]
Private, for-profit	-0.678	-0.676	-0.69	-0.686
	[-0.803,-0.555]	[-0.798,-0.552]	[-0.809,-0.565]	[-0.81,-0.562]
Open admissions policy	-0.207	-0.209	-0.206	-0.204
	[-0.286,-0.123]	[-0.292,-0.126]	[-0.286,-0.124]	[-0.288,-0.119]
log(Total enrollment)	0.925	0.924	0.924	0.924
	[0.895,0.954]	[0.895,0.953]	[0.896,0.952]	[0.894,0.954]
Prop. non-white	-0.73	-0.741	-0.739	-0.745
	[-0.879,-0.576]	[-0.889,-0.586]	[-0.89,-0.588]	[-0.897,-0.593]
Prop. women	-0.35	-0.341	-0.348	-0.363
	[-0.565,-0.134]	[-0.568,-0.119]	[-0.567,-0.127]	[-0.579,-0.145]
Prop. Pell grant	1.424	1.418	1.439	1.432
	[1.222,1.635]	[1.206,1.636]	[1.233,1.655]	[1.216,1.634]
Prop. part-time	0.976	0.981	0.977	0.972
	[0.819,1.127]	[0.834,1.133]	[0.835,1.122]	[0.82,1.123]
Prop. 25 years and older	0.81	0.821	0.798	0.799
	[0.614,1.012]	[0.612,1.027]	[0.601,0.995]	[0.597,0.996]
log(Pop. density)	0.001	0.007	-0.003	-0.003
	[-0.026,0.029]	[-0.021,0.035]	[-0.031,0.025]	[-0.032,0.026]
2013	0.151	0.153	0.14	0.14
	[0.092,0.211]	[0.093,0.213]	[0.08,0.201]	[0.081,0.199]
2014	0.251	0.25	0.236	0.242
	[0.184,0.317]	[0.186,0.314]	[0.172,0.301]	[0.175,0.309]
Unique institutions	2975	2975	2975	2975
N	6349	6349	6349	6349

Notes. Bayesian point estimates represent posterior mean values. Values in the square brackets are 95% credible intervals. Intercepts (not reported) were allowed to vary at the state level. First level covariates not reported are indicators for USDA urban/rural community codes. Second level covariates include state unemployment rate, statewide average appropriations per FTE student, the proportion of public open admissions institutions in the state that are two-year institutions, and a population-weighted measure of the average distance to the nearest open admissions institution in the state. Parameter distributions in each model are the combination of four independent MCMC chains of 1000 draws each (with 1000 initial draws discarded as burn-in) for a total of 4000 draws. All models were estimated using the Stan NUTS sampler with a normal likelihood sampling statement. The outcome measure in all models is the log number of students at each institution who enrolled in some distance education courses.

Table B.5: Single level Bayesian beta regressions of percentage of students who enrolled in some distance education courses on broadband measures using all institutions

	(1)	(2)	(3)	(4)
Download speed	0.223			0.424
	[-0.088,0.549]			[0.06,0.799]
Download speed^2	-0.017			-0.028
	[-0.039,0.005]			[-0.054,-0.003]
Upload speed		-0.198		-0.309
		[-0.357,-0.034]		[-0.496,-0.118]
Upload speed^2		0.016		0.026
		[0.001,0.032]		[0.008,0.044]
# Providers			-0.006	0.024
			[-0.079,0.068]	[-0.048,0.1]
# Providers^2			0.002	-0.001
			[-0.006,0.009]	[-0.009,0.007]
Two year institution	-0.052	-0.045	-0.048	-0.051
	[-0.114,0.01]	[-0.11,0.018]	[-0.109,0.01]	[-0.112,0.01]
Has on-campus housing	0.047	0.04	0.046	0.04
	[-0.023,0.116]	[-0.029,0.107]	[-0.019,0.111]	[-0.027,0.107]
Private, nonprofit	-0.588	-0.584	-0.588	-0.585
	[-0.661,-0.514]	[-0.656,-0.514]	[-0.659,-0.517]	[-0.659,-0.514]
Private, for-profit	-0.454	-0.446	-0.458	-0.446
	[-0.553,-0.354]	[-0.544,-0.348]	[-0.558,-0.359]	[-0.545,-0.343]
Open admissions policy	-0.194	-0.196	-0.193	-0.192
	[-0.261,-0.126]	[-0.264,-0.13]	[-0.26,-0.126]	[-0.258,-0.124]
log(Total enrollment)	-0.047	-0.047	-0.047	-0.047
	[-0.07,-0.024]	[-0.07,-0.024]	[-0.07,-0.024]	[-0.071,-0.024]
Prop. non-white	-0.446	-0.446	-0.44	-0.428
	[-0.552,-0.341]	[-0.551,-0.342]	[-0.541,-0.339]	[-0.535,-0.322]
Prop. women	-0.356	-0.349	-0.357	-0.364
	[-0.54,-0.165]	[-0.542,-0.156]	[-0.556,-0.172]	[-0.554,-0.178]
Prop. Pell grant	0.975	0.945	0.971	0.938
	[0.821,1.138]	[0.785,1.105]	[0.814,1.13]	[0.78,1.097]
Prop. part-time	0.708	0.708	0.713	0.711
	[0.598,0.82]	[0.594,0.821]	[0.596,0.83]	[0.594,0.83]
Prop. 25 years and older	0.627	0.638	0.623	0.619
	[0.464,0.79]	[0.486,0.793]	[0.463,0.783]	[0.453,0.785]
log(Pop. density)	-0.018	-0.01	-0.017	-0.015
	[-0.039,0.003]	[-0.03,0.011]	[-0.037,0.003]	[-0.036,0.007]
2013	0.135	0.139	0.127	0.131
	[0.087,0.183]	[0.09,0.189]	[0.077,0.176]	[0.083,0.182]
2014	0.206	0.214	0.195	0.209
	[0.151,0.261]	[0.158,0.269]	[0.14,0.247]	[0.151,0.265]
(Intercept)	-1.552	-1.552	-1.552	-1.553
	[-1.574,-1.531]	[-1.574,-1.53]	[-1.573,-1.531]	[-1.574,-1.531]
Unique institutions	2975	2975	2975	2975
N	6349	6349	6349	6349

Notes. Bayesian point estimates represent posterior mean values. Values in the square brackets are 95% credible intervals. Covariates not reported include indicators for USDA urban/rural community codes. Parameter distributions in each model are the combination of four independent MCMC chains of 1000 draws each (with 1000 initial draws discarded as burn-in) for a total of 4000 draws. All models were estimated using the Stan NUTS sampler with a reparameterized beta likelihood sampling statement. The outcome measure in all models is the percentage of students at each institution who enrolled in some distance education courses.

Table B.6: Varying intercept Bayesian beta regressions of percentage of students who enrolled in some distance education courses on broadband measures using all institutions

	(1)	(2)	(3)	(4)
Download speed	0.212			0.378
	[-0.103,0.544]			[0.011,0.75]
Download speed^2	-0.016			-0.026
	[-0.039,0.006]			[-0.052,-0.001]
Upload speed		-0.155		-0.24
		[-0.32,0.01]		[-0.431,-0.043]
Upload speed^2		0.014		0.022
		[-0.002,0.03]		[0.002,0.04]
# Providers			0.026	0.038
			[-0.049,0.101]	[-0.039,0.113]
# Providers^2			0.001	0
			[-0.007,0.009]	[-0.008,0.008]
Two year institution	-0.063	-0.059	-0.064	-0.065
	[-0.127,0.001]	[-0.122,0.005]	[-0.128,0]	[-0.128,-0.003]
Has on-campus housing	0.04	0.036	0.038	0.036
	[-0.029,0.108]	[-0.034,0.104]	[-0.027,0.104]	[-0.029,0.1]
Private, nonprofit	-0.614	-0.61	-0.614	-0.613
	[-0.686,-0.539]	[-0.683,-0.537]	[-0.689,-0.54]	[-0.687,-0.541]
Private, for-profit	-0.485	-0.482	-0.496	-0.488
	[-0.585,-0.384]	[-0.585,-0.379]	[-0.597,-0.391]	[-0.591,-0.386]
Open admissions policy	-0.205	-0.206	-0.205	-0.203
	[-0.273,-0.137]	[-0.275,-0.138]	[-0.272,-0.137]	[-0.27,-0.137]
log(Total enrollment)	-0.059	-0.059	-0.058	-0.058
	[-0.082,-0.036]	[-0.083,-0.035]	[-0.083,-0.034]	[-0.081,-0.035]
Prop. non-white	-0.62	-0.623	-0.621	-0.626
	[-0.748,-0.495]	[-0.749,-0.498]	[-0.749,-0.499]	[-0.756,-0.497]
Prop. women	-0.374	-0.366	-0.378	-0.387
	[-0.577,-0.173]	[-0.564,-0.174]	[-0.569,-0.182]	[-0.591,-0.186]
Prop. Pell grant	1.017	1.009	1.028	1.019
	[0.84,1.196]	[0.831,1.183]	[0.858,1.198]	[0.842,1.196]
Prop. part-time	0.723	0.727	0.726	0.722
	[0.606,0.843]	[0.611,0.841]	[0.614,0.836]	[0.607,0.84]
Prop. 25 years and older	0.561	0.568	0.552	0.55
	[0.4,0.722]	[0.401,0.735]	[0.392,0.716]	[0.39,0.714]
log(Pop. density)	0.007	0.013	0.002	0.004
	[-0.016,0.031]	[-0.012,0.035]	[-0.022,0.026]	[-0.02,0.028]
2013	0.131	0.131	0.117	0.121
	[0.082,0.179]	[0.083,0.18]	[0.071,0.164]	[0.069,0.172]
2014	0.211	0.21	0.193	0.205
	[0.154,0.268]	[0.157,0.262]	[0.139,0.247]	[0.147,0.262]
Unique institutions	2975	2975	2975	2975
N	6349	6349	6349	6349

Notes. Bayesian point estimates represent posterior mean values. Values in the square brackets are 95% credible intervals. Intercepts (not reported) were allowed to vary at the state level. First level covariates not reported are indicators for USDA urban/rural community codes. Second level covariates include state unemployment rate, statewide average appropriations per FTE student, the proportion of public open admissions institutions in the state that are two-year institutions, and a population-weighted measure of the average distance to the nearest open admissions institution in the state. Parameter distributions in each model are the combination of four independent MCMC chains of 1000 draws each (with 1000 initial draws discarded as burn-in) for a total of 4000 draws. All models were estimated using the Stan NUTS sampler with a reparameterized beta likelihood sampling statement. The outcome measure in all models is the percentage of students at each institution who enrolled in some distance education courses.