Every Little Bit Counts: The Impact of High-Speed Internet on the Transition to College

Lisa J. Dettling[1] - Federal Reserve Board of Governors
Sarena Goodman - Federal Reserve Board of Governors
Jonathan Smith - College Board and Georgia State University

November 2016

Abstract

This paper examines whether high-speed Internet affects students' college applications. Our analysis links the diffusion of residential broadband to the testing and application outcomes of millions of PSAT- and SAT-takers and finds that students with broadband in their zip code perform better on the SAT and apply to a higher number and more expansive set of colleges. Extended results reveal that effects are concentrated among higher-SES students, indicating that while, on average, the availability of broadband improves applications to college, it may also increase pre-existing inequities by primarily benefiting those with more resources.

[1] Email: lisa.j.dettling@frb.gov, sarena.f.goodman@frb.gov, and jsmith@collegeboard.org

The analysis and conclusions set forth are those of the authors and do not indicate concurrence by other members of the Board of Governors, the College Board, or the research staff at either institution. We thank Josh Goodman, Hilary Hoynes, Adam Isen, Melissa Kearney, Michael Lovenheim, Amanda Pallais, Bekah Richards, and seminar participants at the Federal Reserve Board Applied Micro Lunch Seminar, the DC Economics of Education Working Group, the 2016 Federal Reserve System Applied Micro Conference, the 2016 Annual Meeting of the Society of Labor Economists, and the 2015 Association for Education Finance and Policy and Society for Research and Educational Effectiveness conferences for helpful comments and suggestions.

I. Introduction

College planning is complicated: there are thousands of colleges in the United States, each with countless attributes, and students face uncertainty in both admission and completion. Moreover, students encounter hurdles throughout the lengthy college planning and application process. These hurdles can be procedural, such as requirements that students register for exams by certain deadlines, or informational, such as the wealth of college characteristics that students are expected to obtain and distill to help produce a good fit. These complications, along with the many steps involved in college applications (Klasik, 2012), have contributed to widespread inequality in college access (e.g., Bowen, Chingos, and McPherson, 2009; Hoxby and Avery, 2013; Smith et al, 2013; Pallais and Turner, 2006; Spies, 2001; Ellwood and Kane, 2000).

While college access efforts, programs, and organizations are designed to help students overcome the optimization problem they face, students often do not have full or correct information to make this consequential decision (e.g., Dillon and Smith 2013; Hoxby and Turner, 2015; Betts, 1996; Altonji, 1993). A natural question follows - how can we encourage students to apply to and attend colleges that best fit their needs? Prior successful interventions have aimed to increase information provision (Avery and Kane, 2004; Carrell and Sacerdote, 2013) and change the way that information is delivered (Bettinger et al., 2012; Hoxby and Turner, 2013 and 2015). Related work has found that student decisions are quite sensitive to small changes in information or costs.[2] In theory, a technology that increased the availability or improved the presentation of information could also generate large changes in college-going.

[2] Evidence of students' responsiveness to small changes in information or costs include interventions that change: 1) rules of thumb (Pallais, 2015), 2) the salience of college rankings (Luca and Smith, 2013), 3) financial aid offers (Cohodes and Goodman, 2014), 4) application fees and essays (Smith, Hurwitz, and Howell, 2015), and 5) admissions exam taking (Bulman, 2015; Goodman, 2016; Klasik, 2014; Hurwitz et al., 2015).

Between 2000 and 2013, the landscape of information technology was profoundly altered, as broadband Internet rapidly diffused across the United States. Over that time, reported household usage rates surged from 3 to 70 percent.[3] Recent research has linked broadband technology to changes in a host of outcomes, including academic achievement (Vigdor et al., 2014; Faber, Sanchis-Guarner, and Weinhardt, 2015), labor force participation (Dettling, forthcoming), voter turnout (Falck, Gold and Heblich, 2014), and crime (Bhuller et al., 2011). The estimates from each study reflect how well broadband improved upon existing technologies, with the ease and speed of information acquisition that it enabled often emphasized as an important mechanism.

In this paper, we tie the two strands of literature together and examine whether the dramatic and conditionally-random national diffusion of broadband – which could similarly make it easier and less costly to obtain information about the college application process – affected students' applications. Our empirical strategy links the universe of administrative test-taker data from the College Board to a new zip code-level measure of residential broadband Internet availability.[4] This measure, which we derive from data published by the Federal Communications Commission (FCC) on the number of broadband Internet Service Providers (ISPs) in each zip code between 1999 and 2007, is the first that reflects aggregate trends in survey-reported usage over the time that the prevalence of broadband was skyrocketing. Specifically, our measure outperforms other potentially viable measures in matching national usage rates (Figure 1) and, of particular relevance to our research question, predicting teen usage (Appendix Table 1).[5]

[3] High-speed Internet usage rates were obtained from PEW Research at www.pewinternet.org/data-trend/internet-use/connection-type/.
[4] We focus on the effects of *residential broadband* availability. Other forms of Internet access (e.g., dial-up, school-based) could also affect the outcomes we consider; thus, our estimates may understate the full role of Internet technology in college applications.
[5] We also produce estimates of the "effects of broadband availability" under these other measures. Results become more similar to our main estimates the better the measure fits usage trends and the more predictive the measure is of teen usage. Measures that perform poorly on these two metrics tend to be statistically insignificant.

Our analysis compares outcomes across millions of students with Preliminary SAT/National Merit Scholarship Qualifying Test® (PSAT) and SAT records, leveraging variation in whether broadband was available in a student's zip code during her junior year. We estimate the effects of broadband availability on a student's SAT score (used by colleges in admissions) and characteristics of her application set derived from Score Sends, including whether she applied to: a four-year college, a larger number of colleges, a selective college, an academically matched college, a top liberal arts college, her state's flagship university, and an out-of-state college.[6]

Our findings indicate that the availability of broadband Internet in a student's zip code unambiguously improves her SAT score and her application set. On average, her test score improves by 0.7 scale point and her application portfolio increases in both size and quality, with the magnitude of these changes ranging from 0.2 to 0.4 percentage point, depending on the outcome we consider. Some results are particularly striking when we scale by mean application rates; for example, 7.2 percent of the SAT-taking population apply to a top liberal arts college, but we find that high-speed Internet availability increased that rate by almost 0.2 percentage point (or 3 percent). Though a student's performance on the SAT and application set are undoubtedly intertwined, these improvements in applications appear to be largely independent from the increase in scores that we estimate. Further, when we evaluate the mechanisms that could explain our full set of results, it seems that the availability of broadband improves applications primarily by reducing the direct effort and informational costs involved.

Our identification strategy requires that the availability of broadband is exogenous to student applications, which we probe in several ways. First, because applications are only observed if students take the SAT, we confirm that broadband availability does not correlate with test-taking,

[6] Score Sends are a measure of student interest that are frequently used as proxies for applications (e.g., Card and Krueger, 2005; Pallais, 2015).

student characteristics (e.g., race, family background, and PSAT scores), or the availability of a competing exam (the ACT). Second, we provide suggestive evidence that broadband availability is not picking up excluded variables or trends that are broadly driving changes in applications – namely, we show that broadband availability does not correlate with zip code income and that applications trended similarly in early- and late-adopting zip codes in the late 1990s (before broadband was available). Third, we examine broadband availability within an event-study framework, and the resulting estimates are generally indistinguishable from zero in years prior to availability, jump in the year broadband becomes available, and are positive thereafter.[7]

Even as broadband diffused across the United States, not every type of student necessarily benefited equally. On one hand, broadband take-up rates, computer accessibility, and Internet use vary by socioeconomic status (Appendix Figure 1). On the other, teens from disadvantaged backgrounds may be the least informed of their postsecondary options and, hence, may stand to gain most from a technology that reduces information costs. When we examine responsiveness by group, effects are concentrated among students in urban areas and in high-income areas, white students, and students with more-educated parents, likely reflecting differences in how low- and high-SES households engage with the Internet. Thus, while, on average, broadband appears to improve applications, it may also widen existing inequities, favoring those with more resources.

The contribution of our paper is fourfold. First, we provide the first causal estimates of the effects of high-speed Internet availability on college admissions testing and applications.[8] Second, we add to a literature that finds that small cost reductions and changes in information

[7] We also probe effects of senior year broadband availability as well as within PSAT subgroups and show that estimates conform to theoretical predictions – namely, that applications (but not SAT scores) improve if availability is measured in senior year and that students in lower PSAT score groups drive our main estimates.
[8] A related literature generally does not find positive effects of computer and Internet technology on academic achievement (e.g., Faber, Sanchis-Guarner, and Weinhardt, 2015; Vigdor, Ladd, and Martinez, 2014; Belo, Ferreira, and Telang, 2013; Fairlie and Robinson, 2013).

can improve postsecondary outcomes. We demonstrate that significantly many students, if given the opportunity, appear able to obtain and distill information on colleges and universities online. We also show that some students are potentially being left behind by the information age because unobserved barriers limit access and/or they require more guidance on how to benefit from access. Third, given that universal broadband Internet is a central policy goal, our finding that broadband Internet availability can change college-going outcomes, but that the benefits may unequally accrue to higher-resource students, suggests a more intensive intervention than deployment alone may be necessary to realize the full benefits of increased availability in underserved areas. Last, we make a methodological contribution by developing a way to measure broadband Internet availability at the local level that matches aggregate usage patterns.

II. Conceptual Framework

Despite the existence of substantial returns to college quality (e.g., Card, 1995; Black and Smith, 2006; and more recently, Zimmerman, 2014), particularly among disadvantaged students (Dale and Krueger, 2002, 2011), three key findings have emerged from the economic literature on college choice that suggest that student applications might benefit from an advancement in information technology. First, many students, especially disadvantaged students, do not apply to or attend a college commensurate with their abilities. Second, information constraints appear to play a sizable role in this “under-match.” Third, students exhibit large responses to defaults, nudges, and small changes in costs, even those associated with effort or time, suggesting that they are broadly not well informed of optimal strategies in the application process.[9]

[9] Notably, Hoxby and Turner’s 2013 experiment produced large enrollment effects from targeted mailings among college admissions test-takers, and their corresponding survey revealed that the untreated students were dramatically less informed about college quality – particularly at top-ranked liberal arts schools and out-of-state schools.

In theory, the diffusion of broadband could operate similarly to prior successful interventions in this area, by reducing frictions for high school students as they transition to college.[10] Prospective applicants could use the Internet to quickly and easily conduct tailored searches for information that once required conversations with high school counselors, parents, and peers, who may be differentially knowledgeable of the college-going landscape, students' needs and interests, or more practical components of the application process (e.g., exam registration and preparation, soliciting and submitting applications to schools or for financial aid and scholarships). College rankings, costs, and other characteristics, admissions strategies, requirements, and deadlines, study guides, practice exams, application forms, and information on financial aid and scholarships are all available online. Thus, by making useful information more accessible, broadband could reduce the time, effort, or monetary costs of applying to college.

However, compared with prior interventions, which were targeted toward particular students or dispensed particular types of information, the nature of the relationship between broadband availability (upon which our analysis is based) and applications is more complex. First, the relationship between availability and use is potentially mediated by several additional factors, including household take-up, connection speeds, and access to a computer. Second, the relationship between use and useful is not even necessarily positive and results in ambiguous predictions. In this respect, our setting is quite similar to Gentzkow and Shapiro (2008), who argued that the value-add of preschool television depended critically on how useful it was relative to the alternative activities that it crowded out. If students with broadband successfully access and distill useful information, applications would improve. If students do not use the

[10] Broadband is likely an improvement over the previous technology, dial-up Internet, as it offers much faster search speeds and consequently enables students to more easily access and download large amounts of content. It is also "always on," eliminating the time cost to connect. There is growing evidence that broadband affects outcomes despite the existence of dial-up (e.g., Falck, Gold and Heblich, 2014; Bhuller et al., 2011; Dettling, forthcoming).

Internet in this manner (or do so unsuccessfully), we would see no effect on student outcomes. If students primarily use the Internet for leisure activities, and it serves as a distraction leading them to substitute time towards leisure, or they use the Internet instead of more productive forms of information acquisition that already exist, we would see worse outcomes. Indeed, survey evidence is somewhat mixed on these topics, though college search does appear to be a popular use of the Internet.[11] Ultimately, whether broadband availability, on average, generates positive, negative, or no effects on applications is an empirical question we strive to answer through the paper. In Section V, we return to these theories and explore the underlying mechanisms at work.

Driven in part by these nuances, we also investigate heterogeneity by demographics and geography. Specifically, we examine whether minority students, students in low-income or more-remote areas, or students with less-educated parents respond to broadband availability differently, under two broad hypotheses. First, per the under-match literature, effects might be concentrated among students with the least resources, who have the fewest inroads to elite colleges. Such students may have the least access to alternative technologies that can assist in their college search. Second, in direct contrast, effects might be concentrated among students with the *greatest* resources, who, despite likely having the most inroads to elite colleges without broadband, also may have higher broadband take-up and more productive use of the Internet. Per the former, there are large gaps in home computer and broadband access by socioeconomic

[11] In 2005, 45 percent of Internet users reported searching for information on prospective college or universities for themselves or a family member (PEW 2005). (This rate is similar to reported use rates for banking online (44 percent) and job search (44 percent), and higher than reported rates for reading a blog (27 percent), playing online games (32 percent), looking at online classifieds (36 percent), or using social networking sites (7 percent).) In 2012, 75 percent of teachers reported that the Internet and digital search tools had a mostly positive impact on students' research habits; however, 64 percent reported that digital technologies do more to distract students than help them academically (Purcell et al., 2012). Finally, according to Google Correlate data spanning 2004 to 2016, the top 10 correlates with a search for the word "SAT" include: "SAT test," "SAT admission," "SAT practice," and "taking the SAT," speaking to the many ways in which prospective applicants – or potentially SAT-takers (who are the focus of our study) – use the Internet to aid in their search for college.

status, within our sample period and in 2009—by which point local broadband had become ubiquitous—indicative of phenomena at work beyond questions of availability (Appendix Figure 1). Per the latter, Vigdor, Ladd, and Martinez (2014) found that, for younger students in North Carolina, Internet access increased the achievement wedge between low- and high-income students, which they attributed to a digital divide in how productively the Internet was being used by these groups at home. If these divides prevail in our setting, the mere availability of broadband may not benefit traditionally underserved populations.[12]

III. Data

The empirical strategy relates zip code-level broadband Internet availability to individual-level testing and application data for students who took SAT exams. This section describes our main data sources and how we construct our relevant variables.

a. Testing and Applications

Our primary data source is the College Board (CB), an organization that administers the Preliminary SAT/National Merit Scholarship Qualifying Test® (PSAT) and the SAT to high school students. The PSAT is an assessment taken prior to the SAT that serves as a qualifying exam for a nationally competitive scholarship program. Approximately 3.5 million students take the PSAT each year, typically either in the fall of their sophomore or junior year or both. The SAT is primarily a college admissions exam, and it yields a key metric on admissibility as well as application and demographic information for a majority of college-bound students. Over 1.5 million students in the graduating high school class of 2014 took the SAT, typically as juniors or seniors and often as both. The SAT contains math and critical reading sections, with each scored

[12] This prediction is not entirely inconsistent with the under-match literature, as many successful interventions were conducted after broadband had become commonplace.

on a scale between 200 and 800, for a maximum composite score of 1600.[13] The PSAT has a similar format, but the section scoring is between 20 and 80.

CB data also contain self-reported information on high school graduation cohort, student race, gender, cumulative high school GPA, home zip code, parental income and education, and high school. Along with exam scores and demographics, the CB data identify colleges to which students send official copies of their SAT scores (Score Sends), which serve as good proxies for college applications (Card and Krueger, 2005; Pallais, 2015). When registering for the SAT, the student has the option to send her scores to four colleges for no additional cost. Scores may also be sent at a later date for a fee of approximately $11 per Score Send.[14] For every Score Send, we merge characteristics of each college, including its quality (average freshman SAT score), control (public or private), level (two- or four-year), type (e.g., liberal arts, state public flagship), and location (state), using the Integrated Postsecondary Education Data System (IPEDS).

Our final sample is composed of the 7,452,302 students in the graduating high school cohorts of 2001 to 2008 who took both the PSAT and SAT.[15] We select these cohorts in order to examine broadband availability in both junior and senior year of high school. (Consistent measurement of broadband service is available from December 1999 to December 2007.) While the outcomes we study can be derived exclusively from SAT information, there are two main benefits to imposing the additional restriction that students take the PSAT. First, the analysis will rely on comparisons in outcomes that can only be observed if students take the SAT, which might not be valid if there are responses along the extensive margin (i.e., election of the SAT).

[13] A writing section was introduced in 2005. For consistency across cohorts, and because colleges typically do not rely on the writing section, we focus on the composite score formed from the math and critical reading sections.
[14] The cost changed slightly over the sample period.
[15] The sample includes only students with a valid zip code. A few students were excluded because all of their demographic data were missing or because they live in a zip code with no information on broadband availability.

The additional restriction will allow us to test for the presence of such responses using information available from the PSAT.[16] Second, the PSAT is generally taken before students are "treated" in our setting, and the resulting score is not a direct input into college applications; thus, the PSAT score offers a measure of a student's ability (or admissibility) that is unlikely to be influenced by treatment and can be included as a control in the main analysis.

Table 1 describes our sample. The average combined PSAT score on both sections is just over 99 and, similarly, the average SAT score, which combines the results of these same testing areas, is just over 1000. About 79 percent of our sample submitted an application to at least one four-year college. In addition, 40 percent applied to at least five colleges, which, importantly, is one more than the number of free Score Sends.[17] Approximately 50 percent applied to a selective college (average SAT score greater than 1200), though almost 70 percent of students applied to an academically matched college (average SAT score at least as good as their own or 1300, whichever is lower). About 30 percent of test-takers applied to the state flagship and 50 percent to an out-of-state college; notably, less than 10 percent applied to a top liberal arts college. Demographically, the sample is 67.2 percent white, 10.7 percent black, and 9.5 percent Latino/Hispanic, and is 45 percent male. High school GPA is derived from a categorical variable where 0 is a non-response, 1 is the lowest, and 12 is the highest. We use a set of GPA dummy variables in the analyses, but the table reports the average of the continuous version (3.363).

Finally, we add zip code-level economic characteristics to our data in order to control for changes in local economic conditions. Mean adjusted gross income is $77,444, which is measured at the zip code-level and was obtained from the IRS Statistics of Income (SOI) data.[18]

[16] We will also test for responses along the PSAT-taking margin.
[17] The modal number of Score Sends is four and Pallais (2015) shows that students tend to apply to the number of free Score Sends. Therefore, sending at least five Score Sends is a deliberate act.
[18] We interpolate missing years in this data.

Population and housing data at the zip code-level were obtained from the 2000 Census and made time-variant by merging with zip code-level trends in SOI counts of filers and households. We capture local labor market trends by including information on county-level unemployment rates, collected from the Bureau of Labor Statistics. We also include trends in county-level house prices, obtained from the FHFA house price index and the 2000 Census.[19]

b. Broadband Availability

Our goal is to estimate how broadband availability affects students' applications.[20] Since there is no measure of broadband availability in the CB data, we construct one from outside sources, combining zip-code level ISP counts published by the FCC with aggregate broadband usage rates from population surveys. Because individual households have very little control over whether and when providers enter their zip code, and very little impact on aggregate usage, our measure is arguably exogenous to student use of broadband. In this section, we describe the measure and save our examination of its exogeneity for the analysis section.

Broadband ISP coverage is published in FCC Form 477 Filing data.[21] The FCC requires every facilities-based provider with at least 250 high-speed lines to report basic information about its service offerings and end users twice a year. The FCC releases summary statistics to the public aggregated to the zip code-level, consisting of a list of zip codes with the number of ISPs

[19] We construct a county-level measure of house prices by combining county-level median home prices from the 2000 Census with the Federal Housing Finance Agency house price index, as in Dettling and Kearney (2014). Urban counties use the Metropolitan Statistical Area (MSA) version of the index and rural counties use the rural index.
[20] We might wish to go beyond questions of availability and estimate 2SLS effects of household or student broadband usage, which, while likely endogenous to educational outcomes, captures a student's ability to access online content. Unfortunately, usage rates cannot be constructed for all years at the subnational level. The PEW data are available frequently but only at the national level. The CPS data include state identifiers but are only available in 2000, 2001, 2003, 2007, and 2009. Later, we relate availability – aggregated to the state-level – to state-level teen usage that can be derived from the CPS and use this relationship to form suggestive estimates of the effects of use.
[21] Small providers, many of whom serve sparsely populated areas, are not required to report to the FCC but sometimes provide information on a voluntary basis. In our analysis, we provide separate treatment for rural and urban areas. The FCC data can be downloaded from https://www.fcc.gov/encyclopedia/form-477-data-zip-codes-number-high-speed-service-providers.

who have at least one subscriber within the zip code receiving speeds of 200 kbps or more.[22] The data are available bi-annually from December 1999 to June 2008, and to protect confidentiality, do not distinguish between one, two, or three providers in a zip code. Over that time, there is considerable variation both across and within zip codes.

While there is general consensus that the FCC data are the best available data measuring broadband roll-out over our period of study, the raw FCC data alone do not adequately reflect households' access to broadband (GAO, 2006; Connolly and Preiger, 2009). The reasons are twofold: first, a provider need only have one subscriber in a portion of a zip code to be counted in the data, so the counts themselves are not economically meaningful and tend to overstate household availability, and second, the binned nature of the data make zip codes with few providers difficult to compare.[23] To illustrate, Figure 1 compares nationally aggregated coverage rates implied by the FCC data to survey-reported national usage rates of high-speed Internet, published by PEW Research.[24] The figure offers a reasonable litmus test for how well the raw FCC data capture market penetration – i.e., how provider counts translate into usage – at least at the national level. There are large discrepancies between survey-reported usage and the fraction of the population residing in a zip code with at least one provider, in both levels and trends. The next available cutoff – four or more providers – does little to improve these discrepancies.

The inability of the raw FCC data to capture national trends in usage is not surprising given the vast heterogeneity in geographic and population sizes across the roughly 32,000 zip codes in the United States. Consider, for example, two zip codes that had one to three providers in 2000:

[22] A "subscriber" can be either a residential or small business customer. Larger businesses and institutions are not included as they typically use an alternative technology.
[23] GAO (2006) concludes that defining access according to provider presence alone "…may overstate deployment in the sense that it can be taken to imply there is deployment throughout the zip code even when deployment is very localized." A case study is offered for Kentucky, where 95 percent of residents had a provider in their zip code, but only 70 percent had access.
[24] These rates are extremely similar to those reported in the CPS for 2000, 2001, 2003, 2007 and 2009.

82332 is a rural zip code in Savery, WY (134 residents in 1,422 square miles), and 10030 is an urban zip code in New York, NY (25,847 residents in 0.30 square mile). By 2008, 82332 had 4 providers and 10030 had 11. It seems unlikely that all residents had equal access to broadband across the two zip codes, suggesting a "one size fits all" measure is inappropriate.[25]

Despite the challenges posed by the raw FCC data, prior studies offer intuition that we use to guide a conversion of the number of ISPs in a zip code into a comparable metric across zip codes and time that captures the average resident's ability to access high-speed Internet in her home. First, Connolly and Preiger (2009) point out that because the data were collected in the same manner over a relatively long time period, changes within zip codes over time can likely be meaningfully interpreted. In addition, GAO (2006) finds that ISPs rarely overlapped service territories; thus, the number of ISPs in a zip code should roughly translate into the extent to which that zip code is covered, relative to its size. Finally, GAO (2006), together with Faulhaber (2002), Greenstein and Price (2007), and Grubesic and Murray (2002), indicate that supply-side constraints may have been structurally different in urban and rural areas.[26]

Following these observations, we propose a measure that first scales the number of providers in a zip code by its size and then converts the scaled measure into an indicator variable, after it crosses a given threshold.[27] Since there is no theoretical guidance for how to determine a

[25] Prior research using the FCC data sidesteps comparability issues by limiting analyses to fairly homogeneous zip codes, either by investigating outcomes within a single state – where zip codes are quite similar structurally – or by removing high- and low-density zip codes (Vigdor, Ladd, and Martinez, 2014; Xiao and Orazem, 2010), neither of which would be desirable in our setting.

[26] Specifically, in rural areas, where zip codes are much larger and less densely populated, coverage was constrained by the cost of extending additional lines long distances to reach relatively few customers (GAO, 2006), while, in urban areas, population density was the major constraint because too many customers using a single line at once could exhaust the system (Faulhaber, 2002; Greenstein and Price, 2007; Grubesic and Murray, 2002).

[27] We prefer to dichotomize the measure for several reasons. First, a dichotomous measure has a more straightforward interpretation in terms of "availability" than the scaled measure, as the literature provides no guidance on what it means for a zip code to have, say, one additional provider per person or square mile. Second, under a linear specification of availability, zip codes with very many ISPs (for their size) will necessarily be given the most weight in our analysis, which is undesirable given our goals. Last, the ISP counts are published in bins, so they are not truly continuous to begin with; linearizing them necessarily requires judgment.

zip code's size or for what an appropriate threshold might be, we develop an algorithm that chooses these metrics empirically based on their ability to fit the national usage data in Figure 1, allowing for a potential separate treatment of urban and rural zip codes. The best-fit measure defines availability as "at least one provider per 12 square miles" in rural zip codes and "at least one provider for every 2,700 people" in urban zip codes.[28] Additional details on this measure's construction are available in the appendix, where we also present estimates using alternative measures of broadband availability. (We will summarize these estimates in Section IV.)

Figure 1 plots broadband availability using our preferred measure, aggregated to the national level. Unlike the providers-only-based measures, our measure closely follows levels and trends in survey-based usage. In 2000, both measures are close to zero, and by 2008, both are around 60 percent. To confirm that our measure translates into teen usage, we turn to the Current Population Survey (CPS), which collected information on broadband usage in 2000, 2001, 2003 and 2007, and find that the fraction of the state population with broadband in their zip code indeed predicts state-level teen usage rates (Appendix Table 1). In fact, among other potentially viable aggregated measures of availability, ours performs the best on this metric, with some measures (such as the "at least one ISP" measure) even negatively correlated with usage.

Appendix Figure 2 maps availability in December 1999 and December 2007, where dark gray zip codes have broadband and light gray zip codes do not. There is clear cross-time and cross-zip code variation in coverage. In 1999, few zip codes had broadband, and those that did were clustered in the major population centers on the East Coast. By 2007, coverage had

[28] We define "best fit" as minimizing the root mean squared error between each measure and the survey-reported trends. Zip code population data are based on a combination of Census 2000 population data and SOI population data. SOI data provides a count of the number of income tax filers, which was used to create population trends to move the 2000 Census zip-code population total forward. Zip-code land area estimates are from the 2000 Census.

expanded across the United States, although many areas still did not have coverage. Around 30 percent of our sample resided in a zip code with broadband during their junior year (Table 1).

IV. Analysis

In this section, we first propose a framework to analyze how the availability of high-speed Internet affects college applications. We next present and discuss our main estimates of the effects of broadband availability on SAT scores and, drawing from prior research, on application outcomes that capture deviations from defaults or perceived recommendations. Then, because the analysis relies on College Board records, we show that, within our framework, broadband availability does not coincide with observable changes in sample composition. We also examine the assumptions that underlie a causal interpretation of these estimates, and the sensitivity of our results to other specifications of broadband availability. Finally, we investigate whether responsiveness appears to vary by demographic or socioeconomic group.

a. Empirical Specification

Our estimating equation is a generalized difference-in-differences:

$$y_{izc}=\alpha+\beta*\text{broadband}_{izc}+\mathbf{A_i}\boldsymbol{\theta}+\mathbf{X_{zc}}\boldsymbol{\lambda}+\gamma_z+\gamma_c+\varepsilon_{izc}$$

where y_{izc} is a binary variable capturing an application outcome or is an exam scale score for student i from zip code z in cohort c and broadband_{izc} is an indicator for broadband availability in her junior year. $\mathbf{A_i}$ is a vector of student characteristics, including demographics (e.g., gender, race), high school GPA, and PSAT math and verbal scores, the inclusion of which captures fixed differences in applications according to a student's "latent ability" (or, more simply, an early signal of her college admissibility).[29] $\mathbf{X_{zc}}$ is a vector of cohort-varying zip code-level economic

[29] Results are not sensitive to the exclusion of the PSAT score as a control (Appendix Table 2), and, when we examine the stability of our sample, we demonstrate that broadband availability is not correlated with either PSAT-taking or composite PSAT scores. Additionally, in Appendix Table 3, we add controls for student survey responses

conditions (e.g., local income, unemployment rates, house prices, population density) that might correlate with broadband availability and applications. The cohort fixed effects (γ_c) absorb national trends in broadband adoption and the availability of online content, and the zip code effects (γ_z) absorb fixed differences between zip codes that received broadband earlier and later.

The key parameter of interest is the β coefficient, which reflects the difference in an application outcome, y, attributable to broadband availability. For example, when y is "applied to a four-year college," our estimate represents the effect of broadband availability on the likelihood a student applies to a four-year college. This characterization of β holds under the assumption that, all else equal, broadband availability is exogenous to students' applications. We formally investigate the validity of the identifying assumption after we present the main results.

Finally, we make two notes regarding interpretation. First, the β coefficients reflect the effects of *potential* access, as opposed to use, because $broadband_{izc}$ measures local availability rather than student-specific adoption. (In other words, our estimates are interpretable as "intent to treat" effects of availability, not "treatment on the treated" effects of use in college search.) This parameter is arguably the most relevant for policy, since how well broadband availability translates into effective use depends on how many students use it, how often they use it, and in what ways, all of which are factors that cannot be controlled through policy nor even reliably measured. Moreover, local peer effects—wherein the availability of broadband affects the college-going culture of a student's neighborhood or high school, even when she does not have access at home—could amplify or reduce the effectiveness of use. Such effects are difficult to disentangle from the pure effect of use, complicating the interpretation of estimates; whereas, the effect of availability subsumes peer effects, and estimates are straightforward to interpret.

---

to parental education and income, questions that do not appear on the PSAT survey and that are associated with considerable non-response on the SAT survey. Results are qualitatively unchanged to their inclusion.

Nonetheless, we will provide a suggestive rescaling of our coefficients (under some assumptions regarding use) to compare them with those from more-intensive or more-targeted interventions.

Second, we estimate the effects of broadband availability over a very short window of exposure, i.e., beginning in junior year, in order to most cleanly identify how reducing informational frictions in the college search affects the applications students send.[30] While such frictions may be present at even earlier stages, other mechanisms—that potentially interact with these frictions—are more likely to also be operating over longer horizons. For instance, young students with broadband may improve their coursework, which then may result in better applications. Some extended results will suggest that effects cumulate over time, but we will generally caution against this interpretation and leave this possibility for future work to address.

b. Main Outcomes: SAT Scores and Applications

First, we examine the extent to which broadband availability affects SAT scores. Then, drawing from existing research on college quality and under-match, we estimate effects on applications. Each application outcome considered is intended to capture a deviation from some default or perceived recommendation, such as applying to more schools than the number allotted with a test registration (i.e. applying to five or more), applying to a four-year school, applying to a selective school, applying to a match school, applying to a state flagship, applying out of state, and applying to a highly ranked liberal arts school. We construct each application measure as an indicator variable and we test, in line with the literature, whether the likelihood a student will deviate from perceived defaults increases with more information.

[30] Technically, because our estimation sample includes students who may have had broadband in their zip codes during periods other than junior year, our coefficients could be larger than those we would recover if we had limited the analysis to zip codes in which broadband did or did not *become available* junior year. However, the event-study results and the evidence we bring to bear on the stability of our sample over time both imply that longer periods of exposure are not driving our estimates.

Table 2 displays estimates of the effects of broadband availability. Students in zip codes with broadband outperform their peers by 0.7 SAT point (or 0.3 percent of a standard deviation). The remaining coefficients imply applications also improve: we estimate statistically significant gains in the probability a student applies to more than the default number of colleges, a four-year college, a selective college, a college commensurate with her own score (i.e., a match college), a top liberal arts college, and an out-of-state college ranging, from 0.2 to 0.4 percentage point. In some instances, these changes reflect meaningful deviations from mean behavior; for example, only 7.2 percent of SAT-takers apply to top liberal arts colleges, and broadband availability induces a 0.17 percentage point change over that baseline.

While positive in sign, the effect on whether students apply to an in-state flagship is statistically indistinguishable from zero.[31] This null result could reflect disparities in student groups to whom the gains from broadband tend to accrue, or could speak to more information-based phenomena. For instance, students who under-match often apply to public, in-state colleges (Hoxby and Avery, 2013), so we might not expect large effects among this class of institutions. Moreover, flagship colleges tend to be large and well-known; thus, if simple awareness was preventing applications, we similarly might not expect large effects.[32]

As noted earlier, our estimates do not represent the effect that broadband has on students who use it, but rather the effect of making broadband more readily available. To aid in the interpretation of our results, we may wish to scale our estimates by the increase in use that availability induces. Unfortunately, we do not observe a very good measure of use for our setting

[31] While not shown, the estimated effect on application to a very selective college (i.e., a college where the average SAT score is greater than 1300) is similarly indistinguishable from zero—0.00055 (0.00074)—potentially because broadband may not be as large a technological gain for the applications of very elite students. Estimates by PSAT score group, described later, support this interpretation.
[32] Hoxby and Turner (2015) find that the reasons students give for not applying to the state flagship seem more related to the social environment at the school than to unawareness of its academic quality.

– e.g., college-related Internet search activity by teens – so we instead scale by the increase in statewide teen general usage predicted by statewide availability (i.e., 0.14 p.p. from Appendix Table 1). We stress that the scaled estimates are merely suggestive because the "first stage" is derived from very few data points and, more importantly, the exclusion restriction (i.e., availability only affects outcomes through usage) likely does not hold, particularly in the presence of spillovers and peer effects. Nonetheless, the scaled estimates offer a rough magnitude by which to compare the implied effects of use against those from prior literature.

Taken at face value, the scaled estimates suggest that broadband use increases SAT scores by 5 scale points and applications by between 1.4 and 3.0 percentage points, depending on the outcome. By comparison, this implied effect on scores is about 6½ percent of the 78-point effect that Angrist et al. (2016) obtain in their analysis of Boston charter schools.[33] The implied effects on applying to at least 5 schools and to a liberal arts school are about one-quarter and one-third, respectively, the size of those reported by Hoxby and Turner (2013) in their semi-customized mail-based experiment. Altogether, in relation to these frontier studies of interventions that were relatively intensive and, in the case of Hoxby and Turner (2013), relatively targeted, the implied effects of broadband use that we estimate appear to be both economically meaningful—particularly with respect to its influence on application sets—and reasonably sized.

c. Validity of Research Design

i. Analysis of Test-taking Rates and Sample Composition

Our analysis relies on comparisons of outcomes among students who take the SAT exam. For most students, the SAT is optional, taken in preparation to apply to college; thus, relative to the

[33] Angrist et al. (2016) exploit variation from random admission to Boston charter schools. The schools admitted students in 9th grade, so exposure to treatment before the SAT was generally longer than in our setting. According to the authors, relative to their counterfactual high school experience, charter students had longer class days, improved teacher quality, an increased focus on core instruction, and/or more-able peers.

full population of U.S. students, our sample is likely positively selected from the distribution of student ability. Moreover, there exists a competing exam for college-bound students, the ACT, that students can elect to take instead. Thus, caution should be used when generalizing our estimates to broader student populations.

A higher-order concern is whether the opt-in nature of the SAT complicates the *within-sample* interpretation of estimates, since our outcomes can only be observed when students take the SAT. While it would be interesting if the external margin of SAT-taking was affected by broadband availability, if there were shifts along this margin—especially if the quality of test-takers changed—applications and scores across zip codes with and without broadband would not necessarily be comparable.

To assuage this concern, we investigate whether SAT-taking or the composition of our sample is correlated with broadband in Table 3. In the first two columns, we examine SAT-taking under two specifications of the at-risk (i.e., non-selected) population of SAT-takers.[34] First, we leverage information available from the PSAT. The PSAT is the qualifying test for the National Merit Scholarship Program and is often mandatory within a state or district. Using the full set of PSAT takers, we specify y_{izc} as an indicator value that takes the value of 1 if a student also takes the SAT exam and 0 otherwise. The estimated effect of high-speed Internet on SAT-taking is extremely small (0.00049) and statistically indistinguishable from zero (column 1). While far less likely, interaction with the PSAT could evolve with broadband.[35] Thus, we also

[34] Ideally, to examine selection, we would use all students who are at risk of applying to college when broadband becomes available (e.g., high school juniors by home zip code over time), but we do not observe such a measure in our data and, to our knowledge, these data are not available elsewhere.

[35] The PSAT is generally administered to high school sophomores and juniors in the fall, leaving a very small window for students to alter their PSAT-taking decisions in response to broadband availability in December of junior year. Further, students induced into PSAT-taking would generally not have been otherwise college-bound and thus are extremely unlikely to apply broadly or to schools in the selectivity ranges we consider, so their inclusion would bias estimates toward zero.

specify y_{izc} as the fraction of high school junior-aged individuals in a zip code that take the SAT and, for completeness, the fraction that take the PSAT, whereby the denominator is inferred from the size of the relevant cohort in the 2000 census, aged forward or backward.[36] Again, the estimates are extremely small and statistically insignificant (columns 2 and 3). In sum, the likelihood a student appears in our sample does not seem to coincide with broadband availability.

Still, there may be changes in the composition of SAT-takers underneath this non-effect. For instance, the number of more-able students induced to take the exam could directly offset the number of less-able students who are discouraged from taking the exam. In the right columns of Table 3, we investigate whether the composition of SAT-takers appears to vary with broadband along several characteristics, including: minority status, parents' education, and student PSAT scores. Across the board, we obtain results indistinguishable from zero, implying that the composition of our sample in zip codes with and without broadband is likely stable as well.

Finally, even if our analysis sample is observably unchanged, the existence of a competing admissions exam could introduce bias to our estimates if students' reliance on the SAT for admissions purposes systematically changed with the availability of broadband. For example, students with broadband may newly adopt a strategy that entails taking both the ACT and SAT and then sending only their highest score to colleges. We examine this concern in two ways and present the results in the appendix. First, we exploit the institutional background of the two exams—namely, that, while colleges throughout the United States mostly accept either exam, the tests have historically prevailed in particular geographic regions, with students on the coasts and in the South likely to take the SAT and those in the Midwest, the ACT—and show that estimates

[36] We define the at-risk PSAT cohort as individuals aged 15-17 and the at-risk SAT cohort as those aged 16-18, in a given zip code and year. Due to data limitations, these cohort measures do not adjust for migration or mortality.

are generally not sensitive to the exclusion of ACT states (Appendix Table 4).[37] Second, we show that broadband availability is uncorrelated with the number of ACT test centers in a zip code (Appendix Table 5), which—given that the presence of a local test center is highly related to exam-taking—suggests that availability is uncorrelated with ACT-taking behavior as well.[38]

From this set of results, we infer that student interaction with the CB likely does not coincide with broadband availability in a way that would complicate the interpretation of our estimates.

ii. Examination of Identifying Assumption

For a causal interpretation of our estimates, our identifying assumption is that, all else equal, broadband availability is exogenous to students' applications. Given our framework, there are two conditions under which this assumption would be violated: student-demand-related pressures on providers to enter zip codes, and omitted variables that co-vary with applications and entry. Per the former, there is abundant evidence that supply-side constraints restricted broadband access, and that the supply of broadband lagged consumer demand. To provide high-speed Internet, an ISP – typically the existing phone or cable company – had to make substantial infrastructure investments, retrofitting existing phone and cable lines and installing new switches and servers (Faulhaber, 2002; Greenstein and Prince, 2007; Grubesic and Murray, 2002). There is a general consensus that the costs slowed rollout and access did not keep up with consumer demand (Greenstein and Prince, 2007; Faulhaber, 2002). Dettling (forthcoming) discusses how variation in the underlying housing infrastructure and the availability and quality of existing telephone and cable wiring made these infrastructure investments differentially costly across locations, which created differences in the roll-out of broadband across locations that were

[37] For completeness, we re-estimate the specifications examining selection into SAT-taking (i.e., the top panel of Table 2). The coefficients (standard errors) on "PSAT Score" and "Took the SAT" are -0.06520 (0.03571) and 0.00175 (0.00098), respectively, neither of which is statistically different from zero (at the 5% level).
[38] Locations of active ACT test centers between 1999 and 2007 were generously provided by George Bulman.

unrelated to consumer demand-related pressures.[39] Based on these known barriers to entry, it seems unlikely that, net of the fixed effects and set of economic controls in our model, student-demand-related pressures induced the provision of broadband.

For the latter to be plausible, our extensive set of student- and zip code-level controls must not adequately capture factors that correlate with broadband availability and applications. The remainder of this subsection extensively investigates this possibility.

We begin by testing for an effect of broadband availability on zip code income, measured using adjusted gross income (which is one of our zip code-level control variables) and wage and salary income (which is not)—which might suggest that the differences in applications we detect are driven by variation in conditions that correlate with broadband availability rather than broadband availability itself—and find none (Appendix Table 6). Further, when we include the other time-varying zip code-level control variables that we used in our main specification (i.e., population density, unemployment rates, and home prices), the coefficients are only marginally statistically significant and, more importantly, are wrong-signed, which, if anything, would indicate that broadband availability is *negatively* correlated with income. Thus, it does not appear there are important changes in local economic well-being that are generating our results.

Next, we provide two checks of whether application trends in zip codes appear to have been evolving similarly over time, save for the availability of broadband. First, we compare application trends in zip codes before 1999—when broadband began to become available to

[39] There are two main transmission modes for broadband in the United States: cable-based and telephone-based digital subscriber line (DSL) service. The main issue that constrained cable companies was capacity. While they had previously installed some lines for digital cable service that could be used to provide broadband, each additional customer on a single line reduced the "downstream" capacity, meaning that multiple simultaneous users would exhaust the system. To provide reliable service, cable companies needed to add more lines closer to residences. DSL providers were mostly constrained by the need to upgrade the existing telephone wiring, much of which was old and had been split too often to be capable of carrying high-speed two-way traffic. Dettling (forthcoming) demonstrates that infrastructure constraints generated variation in the timing of ISP entry.

residential customers—according to when they would later receive broadband. Specifically, we classify zip codes according to whether broadband became available during the early (1999–2001), middle (2002–2005), or late (2005–2007) years of our sample period, and chart the evolution of applications for the 1996 to 1998 graduation cohorts in Figure 2.

The figure reveals marked differences before broadband technology became commonplace, whereby earlier-adopting zip codes tended to have better outcomes. These differences are not surprising—ISP entry is correlated with wealth and wealth is correlated with applications—and underscore our inclusion of zip code effects. The persistence of these differences is what is most relevant for our specification; since trends in the prior period look similar, we might infer that applications—without broadband—continued to evolve similarly over the sample period as well.

Second, in order to comprehensively examine the pre-trends together with the dynamics following the introduction of broadband, we follow Autor (2003) and estimate an event-study version of our estimating equation. Specifically, we replace $broadband_{izc}$ with a series of dummies for the years prior to and following the introduction of broadband over our sample period. Note that a more tractable interpretation of this specification in our setting is that the estimates represent the effect of broadband according to where students were in their educational careers when it became available (e.g., broadband availability for two years equates to availability since freshman year of high school, and availability in two years equates to broadband becoming available during freshman year of college). As such, we would expect to see: no systematic differences in applications among students who were already freshmen in college or older by the time broadband was available in their home zip code (i.e., pre-trends), a considerable jump among high school juniors, and similar (or larger) differences among students who were younger than juniors when broadband became available. For consistency and ease of interpretation, the omitted dummy in each equation corresponds to senior year availability.

Figure 3 summarizes the results, where the x-axis plots the year of schooling when broadband became available (from oldest to youngest) and the y-axis displays the estimated coefficients, with confidence intervals, on broadband availability in the year listed. Students who were beyond application age when broadband became available do not exhibit statistically different outcomes; thus, again, there is little evidence of anticipatory effects or heterogeneity in application pre-trends.[40] By contrast, coefficients are positive and highly significant for students who were application age or younger when broadband became available, with a clear jump in applications coincident with junior year availability, altogether affirming our design. Note that, while we caution against the interpretation of estimates over very long horizons, especially since the composition of students using broadband within a zip code could change over time, the coefficients appear to swell the longer that broadband is available, suggesting that the benefits of broadband may be larger if it is introduced earlier in the application cycle.

Finally, as an additional check on our design, we examine whether our estimates move as expected over other timings of availability and definitions of our sample. First, we estimate how broadband availability in senior year of high school affects applications. By that point, students would already have taken their SAT and, more generally, would have much less time to research schools or otherwise improve their applications. Indeed, the estimates reflect these differences (Appendix Table 7). Students with broadband in senior year also exhibit improvements in applications, but the coefficients are smaller. And, while still positive, we no longer detect

[40] The one exception is that the policy leads for “Apply 5+” appear to indicate pre-trends, which may be a statistical artifact of examining numerous outcomes.

significant effects on the probability a student applies to a selective school or out-of-state school. Most importantly, there is no detectable effect on SAT scores, consistent with the later timing.

Second, we divide the sample into PSAT score groups – top quartile scorers and non-top quartile scorers – and examine responsiveness within each group. The intuition behind this exercise is twofold. First, test preparation that is available on the Internet is likely less valuable to high PSAT scorers, implying that those with the highest PSAT scores will exhibit the smallest responses (at least in terms of SAT scores). Second, high PSAT scorers are the most likely to already have sound information from teachers, counselors, and recruitment campaigns by colleges.[41] These lines of intuition together imply relatively small effects among high-scorers.

The estimates again conform to predictions (Appendix Table 8). Among the top PSAT scorers, effects on SAT scores and nearly all application outcomes are statistically indistinguishable from zero. Non-top scorers appear to drive our results. These findings echo those from prior studies. First, Angrist et al. (2016) similarly found that the largest gains from charter schools, in achievement and in college enrollment, were concentrated among students with lower initial scores. Second, despite their differences, both groups are more likely to apply to liberal arts schools when broadband is available, which implies that students have particularly little knowledge or understanding of such institutions, as in Hoxby and Turner (2015).

iii. Alternative Measures of Broadband Availability

Recall that we chose our measure of broadband availability over others that could be constructed from the ISP data based on its ability to fit national usage trends and to predict teen usage (Appendix Table 1). Appendix Tables 9 and 10, described fully in the appendix, provide

[41] PSAT scores are used by colleges to identify students to recruit not only because the score is one of the earliest signals of ability/admissibility that is measured on a national scale, but also because students with the highest scores might qualify for the National Merit Scholarship.

estimates under alternative definitions of broadband availability—by altering the thresholds (e.g., one provider per 2000 people or per 10 square miles) and using a linear measure (providers, providers per person, providers per square mile)—both overall and within the urban and rural subsamples. In general, the results become more similar to our main estimates the better the measure fits usage trends and the more predictive the measure is of teen usage. Measures that perform very poorly on these two metrics (e.g., the linear measures) tend to generate statistically insignificant results, which we attribute to attenuation bias due to measurement error.

d. Heterogeneous Response

Now, we examine responsiveness within particular segments of the student population. As described in Section 2, the under-match literature suggests that disadvantaged students would benefit most from broadband and exhibit the largest responses. However, since the β's reflect the effects of potential access, responsiveness may be mediated by other factors that correlate with SES (Appendix Figure 1), whereby disadvantaged students would exhibit the smallest responses.

Table 4 displays estimates by group. Improvements in applications are concentrated among higher-income, more-educated, white, and urban students. By contrast, the estimates for low-SES students imply large changes in SAT scores but a pervasive null effect on applications. These patterns may reflect how effectively students in each group use the Internet. In general, they suggest that benefits of broadband are lopsided and favor those who already had more resources.

These results may be specific to our setting, when broadband was first becoming available to households. It may be that in its early stages, the benefits primarily accrued to resource-rich students, but as it became more diffuse, lower-resource students experienced large benefits as well. We probe this hypothesis in an event-study framework mirroring Figure 3, splitting the sample by parental education. While we again caution the interpretation of estimates over long horizons of availability, a full read of these estimates might imply that low-SES students realize increased score gains over time but that their applications only improve if they essentially grew up with broadband in their zip code (Appendix Figure 3). (By contrast, for high-SES students, the gains in both scores and applications are immediate and persistent.) This pattern could imply: 1) there is learning among low-SES students in how to use the Internet to improve outcomes; 2) low-SES students are very late adopters; 3) low-SES students alter their admissibility over time in important ways we cannot observe (e.g., through changes in their coursework); or 4) the benefits of broadband disseminate across zip codes over time, even to non-users.

V. Discussion of Mechanisms

While we cannot observe the primary mechanism that generates our results, there are several leading explanations for how the availability of broadband could affect college applications—namely, 1) by reducing the direct effort and informational costs associated with applications (e.g., in submitting applications; in applying for financial aid; in researching school characteristics, admissibility requirements, or application strategies); 2) by removing distance costs from attending far-away colleges; 3) by improving admissibility through online test preparation for the SAT; 4) by improving admissibility through increased student investment in their broader academic records or other aspects of their application (e.g., search engines may improve the quality of or reduce time spent on class assignments, students may learn about the importance of extracurriculars in applications); or 5) deleteriously, either by distracting students or by offering an imperfect, less-valuable substitute for other options, like guidance counseling, that students then take up. Here, we describe the evidence for or against each.

First, our estimates are generally positive (or zero), implying that any deleterious effects (i.e. #5) are negligible, on net. Note that we can rule out deleterious mechanisms on average as well as for the subgroups we consider in our heterogeneity analysis.

Second, while broadband availability increases SAT test scores, potentially stemming from increased test preparation, there appear to be meaningful improvements in applications independent of such increases. First, when we consider the senior year timing of broadband, we observe positive application effects without test score effects. Second, when we add a separate (endogenous) control for SAT score to our main specification, the estimates of β shrink by no more than 20 percent in each instance (Appendix Table 11).[42] Thus, likely only a small share of the improvement in applications is attributable to increases in student admissibility via test preparation (i.e. #3), with the rest driven by a separate phenomenon.

Third, while we do not observe students' complete records, we can make some related observations. One, student SAT-taking rates do not appear to change with broadband, so if students were adjusting other aspects of their application, like their academic records, there is no adjustment along a rather-obvious margin. Two, the event-study analysis reveals large responses specifically among juniors, even though meaningful changes in academic records, which would be complementary to applications, likely require more time than the two semesters of high school before applications are due (January 1 or so of senior year).[43] Last, as a suggestive exercise, we estimate the effect of broadband availability on the GPA that students report on their SAT in-take survey, and the coefficient is only weakly significant and, unlike applications,

[42] SAT scores are likely positively correlated with broadband and our application outcomes, so the estimated effect of broadband on applications in a regression that includes SAT scores should, if anything, be downwardly biased.
[43] The event-study analysis might suggest that improvements in student records may be a meaningful part of the story for longer periods of broadband availability.

not stable to the inclusion of SAT scores.[44] All told, it seems unlikely that the application effects are driven by non-CB aspects of admissibility that improve with broadband availability (i.e., #4).

Finally, while we detect effects on out-of-state applications, which could imply that there is scope for college distance alone to be a relevant margin of adjustment (i.e., #2), the broader set of evidence points to changes in this outcome via an information or effort cost angle (i.e., #1). The full set of estimates reveal a systematic pattern of students with broadband straying from defaults, of which applying to only in-state schools is a subset. Further, while most of our estimates maintained statistical significance under the senior year timing of broadband availability, the coefficient on out-of-state applications is one of two that did not. Finally, the largest responses, even focusing exclusively on the out-of-state estimates, occur within the highest-resource groups (who likely would not be particularly constrained on distance).

Thus, given our full set of results, together with the evidence we cite on uses of the Internet (e.g., Appendix Figure 1, footnote 12), our preferred interpretation of how broadband improves college applications is the first—by reducing the direct effort and informational costs involved.

VI. Conclusion

This paper examines the effect of broadband on college applications. We find that the availability of broadband in a student's zip code during her junior year unambiguously improves her SAT score and her application set. These findings are consistent with related studies that document considerable effects of interventions late in adolescence on students' postsecondary outcomes. Specifically, our results imply that students with broadband are more likely to diverge from defaults and perceived recommendations: they send more applications than are allotted to them upon registration for the SAT, and they are more likely to apply to more-selective schools,

[44] The GPA effects are merely suggestive because GPA is self-reported, cumulative, school-specific, and potentially sensitive to the time at which students take (or register for) the SAT.

schools outside their state, highly-selective liberal arts schools, and schools commensurate with their abilities. The largest responses are concentrated among students from relatively high-resource backgrounds. A careful examination of the mechanism that best fits our full set of findings suggests that the advent of broadband technology considerably reduced the direct effort and informational costs associated with applications to college.

Our results have clear implications for broadband policy, both within an educational setting and more broadly. Per the former, we document that students can benefit from content available online to improve their outcomes. However, the primary beneficiaries of broadband availability appear to be higher-resource student groups, implying that the digital divide may be neglecting students who already tend to have fewer inroads to elite academic institutions. If this gap is due to differences in student ability to access and use broadband – because parents do not take up broadband or because students do not have access to devices – policies that increase broadband access and affordability may reduce inequality in postsecondary outcomes. If these gaps are due to differences in the ability to find and digest the relevant online content, policies that provide guidance on how the Internet can enhance opportunities, potentially through in-school programs that encourage and monitor Internet searches, may be effective. We leave it to future work to uncover which margins are most relevant. Per the latter, our results underscore key design features of general broadband policies. Because we find that the availability of broadband does not benefit all types of households equally, multifaceted broadband initiatives (e.g., BroadbandUSA), which entail steps to improve connectivity, access, *and* use, will likely see much larger gains than “Last Mile” initiatives, which focus solely on connectivity.

Further, even though the introduction of broadband technology is a rare event, unlikely to be repeated in its exact form, the effects we detect speak to broader policy issues. First, our results offer new evidence on informational frictions in college applications. While many interventions target children from disadvantaged backgrounds, children from well-off families also appear to behave as though they are information constrained. Second, we document a pattern of diffusion of information technology, whereby, even within coverage areas, the highest SES households appear to be the earliest adopters (or most effective users). If this pattern translates into real differences in productivity, stemming either from differential familiarity with the technology or skill differences that result from its use, we might see structural shifts in the labor market similar to those that have been attributed to the advent of related technologies (e.g., Skinner and Staiger, 2015; Autor and Dorn, 2013; Goldin and Katz, 1998; Autor, Katz, and Krueger, 1998).

An important lingering question is how the application improvements we estimate translate into improved attendance. While we cannot observe enrollment outcomes for our sample, we can use comparable estimates from the literature to derive anticipated effects. In the national intervention launched by Hoxby and Turner (2013), 43 percent of the detected change in applications to "peer" institutions among students who were sent material translated into improved attendance. If we apply those estimates to our "match" application coefficient, we can expect a 0.12 percentage point increase in the likelihood that college-bound students enroll at schools commensurate with their ability. Assuming no supply-side constraints or general equilibrium effects, which are both beyond the scope of this paper, this increase in enrollment decreases the extent of under-match by about 3,500 students per year.[45]

[45] Statistic derived from the Digest of Education Statistics 2013 count of first-time freshmen.

References

Aigner, D. (1973). “Regression with a Binary Independent Variable Subject to Errors of Observation,” *Journal of Econometrics*, 1: 49-60.

Altonji, J. (1993). “The Demand and Return to Education when Education Outcomes are Uncertain,” *Journal of Labor Economics*, 11(1): 48-83.

Autor, D. (2003). “Outsourcing at Will: The Contribution of Unjust Dismissal Doctrine to the Growth of Employment Outsourcing,” *Journal of Labor Economics*, 21(1): 1-42.

Autor, D. and D. Dorn (2013). “The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market,” *American Economic Review,* 103(5): 1553-97.

Autor, D., L. Katz, and A. Krueger (1998). “Computing Inequality: Have Computers Changed the Labor Market?” *The Quarterly Journal of Economics,* 113(4): 1169-1213.

Avery, C. and T. Kane (2004). Student Perceptions of College Opportunities: The Boston COACH Program. In C. Hoxby, ed. College Choices: The Economics of Where to Go, When to Go, and How to Pay for It. University of Chicago Press.

Bailey M. and S. Dynarski (2011, December). Gains and Gaps: Changing Inequality in U.S. College Entry and Completion. Working Paper 17633, NBER.

Belo, R., P. Ferreira, and R. Telang (2013). “Broadband in School: Impact on Student Performance,” *Management Science*. 60(2): 265-282.

Bettinger, E. P., Long, B. T., Oreopoulos, P., & Sanbonmatsu, L. (2012). “The role of application assistance and information in college decisions: Results from the H&R Block FAFSA experiment,” *Quarterly Journal of Economics*, 127, 1205–1242.

Betts, J. (1996). “What Do Students Know about Wages? Evidence from a Survey of Undergraduates,” *Journal of Human Resources* 31(1): 27-56.

Bhuller, M., T. Havnes, E. Leuven and M. Mogstad (2011) “Broadband Internet: An Information Superhighway to Sex Crime?” *IZA Discussion Papers Series* No. 5675.

Bound, J., C. Brown, and N. Mathiowetz (2001). “Measurement Error in Survey Data.” In J. J. Heckman and E. E. Leamer (eds), *Handbook of Econometrics*, V: 3705-3843.

Bowen, W., M. Chingos, and M. McPherson (2009). *Crossing the Finish Line: Completing College at America’s Public Universities*. Princeton University Press.

Bulman, G. (2015). “The Effect of Access to College Assessments on Enrollment and Attainment.” *American Economic Journal: Applied Economics*, 7(4): 1-36.

Card, D. (1995). "Using Geographic Variation in College Proximy to Estimate the Return to Schooling," in *Aspects of Labour Market Behaviour: Essays in Honour of John Vanderkamp,* ed. by Louis N. Christofides, E. Kenneth Grant, and Robert Swidinsky. Toronto: University of Toronto Press, 201-222.

Card, D. and A. Krueger (2005). "Would the Elimination of Affirmative Action Affect Highly Qualified Minority Applicants? Evidence from California and Texas." *Industrial and Labor Relations Review* 58(3): 416-434.

Carrell, S. E. and B. Sacerdote (2013). "Late Interventions Matter Too: The Case of College Coaching in New Hampshire," NBER Working Paper, 19031.

Cohodes, S. and J. Goodman (2014). "Merit Aid, College Quality and College Completion: Massachusetts' Adams Scholarship as an In-Kind Subsidy," *American Economic Journal: Applied Economics*.

Connolly, M. and J. E. Prieger (2009) "Economics at the FCC, 2008-2009: Broadband and Merger Review" *Review of Industrial Organization* 35(4), 387-417.

Dale, S. and A. Krueger (2002). "Estimating the Payoff to Attending a More Selective College: An Application of Selection on Observables and Unobservables." *Quarterly Journal of Economics* 117(4), 1491–1527.

Dale, S. and A. Krueger (2011, June). Estimating the Return to College Selectivity over the Career Using Administrative Earnings Data. Working Paper 17159, NBER.

Dettling, L. (forthcoming) "Broadband in the labor market: The Impact of High-Speed Internet on Married Women's Labor Force Participation," *Industrial and Labor Relations Review.*

Dillon, E. and J. Smith (2013). "The Determinants of Mismatch between Students and Colleges," NBER Working Paper.

Ellwood, D. and T. Kane (2000). "Who is Getting a College Education: Family Background and the Growing Gaps in Enrollment." In *Securing the future,* ed. S. Danziger and J. Waldfogel, 283-324.

Faber, B., R. Sanchis-Guarner, and F. Weinhardt (2015). "ICT and Education: Evidence from Student Home Addresses." NBER Working Paper.

Fairlie, R. and J. Robinson (2013). "Experimental Evidence on the Effects of Home Computers on Academic Achievement among Schoolchildren," *American Economic Journal: Applied Economics* 5(3): 211-40.

Falck, O., R. Gold, and S. Heblich (2014). "E-lections: Voting Behavior and the Internet," *American Economic Review* 104(1): 329-389

Faulhaber, G. R. (2002). "Broadband deployment: Is policy in the way?" In R. H. Crandall and J. H. Alleman (Eds.), Broadband: Should We Regulate High-Speed Internet Access, AEI-Brookings Joint Center for Regulatory Studies, 223–244.

Gentzkow, M and J. Shapiro (2008). "Preschool Television Viewing and Adolescent Test Scores: Historical Evidence from the Coleman Study," *Quarterly Journal of Economics* 123(1): 279-323.

Goodman, S. (2016). "Learning from the Test: Raising Selective College Enrollment by Providing Information," *Review of Economics and Statistics*, 98(4): 671-684.

Goldin, C. and L. Katz (1998). "The Origins of Technology-Skill Complementarity." *The Quarterly Journal of Economics*, 113:683-732.

Greenstein, S. and J. Prince (2007). "The diffusion of the Internet and the geography of the digital divide in the United States" In D.Q. Mansell, Robin and R. Silverstone (Eds.), *Oxford Handbook on ICTs*, Oxford University Press.

Grubesic, T. H. and A. T. Murray (2002). "Constructing the divide: Spatial disparities in Broadband access." *Papers in Regional Science 81*, 197–221.

Hoxby, C. and C. Avery (2012, December). The Missing "One-Offs": The Hidden Supply of High-Achieving, Low Income Students. Working Paper 18586, NBER.

Hoxby, C. and S. Turner (2013). Expanding College Opportunities for High-Achieving, Low Income Students. Discussion Paper 12-014. Stanford Institute for Economic Policy Research.

Hoxby, C. and S. Turner (2015). What High-Achieving Low-Income Students Know About College. Working Paper 20861, NBER.

Hurwitz, M., J. Smith, S. Niu, and J. Howell (2015). "The Maine Question: How is Four-Year College Enrollment Affected by Mandatory College Entrance Exams?" *Educational Evaluation and Policy Analysis.* 37(1): 138-159.

Kane, T. (1995, July). "Rising Public College Tuition and College Entry: How Well Do Public Subsidies Promote Access to College?" Working Paper 5164, NBER.

Klasik, D. (2012). The college application gauntlet: A systematic analysis of the steps to four-year college enrollment. *Research in Higher Education*, 53(5): 506-549.

Klasik, D. (2013). "The ACT of enrollment: The college enrollment effects of required college entrance exam taking," *Educational Researcher*, 42(3): 151-160.

Luca, M. and J. Smith (2013). "Salience in Quality Disclosure: Evidence from U.S. News College Rankings." *Journal of Economics & Management Strategy*. Vol. 22(1): 58-77.

Pallais A (2015). “Small Differences that Matter: Mistakes in Applying to College.” *Journal of Labor Economics* 33(2):493-520.

Pallais A. and S. Turner (2006). Opportunities for Low Income Students at Top Colleges and Universities: Policy Initiatives and the Distribution of Students. *National Tax Journal* 59(2), 357-386.

PEW Research (2014) “Internet use over time” http://www.pewinternet.org/data-trend/internet-use/internet-use-over-time/ Accessed September 30, 2014.

Skinner, J. and D. Staiger (2015). “Technology Diffusion and Productivity Growth in Health Care.” *Review of Economics and Statistics*, 97(5): 951-964.

Smith, J. M. Hurwitz, and J. Howell (2015). “Screening Mechanisms and Student Responses in the College Market,” *Economics of Education Review*, 44:17-28.

Spies, R. (2001). “The Effect of Rising Costs on College Choice: The Fourth and Final in a Series of Studies on This Subject.” Princeton University Research Report Series, No. 117.

United States Government Accountability Office (2006) “Broadband deployment is extensive throughout the United States, but it is difficult to assess the extent of deployment gaps in rural areas” *GAO Report 06-426*

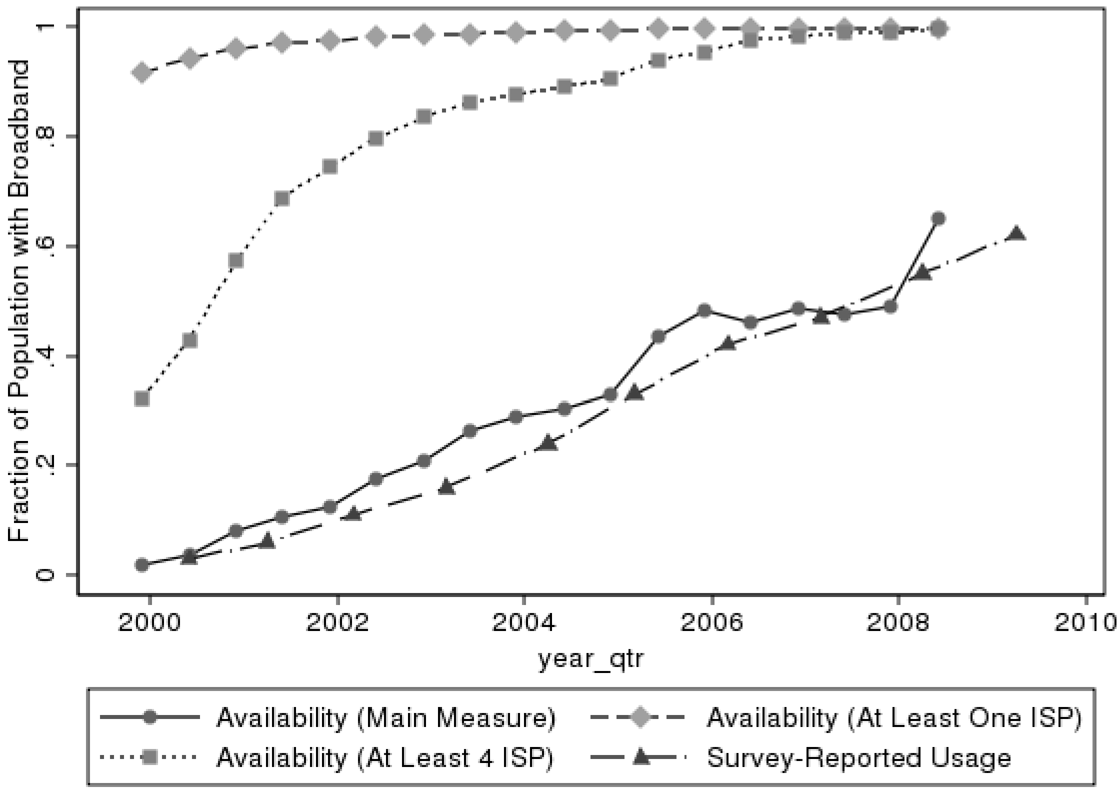
Vigdor, J., H. Ladd, and E. Martinez (2014). "Scaling the Digital Divide: Home Computer Technology and Student Achievement." *Economic Inquiry* 52(3): 1103-1119.

Zimmerman, S. (2014). The Returns to College Admission for Academically Marginal Students *Journal of Labor Economics* 32(4), 711-754.

Zucker, K. and A. Smith (2013). “Home Broadband 2013” August 26, 2013 http://www.pewinternet.org/2013/08/26/home-broadband-2013/

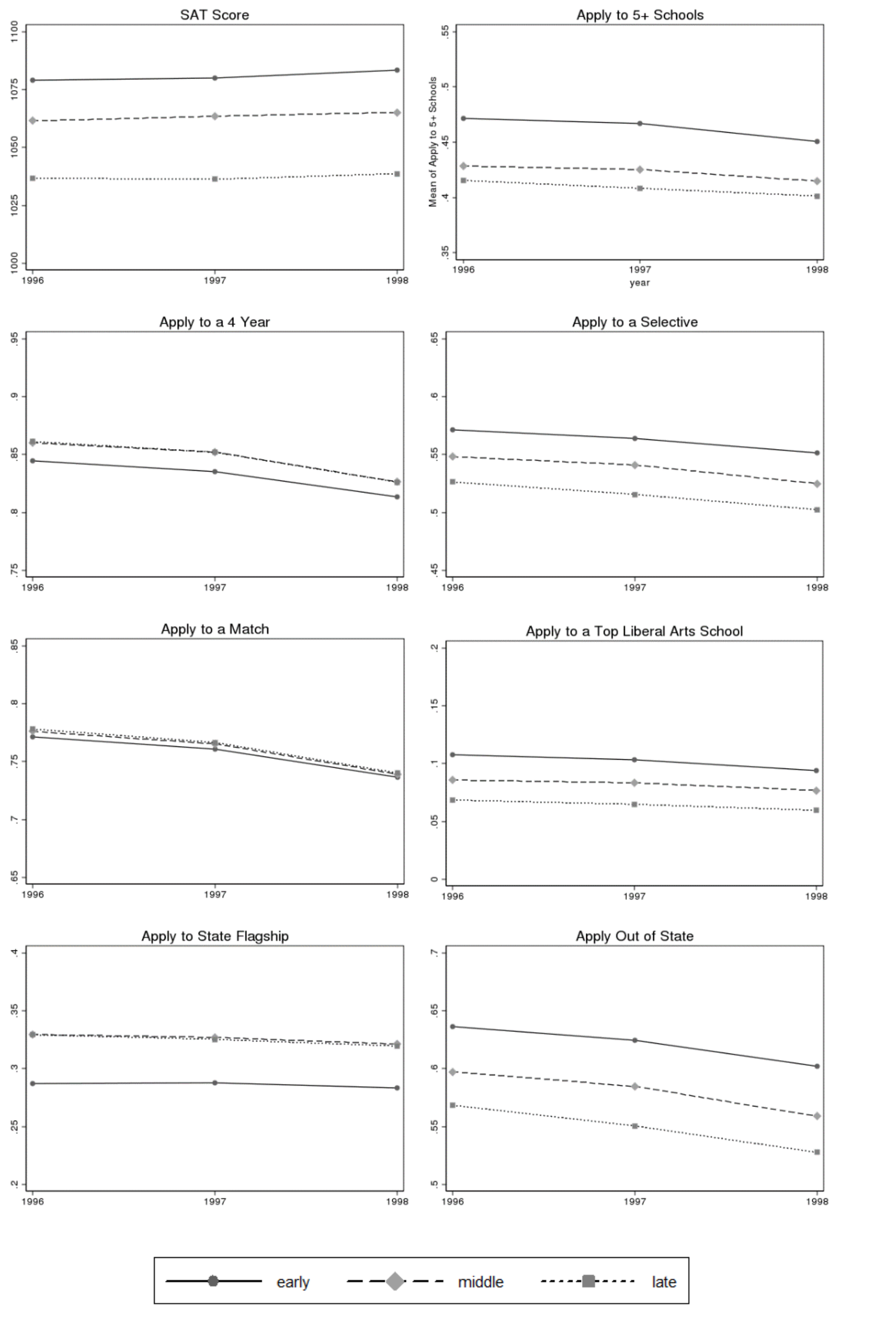
Tables and Figures

Figure 1: Measures of High-Speed Internet Availability and Usage



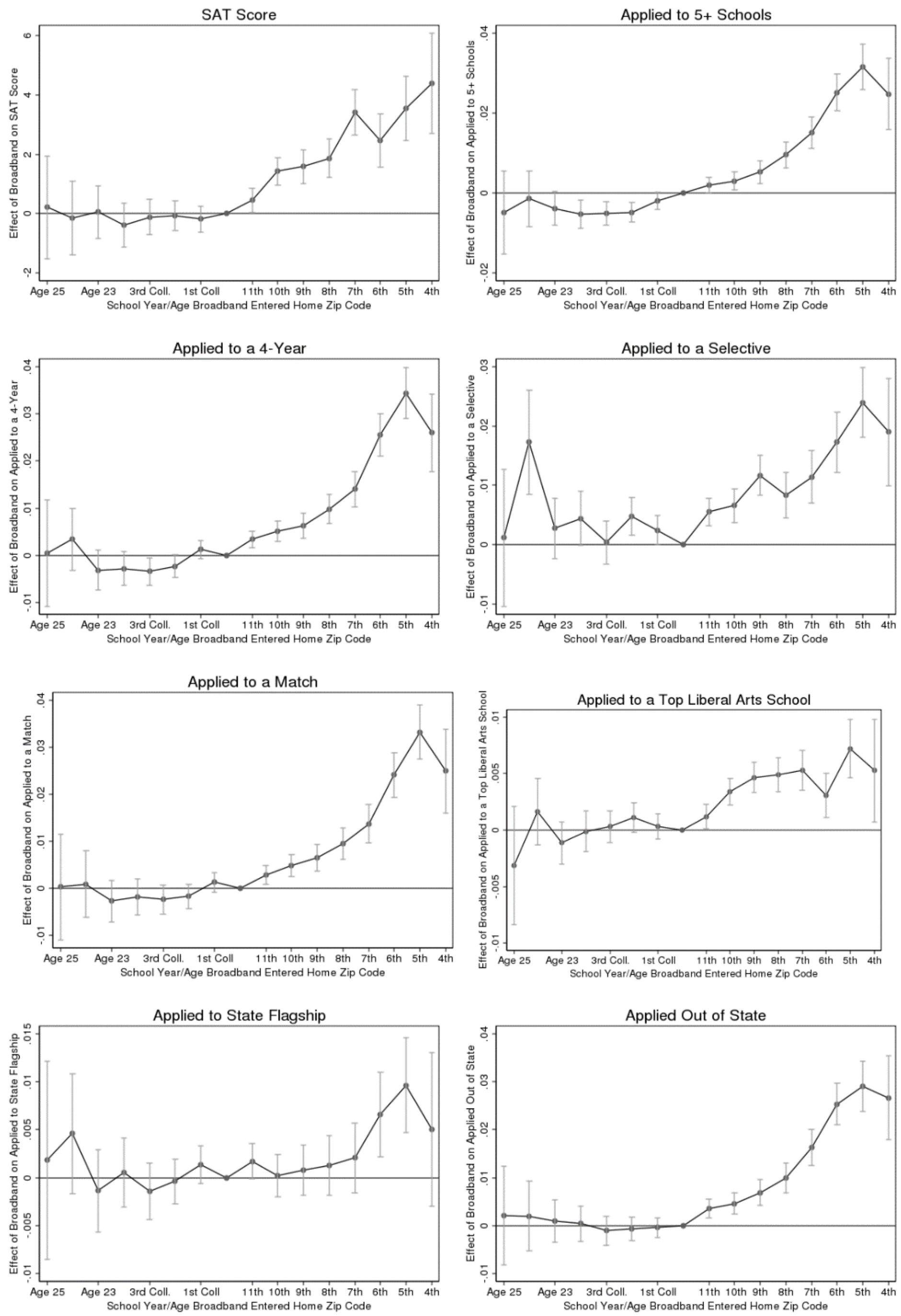
Notes: Displayed are trends in high-speed Internet availability and usage according to different measures. PEW refers to survey-reported national usage rates. The two provider measures (at least1 and at least 4) refer to implied aggregate availability rates, if availability is defined as having at least one or at least four providers in the zip code. These rates are weighted by zip code population. The coverage measure refers to our preferred measure of high-speed Internet availability, as described in the text.

Figure 2: Means of Selected Application Outcomes in 1996-1998 by Year High-Speed Internet Entered Zip Code



Notes: Displayed are population-weighted means of the main application outcomes in the years just prior to the introduction of residential broadband Internet (1996 to 1998). Means were constructed by classifying zip codes according to the year in which high-speed Internet became available in the zip code.

Figure 3: Event Study Estimates of Effect of Broadband Availability on Application Outcomes



Notes: Each graph represents coefficients and 95 percent confidence intervals obtained from OLS regressions whereby the specification is as in Table 2, except the single broadband availability dummy is replaced by a series of dummies for the years prior to and following the introduction of the broadband, which are labelled by year of schooling for ease of interpretation. The dummy for the year immediately preceding introduction (e.g., senior year) is excluded. Standard errors adjusted for clustering.. Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Regressions include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics and a constant.

Table 1: Summary Statistics

	All Students (obs = 7,452,302)			
	Mean	*Std. Dev.*	*Min*	*Max*
Student Demographics and Exam Scores				
White	0.672	0.470	0	1
Black	0.107	0.309	0	1
Latino/Hispanic	0.095	0.293	0	1
Other Race	0.126	0.332	0	1
Male	0.447	0.497	0	1
Parental Education (0: HS; 1: Some college)	0.841	0.365	0	1
Cumulative GPA	3.363	0.610	0	4.33
PSAT Score	99.40	19.83	40	160
SAT score	1049.63	204.85	400	1600
Zip Code/County Level Variables				
Broadband Availability - Dec. of Junior Year	0.3177	0.4656	0	1
Mean Adjusted Gross Income	76,946	53,646	13,370	2,791,601
Population	4,111	9,799	-	132,944
Unemployment Rate	5	2	-	31
Median Home Price	192,769	118,034	-	943,877
Application Outcomes				
Applied to a 4-Year College	0.7897	0.4075	0	1
Applied to at least 5 Colleges	0.4023	0.4904	0	1
Applied to a Selective College	0.5032	0.5000	0	1
Applied to a Top Liberal Arts College	0.0719	0.2583	0	1
Applied to a Match College	0.7071	0.4551	0	1
Applied to the State Flagship	0.2956	0.4563	0	1
Applied to an Out-of-State College	0.4920	0.4999	0	1

Notes: Sample is all SAT takers, high school graduating cohorts 2001 to 2008

Table 2: Effect of Broadband Availability in Junior Year on Application Outcomes

	SAT Score	Applied to 5+ Colleges	Applied to a 4-Year College	Applied to a Selective College	Applied to a Match College	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied to an Out-of-State College
Broadband	0.69798***	0.00327***	0.00357***	0.00361***	0.00292***	0.00172***	0.00109	0.00321***
	(0.16392)	(0.00084)	(0.00081)	(0.00103)	(0.00086)	(0.00040)	(0.00080)	(0.00083)
Observations	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008, except in the first column of the first panel for which the sample is students who took the PSAT test. Estimates are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics.

Table 3: Effect of Broadband Availability in Junior Year on Test-taking and Sample Composition

	Took the SAT (PSAT Takers)	Fraction of Zip Taking SAT	Fraction of Zip Taking PSAT	PSAT Score	Fraction Test-takers Under-represented Minority	Fraction Test-takers Parents Attended College
Broadband	0.00049	0.00015	-0.00074	-0.01789	0.00143	-0.00115
	(0.00081)	(0.00251)	(0.00466)	(0.03232)	(0.00110)	(0.00155)
Observations	12,556,552	223,415	223,415	7,452,302	153,066	150,388

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample in columns 1 includes students who took the PSAT, and in columns 2 includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Estimates in Columns 1-2 are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics. Columns 3- 6 are zip-code level regressions, in 3-4 the outcome is the fraction of the at-risk population in the zip code taking each exam, and in 5-6 the outcome is the fraction of all SAT-takers in the zip-code belonging to each demographic group. Estimates in columns 3-6 include zip-code and year fixed effects and time-varying zip code level controls.

Table 4: Effect of Broadband Availability in Junior Year on Application Outcomes, by Group

	SAT Score	Applied to 5+ Colleges	Applied to a 4 Year College	Applied to a Selective	Applied to a Match College	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied to an Out-of-State College
	Below median income zip							
Broadband	0.55126**	0.00054	-0.00048	0.00319*	-0.00109	0.00018	-0.00001	-0.00025
	(0.26502)	(0.00135)	(0.00135)	(0.00167)	(0.00150)	(0.00051)	(0.00124)	(0.00135)
Observations	3,065,569	3,065,569	3,065,569	3,065,569	3,065,569	3,065,569	3,065,569	3,065,569
	Above median income zip							
Broadband	0.22062	0.00230**	0.00369***	0.00326**	0.00348***	0.00169***	0.00201*	0.00286***
	(0.20831)	(0.00106)	(0.00099)	(0.00133)	(0.00105)	(0.00056)	(0.00104)	(0.00103)
Observations	4,381,979	4,381,979	4,381,979	4,381,979	4,381,979	4,381,979	4,381,979	4,381,979
	Parent's education is high school or less							
Broadband	1.18410***	0.00231	-0.00009	0.00200	-0.00160	0.00114*	-0.00601***	0.00206
	(0.37447)	(0.00192)	(0.00189)	(0.00206)	(0.00204)	(0.00061)	(0.00175)	(0.00186)
Observations	980,650	980,650	980,650	980,650	980,650	980,650	980,650	980,650
	Parent's education is some college or more							
Broadband	0.48472***	0.00288***	0.00313***	0.00384***	0.00291***	0.00165***	0.00178**	0.00261***
	(0.17461)	(0.00092)	(0.00082)	(0.00109)	(0.00089)	(0.00047)	(0.00089)	(0.00090)
Observations	5,202,386	5,202,386	5,202,386	5,202,386	5,202,386	5,202,386	5,202,386	5,202,386

Table 4 Cont: Effect of Broadband Availability in Junior Year on Application Outcomes, by Group

	SAT Score	Applied to 5+ Colleges	Applied to a 4-year College	Applied to a Selective College	Applied to a Match College	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied to an Out-of-State College
				Black/Hispanic				
Broadband	0.50816	0.00203	-0.00083	-0.00074	-0.00139	-0.00038	-0.00289*	0.00007
	(0.36226)	(0.00181)	(0.00168)	(0.00199)	(0.00190)	(0.00062)	(0.00167)	(0.00178)
Observations	1,504,678	1,504,678	1,504,678	1,504,678	1,504,678	1,504,678	1,504,678	1,504,678
				White/Other				
Broadband	0.46961***	0.00298***	0.00320***	0.00422***	0.00274***	0.00208***	0.00209**	0.00271***
	(0.16955)	(0.00089)	(0.00085)	(0.00109)	(0.00090)	(0.00045)	(0.00084)	(0.00088)
Observations	5,947,624	5,947,624	5,947,624	5,947,624	5,947,624	5,947,624	5,947,624	5,947,624
				Rural				
Broadband	0.60706	0.00286	-0.00176	-0.00034	-0.00333	-0.00038	0.00430**	0.00071
	(0.40139)	(0.00212)	(0.00219)	(0.00231)	(0.00228)	(0.00101)	(0.01264)	(0.00218)
Observations	887,587	887,587	887,587	887,587	887,587	887,587	887,587	887,587
				Urban				
Broadband	0.65184***	0.00302***	0.00393***	0.00425***	0.00345***	0.00196***	0.00077	0.00302***
	(0.17790)	(0.00091)	(0.00087)	(0.00113)	(0.00093)	(0.00044)	(0.00087)	(0.00089)
Observations	6,564,715	6,564,715	6,564,715	6,564,715	6,564,715	6,564,715	6,564,715	6,564,715

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Estimates are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics.

For Online Publication

Appendix: Measuring Broadband Availability

a. Construction of Measure

To construct our measure of zip code-level broadband Internet availability, we merged the zip code-level FCC data on the number of ISPs in a zip code, to zip code-level information on (1) land area and (2) population. The land area data are from the 2000 Census. The population data come from a combination of 2000 census data and SOI income tax data, which we used to construct zip-code-level population counts over time. We then constructed ratios of the number of ISPs to zip-code geographic size and population size. We designated each zip code either urban or rural based on the fraction of its population that the Census identified as residing in an urban area, where a zip code is urban if the Census identifies more than 50 percent of its residents this way.

We also collected national and sub-national trends in broadband Internet usage from the CPS/PEW data. The CPS data come from the CPS computer and Internet supplements/school enrollment surveys (August 2000, September 2001, October 2003, 2007, 2009). The PEW Survey is conducted at least annually and the month of the survey was recorded. We interpolated over missing survey months so the timing matched the FCC data (which were collected in June/December). We constructed national trends from both series, and separate urban and rural trends from the CPS series based on whether an individual lived in a metropolitan statistical area.

We then conduct the following exercise:

1) Identify a threshold for "coverage" (one provider per x square miles or one provider per y thousand people)
2) Construct an indicator which takes a value of one if threshold is in each zip code-year
3) Aggregate the indicator to a national level using zip code population weights, leaving a national time series of fraction of the population for which broadband is available
4) Estimate root mean squared error between CPS or PEW measure and #3

We then incrementally increase the threshold in step 1 and repeat steps 2 through 4. We tried the following combinations: one provider per 1-10,000 people (in intervals of 500 people), and one provider per 1-40 square miles (in intervals of one mile). We also allowed different thresholds for urban and rural zip codes (although we did not impose that they had to be different) by comparing to the CPS data. Finally, we minimized the root mean squared error

(RMSE) to identify a particular threshold as the preferred definition of zip code broadband availability. Per this metric, a rural zip code is "covered" when there is at least one provider per 12 square miles and an urban zip code is "covered" when there is at least one provider for every 2,700 people.

b. Measurement Error

Because we assign broadband availability under these assumptions, our measure will be subject to measurement error and our point estimates subject to attenuation bias. To gauge the potential importance of measurement error, we estimate the magnitude of the attenuation bias (e.g., Aigner, 1973) on our coefficients under several plausible assumptions for the misclassification rate (the fraction of zip codes which are misclassified) and the true broadband availability rate (nationally). Though both the misclassification rate and broadband availability rate are ultimately unknowns, we use the PEW broadband usage rate shown in figure 1, and the difference between our measure of availability and the broadband usage rate as proxies. We parameterize this difference in a few ways, including the min, max, and mean difference over the sample period. This leads to a scaling factor of 0.973-0.997. In other words, estimates are biased downwards by no more than 2.7 percent. Of course, it is possible (and likely) that the misclassification rate is somewhat higher than the rate we use, since some false positives will be offset by false negatives in the overall difference we calculate. That said, given that our estimate is already an "intent to treat," and the scaling factor in moving from availability to usage is an order of magnitude larger than this, this issue is somewhat second-order.

c. Validation Exercise: Broadband Availability as a Predictor of Teen Usage

To probe the efficacy of our measure, we can test the strength of the relationship between measures of broadband availability that can be derived from the FCC data and teen high-speed Internet usage. We obtain teen high-speed Internet usage rates from the CPS, which interviewed respondents about their broadband usage in August 2000, September 2001, October 2003, and October 2007. Since zip codes are not available for all CPS respondents, we aggregate availability to the state-level using population weights, which we then match to summarized high-speed Internet usage and demographic data from the CPS for 15-18 year old respondents. To mimic our main estimating equation, we control for similar characteristics, including year and state fixed effects, race/ethnicity, sex, family income, and parental education. We also include measures of population density, home prices, and unemployment rates at the state-level, again to emulate our main equation.

The first column in Appendix Table 1 indicates that the fraction of the state for which high-speed Internet is available by our measure is a positive and significant predictor of teen high-speed Internet usage. Columns 2-8 of Appendix Table 1 displays the results of the CPS analysis using alternative definitions of broadband availability. Of the additional measures we test, only "at least one provider per one square mile" is a statistically significant predictor of usage among teens. Moreover, of the measures we consider, our main measure, given the relatively small standard error, appears to admit the least noise. Of particular note are columns 4-7, which present estimates using the "at least one provider" in a zip code dummy variable, a linear measure of the number of providers, a linear measure of providers per person, and a linear measure of providers per square mile, respectively. None are significantly associated with high-speed Internet usage among teens (and the "at least one provider" measure is also wrong-signed).

d. Estimating Testing and Applications with Other Measures of Availability

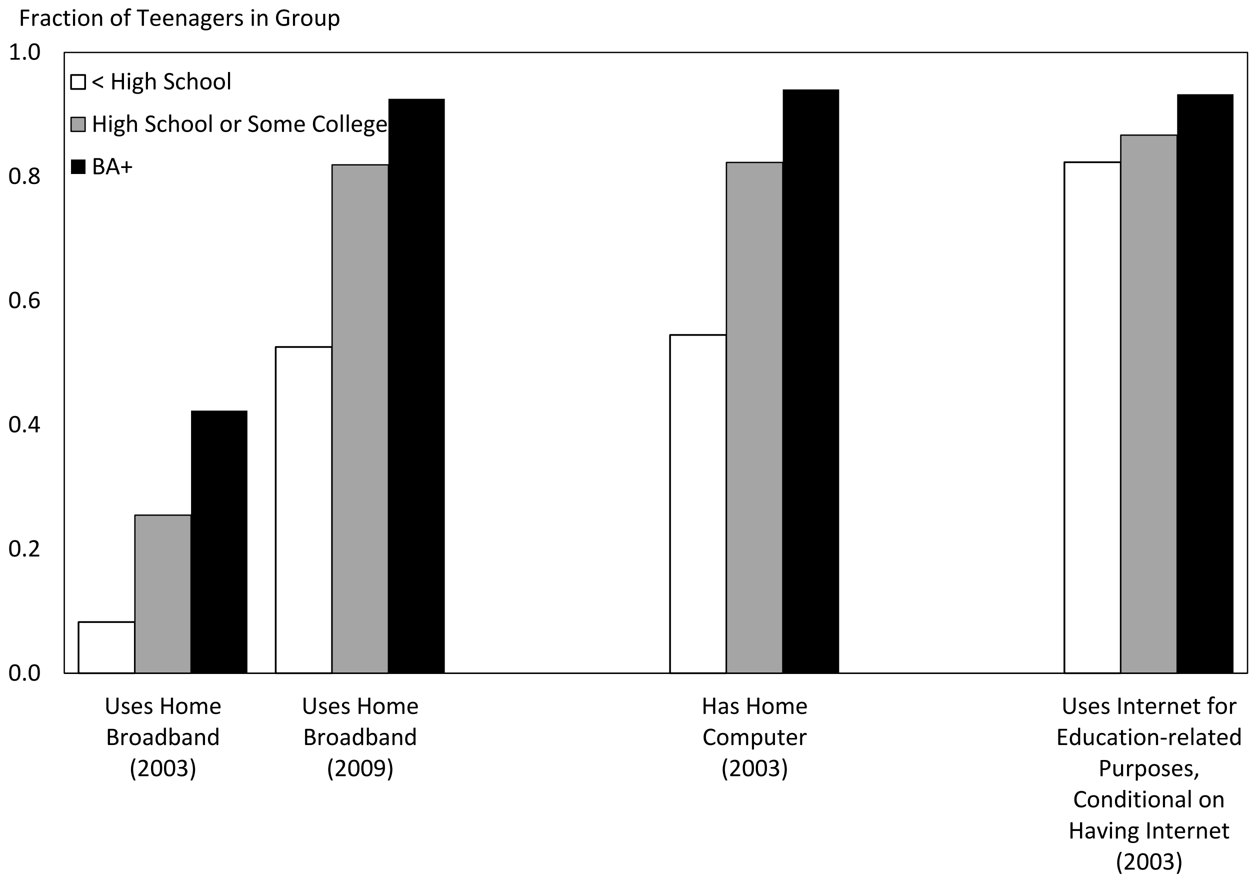
For completeness, we probe the sensitivity of our analysis to alternative measures of broadband availability derived from the FCC data. Because many of these measures are not predictive of teen usage, and all have a higher RMSE than our measure, measurement error is a significant concern. As highlighted in Bound, Brown, and Mathiowetz (2001) in the presence of severe measurement error, coefficients may not only be attenuated, but can even be wrong-signed. Moreover, interpretation is further complicated because some measures of availability are negatively correlated with usage. As such, we exercise strong caution in interpreting the coefficients obtained from the measures that do not fit the usage data.

We will begin our sensitivity analysis by considering slight modifications of our main variable by re-estimating the main results using various alternative thresholds for broadband availability, including one provider per 2000, 2500, 2700 and 3000 people in urban areas, and one provider per 10, 12, and 15 square miles in rural areas. The results are listed in Appendix Table 9, rows 1-12. We see that our main results are little changed, though standard errors tend to increase when we get further from our main specification (as expected). In the final two rows of Appendix Table 9, we consider the linear measures (providers per 1000 people and providers per square mileage measures). In these cases, results are not similar to main results, and sometimes wrong-signed, potentially indicative of measurement error.

Next, we consider the sensitivity of our results to our choice to use a population-based versus mileage-based measure for the urban and rural zip codes. Namely, we apply each of the population-based (i.e. “one provider per 2700 people”) and mileage-based (i.e., one provider per 12 square miles) measures to all zip codes and to the urban and rural subsamples separately in Appendix Table 10. The top panel of Appendix Table 10 demonstrates that the results using the population-based measure are qualitatively similar to those obtained using our main measure for the entire sample. The middle panel presents the results using the mileage-based measure (i.e. “one provider per 12 square miles”). In this case, the results are quite different from the results obtained using our main measure. However, this could be attributable to the fact that this measure fits the usage data poorly, with a RMSE of 2.79 for the entire sample (300 percent larger than the RMSE for our preferred measure). To further examine this, in the bottom panel of Appendix Table 10 we estimate the model using a mileage-based measure that better-fits the data, under the restriction that urban and rural zip codes are subject to the same threshold (one provider per one square mile, which has an RMSE of 1.39). In this case, results are more similar to both the preferred measure and the population-based measure. In sum, these results suggests the choice to use a population-based or square-mileage-based measure does not drive our results.

In Appendix Table 10, we also consider the alternative three measures separately within the urban and rural zip code subsamples. First, it is clear from the RMSEs that the separate population and mileage-based measures used in the preferred measure fit the urban and rural samples better than the alternative measures do, with RMSEs of 0.85 and 0.5, respectively. In addition, the mileage-based measure (i.e., “one provider per 12 square miles”) performs poorly for the urban sample, with an RMSE of 2.89, and leads to results that differ from the population-based results. However, as in the pooled results, the mileage-based measure that fits the usage data more closely (i.e., “one provider per one square mile”) leads to results that are more consistent with the main results. Finally, Appendix Table 10 confirms the null results for the rural subsample in Table 4. In each case, regardless of the fit of the measure, we see little to no effect of availability on our outcomes for the rural subsample. Overall, this subsample analysis indicates that the null (positive) result for the rural (urban) subsample is not driven by the choice to use a mileage-based (population-based) measure for that subsample.

Appendix Figure 1: Teenage Home Computer Access and Internet Use, by Mother's Education

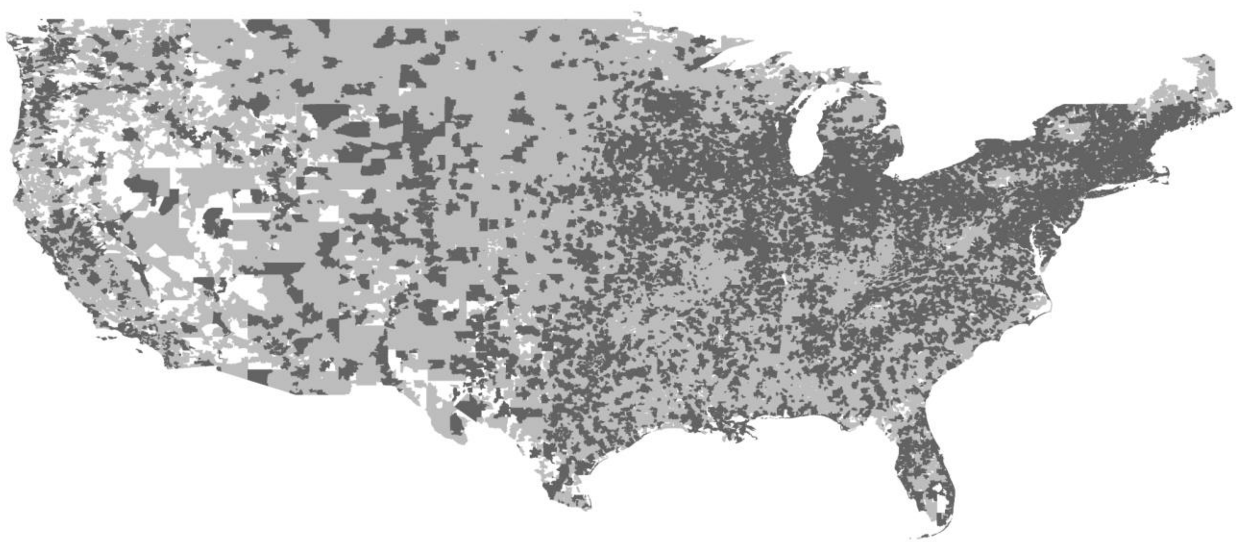


Notes: Displayed are access and use rates among teenagers (15-18 year olds), grouped according to their mother's level of education. Statistics are derived using the October 2003 and 2009 Current Population Survey School Enrollment Supplement. Computer access and detailed use questions were not asked in 2009.

Appendix Figure 2: Zip Code Broadband Access in 1999 and 2007

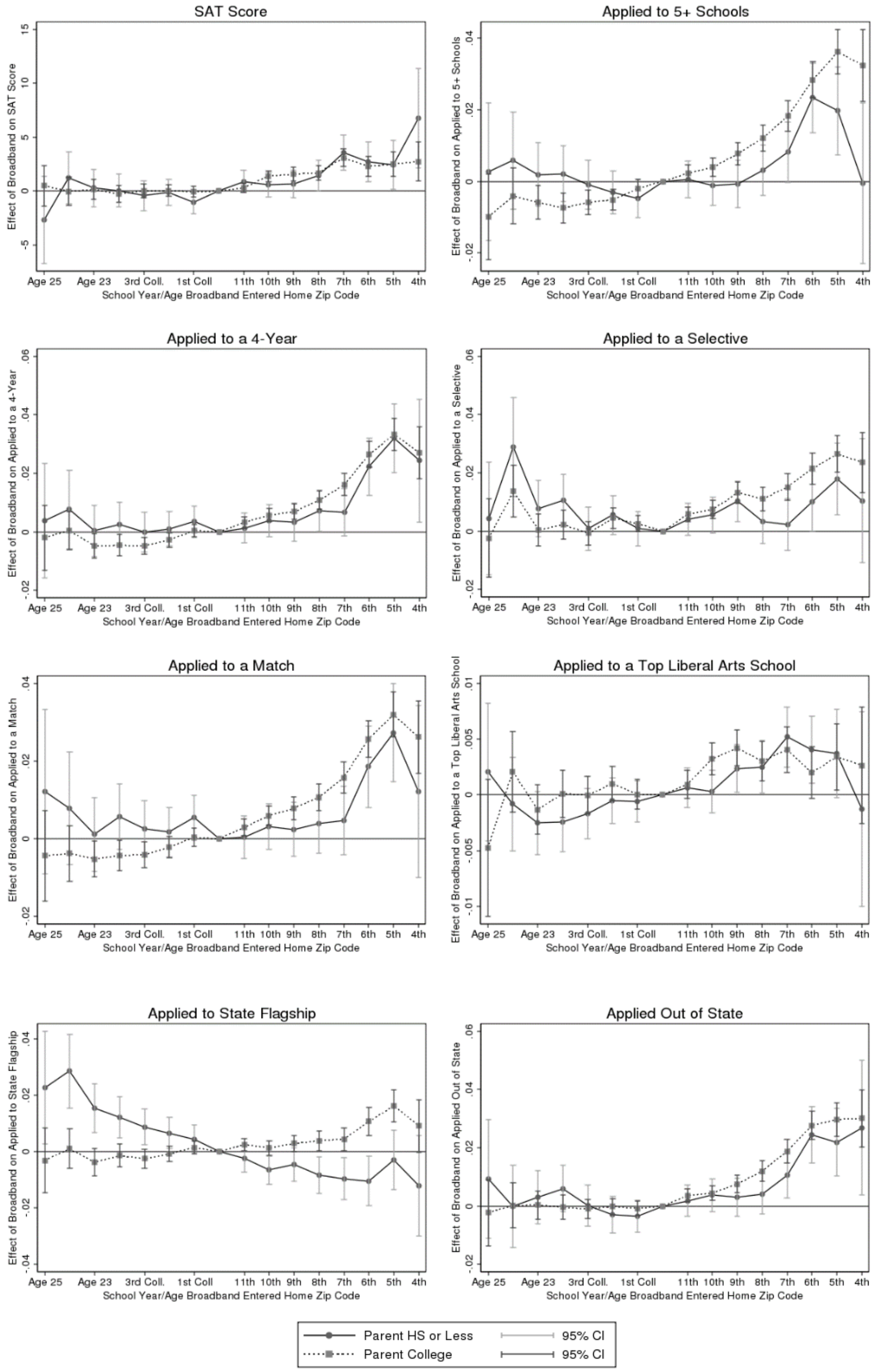
(A) December 1999

(B) December 2007



Notes: Displayed are zip code-level maps of Internet penetration. The light gray shading represents zip codes without coverage and the dark gray shading represents zip codes with coverage, where coverage is defined as our indicator for Internet availability, as described in the text. White areas, which mostly represent unpopulated areas like national parks and bodies of water, are zip codes for which we do not have information and are not included in the analyses.
Sources: FCC and Census.

Appendix Figure 3: Event Study Estimates of Effect of Broadband Availability, by Parent's Education



Notes: Each graph represents coefficients and 95 percent confidence intervals obtained from separate OLS regressions on subsamples of students split by parental education, whereby the specification is as in Table 2, except the single broadband availability dummy is replaced by a series of dummies for the years prior to and following the introduction of the broadband, which are labelled by year of schooling for ease of interpretation. The dummy for the year immediately preceding introduction (e.g., senior year) is excluded. Standard errors adjusted for clustering. Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Regressions include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics and a constant.

Appendix Table 1: Relationship between State-Level Broadband Availability and Teen High-Speed Internet Usage

	Main Measure	1 Provider Per 2700 People	1 Provider Per 12 Square Miles	1 Provider Per Square Mile	At Least 1 Provider	Number of Providers	Providers Per 1000 People	Providers Per Square Mile
Fraction of State with Broadband	0.1447582**	0.101977	0.0428876	0.2231739*	-0.22539	0.0044884	0.043125	0.013305
	(0.06821)	(0.06945)	(0.11755)	(0.12215)	(0.19505)	(0.00555)	(0.035077)	(0.031938)
N	204	204	204	204	204	204	204	204

Notes: Standard errors adjusted for clustering at the state-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes 15-18 year old CPS respondents in 2000, 2001, 2003 and 2007. Estimates are from separate regressions that include state and year fixed effects, as well as controls for the fraction of teens in the state belonging to each gender, race/ethnicity, family income, and parental education group. Regressions are weighted by the sum of the CPS-provided weight in each cell. Also includes controls for time-varying state characteristics. Fraction of state with broadband is calculated by collapsing the zip code-level measures of availability denoted in the column headings, weighting by zip code population.

Appendix Table 2: Effect of Broadband Availability in Junior Year on Application Outcomes, Omitting PSAT Control

	SAT Score	Applied to 5+ Colleges	Applied to a 4-Year College	Applied to a Selective College	Applied to a Match College	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied to an Out-of-State College
Broadband	0.55200*	0.00315***	0.00348***	0.00344***	0.00287***	0.00170***	0.00098	0.00310***
	(0.31590)	(0.00085)	(0.00082)	(0.00106)	(0.00086)	(0.00041)	(0.00080)	(0.00085)
Observations	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Estimates are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, and time-varying zip code characteristics.

Appendix Table 3: Effect of Broadband Availability in Junior Year on Application Outcomes, Controlling for Self-Reported Parental Education and Parental Income

	SAT Score	Applied to at least 5 Colleges	Applied to a 4-Year College	Applied to a Selective College	Applied to a Match College	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied to an Out-of-State College
Broadband	0.64913***	0.00289***	0.00302***	0.00325***	0.00242***	0.00172***	0.00095	0.00271***
	(0.16265)	(0.00083)	(0.00080)	(0.00103)	(0.00086)	(0.00040)	(0.00080)	(0.00082)
Observations	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Estimates are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics.

Appendix Table 4: Effect of Broadband Availability in Junior Year on Application Outcomes, SAT States Only

	SAT Score	Applied to 5+ Colleges	Applied to a 4-Year College	Applied to a Selective College	Applied to a Match College	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied to an Out-of-State College
Broadband	0.70417***	0.00348***	0.00430***	0.00200*	0.00091	0.00148***	0.00012	0.00408***
	(0.18552)	(0.00094)	(0.00092)	(0.00116)	(0.00083)	(0.00044)	(0.00086)	(0.00092)
Observations	6,080,456	6,080,456	6,080,456	6,080,456	6,080,456	6,080,456	6,080,456	6,080,456

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Estimates are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics. SAT states are defined as states in which the SAT is the dominant exam.

Appendix Table 5: Effect of Broadband Availability on ACT Test Centers

	No Controls, Excluding Missing Zip Code-Years	Controls, Excluding Missing Zip Code-Years	No Controls, Missing Zip Code-Years are Zeroes	Controls, Missing Zip Code-Years are Years
Broadband	-0.011	-0.010	0.001	-0.000
	(0.007)	(0.007)	(0.001)	(0.001)
Observations	52,488	50,001	287,217	251,944

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes U.S. zip codes observed between 1999 and 2007. Estimates are from separate regressions that include zip code and year fixed effects. In columns 2 and 4, regressions include time-varying zip code controls for population density, unemployment rates, and home prices. Missing zip code-years refer to observations for which there are no ACT test centers.

Appendix Table 6: Effect of Broadband Availability on Zip Code Income

	Mean Adjusted Gross Income No Controls	Mean Adjusted Gross Income Controls	Mean Wage and Salary Income No Controls	Mean Wage and Salary Income Controls
Broadband	-441.91 (330.61)	-217.22* (114.52)	21.99 (69.53)	-111.66* (63.68)
Observations	284,816	251,944	284,816	251,944

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes U.S. zip codes observed between 1999 and 2007. Estimates are from separate regressions that include zip code and year fixed effects. In columns 2 and 4, regressions include time-varying zip code controls for population density, unemployment rates, and home prices.

Appendix Table 7: Effect of Broadband Availability in Senior Year on Application Outcomes

	SAT Score	Applied to 5+ Colleges	Applied to a 4-Year College	Applied to a Selective College	Applied to a Match College	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied to an Out-of-State College
Broadband	-0.00338	0.00336***	0.00182**	-0.00036	0.00298***	0.00001	0.00138*	0.00010
	(0.17111)	(0.00085)	(0.00083)	(0.00101)	(0.00076)	(0.00039)	(0.00079)	(0.00085)
Observations	7,452,121	7,452,121	7,452,121	7,452,121	7,452,121	7,452,121	7,452,121	7,452,121

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Estimates are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics.

Appendix Table 8: Effect of Broadband Availability in Junior Year on Application Outcomes, by PSAT Score Group

	SAT Score	Applied to 5+ Colleges	Applied to a 4-Year College	Applied to a Selective College	Applied to a Match College	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied to an Out-of-State College
	Top Quartile of PSAT Scorers							
Broadband	0.26413	-0.00006	0.00103	0.00018	0.00069	0.00159*	0.00318**	0.00078
	(0.21187)	(0.00118)	(0.00081)	(0.00118)	(0.00108)	(0.00091)	(0.00127)	(0.00110)
Observations	2,386,354	2,386,354	2,386,354	2,386,354	2,386,354	2,386,354	2,386,354	2,386,354
	Bottom Three Quartiles of PSAT Scorers							
Broadband	0.73333***	0.00447***	0.00412***	0.00491***	0.00368***	0.00076***	-0.00035	0.00415***
	(0.19580)	(0.00098)	(0.00104)	(0.00121)	(0.00108)	(0.00029)	(0.00090)	(0.00098)
Observations	5,065,948	5,065,948	5,065,948	5,065,948	5,065,948	5,065,948	5,065,948	5,065,948

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Estimates are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics. SAT states are defined as states in which the SAT is the dominant exam.

Appendix Table 9: Effects of Alternative Measures of Zip Code-Level Broadband Availability on Application Outcomes

	SAT Score	Applied to 5+ Colleges	Applied to 4-year	Applied to a Selective	Applied to a Match	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied Out of State
				Broadband = 2000 people, 10 square miles				
Broadband	0.64372***	0.00373***	0.00441***	0.00557***	0.00418***	0.00210***	0.00335***	0.00437***
	(0.17948)	(0.00100)	(0.00098)	(0.00115)	(0.00102)	(0.00048)	(0.00091)	(0.00098)
				Broadband = 2500 people, 10 square miles				
Broadband	0.79577***	0.00481***	0.00444***	0.00499***	0.00366***	0.00227***	0.00184**	0.00366***
	(0.16818)	(0.00088)	(0.00086)	(0.00104)	(0.00091)	(0.00041)	(0.00081)	(0.00087)
				Broadband = 2700 people, 10 square miles				
Broadband	0.79577***	0.00481***	0.00444***	0.00499***	0.00366***	0.00227***	0.00184**	0.00366***
	(0.16818)	(0.00088)	(0.00086)	(0.00104)	(0.00091)	(0.00041)	(0.00081)	(0.00087)
				Broadband = 3000 people, 10 square miles				
Broadband	0.66823***	0.00379***	0.00340***	0.00314***	0.00249***	0.00191***	0.00079	0.00272***
	(0.16236)	(0.00082)	(0.00081)	(0.00099)	(0.00086)	(0.00039)	(0.00077)	(0.00082)
				Broadband = 2000 people, 12 square miles				
Broadband	0.66666***	0.00289***	0.00396***	0.00502***	0.00356***	0.00180***	0.00296***	0.00388***
	(0.17788)	(0.00099)	(0.00097)	(0.00114)	(0.00101)	(0.00047)	(0.00091)	(0.00097)
				Broadband = 2500 people, 12 square miles				
Broadband	0.81513***	0.00415***	0.00410***	0.00458***	0.00319***	0.00204***	0.00155*	0.00330***
	(0.16724)	(0.00087)	(0.00086)	(0.00104)	(0.00091)	(0.00041)	(0.00081)	(0.00086)
				Broadband = 2700 people, 12 square miles				
Broadband	0.69798***	0.00327***	0.00357***	0.00361***	0.00292***	0.00172***	0.00109	0.00321***
	(0.16392)	(0.00084)	(0.00081)	(0.00103)	(0.00086)	(0.00040)	(0.00080)	(0.00083)
				Broadband = 3000 people, 12 square miles				
Broadband	0.68832***	0.00321***	0.00312***	0.00280***	0.00209**	0.00172***	0.00054	0.00241***
	(0.16174)	(0.00082)	(0.00080)	(0.00099)	(0.00085)	(0.00039)	(0.00077)	(0.00082)

				Broadband = 2000 people, 15 square miles				
Broadband	0.67808***	0.00240**	0.00354***	0.00474***	0.00312***	0.00171***	0.00292***	0.00370***
	(0.17638)	(0.00097)	(0.00096)	(0.00113)	(0.00100)	(0.00047)	(0.00090)	(0.00096)
				Broadband = 2500 people, 15 square miles				
Broadband	0.82533***	0.00375***	0.00378***	0.00438***	0.00285***	0.00198***	0.00154*	0.00317***
	(0.16645)	(0.00087)	(0.00085)	(0.00103)	(0.00090)	(0.00041)	(0.00080)	(0.00086)
				Broadband = 2700 people, 15 square miles				
Broadband	0.82533***	0.00375***	0.00378***	0.00438***	0.00285***	0.00198***	0.00154*	0.00317***
	(0.16645)	(0.00087)	(0.00085)	(0.00103)	(0.00090)	(0.00041)	(0.00080)	(0.00086)
				Broadband = 3000 people, 15 square miles				
Broadband	0.70038***	0.00287***	0.00285***	0.00264***	0.00180**	0.00167***	0.00055	0.00231***
	(0.16131)	(0.00081)	(0.00080)	(0.00098)	(0.00085)	(0.00038)	(0.00077)	(0.00081)
				Broadband = providers per 1,000 people				
Broadband	0.00798**	0.00002	-0.00006***	0.00001	-0.00005***	0.00002	-0.00001	-0.00008***
	(0.00384)	(0.00002)	(0.00002)	(0.00003)	(0.00001)	(0.00002)	(0.00002)	(0.00001)
				Broadband = providers per square mile				
Broadband	-0.08578	-0.00003	0.00041	0.00114***	0.00077**	0.00007	0.00046*	-0.00015
	(0.06705)	(0.00034)	(0.00034)	(0.00034)	(0.00035)	(0.00021)	(0.00025)	(0.00033)

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Estimates are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics. RMSE refers to the fit of the measure with national trends in Internet usage from PEW.

Appendix Table 10: Effects of Alternative Measures of Zip Code-Level Broadband Availability on Application Outcomes, Urban and Rural Subsamples

	SAT Score	Applied to at least 5 Colleges	Applied to a 4-Year College	Applied to a Selective College	Applied to a Match College	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied to an Out-of-State College
Broadband=1 Provider per 2700 People								
All (RMSE=1.34)								
Broadband	0.53355***	0.00134	0.00276***	0.00342***	0.00228***	0.00139***	0.00073	0.00236***
	(0.16185)	(0.00082)	(0.00079)	(0.00101)	(0.00085)	(0.00039)	(0.00078)	(0.00081)
Urban (RMSE=0.85)								
Broadband	0.65184***	0.00302***	0.00393***	0.00425***	0.00345***	0.00196***	0.00077	0.00302***
	(0.17790)	(0.00091)	(0.00087)	(0.00113)	(0.00093)	(0.00044)	(0.00087)	(0.00089)
Rural (RMSE=2.36)								
Broadband	0.01110	-0.00390**	-0.00298	-0.00102	-0.00343*	-0.00091	-0.00061	-0.00167
	(0.35885)	(0.00180)	(0.00182)	(0.00200)	(0.00194)	(0.00084)	(0.00165)	(0.00189)
Broadband=1 Provider per 12 Square Miles								
All (RMSE=2.79)								
Broadband	-0.22949	-0.00351***	-0.00526***	-0.00114	-0.00586***	-0.00148***	0.00187*	-0.00281**
	(0.21385)	(0.00110)	(0.00112)	(0.00133)	(0.00118)	(0.00050)	(0.00110)	(0.00113)
Urban (RMSE=2.89)								
Broadband	-0.47421*	-0.00468***	-0.00680***	-0.00067	-0.00666***	-0.00156***	0.00143	-0.00469***

	(0.25033)	(0.00128)	(0.00130)	(0.00158)	(0.00138)	(0.00058)	(0.00130)	(0.00131)
Rural (RMSE=0.50)								
Broadband	0.60706	0.00286	-0.00176	-0.00034	-0.00333	-0.00038	0.00430**	0.00071
	(0.40139)	(0.00212)	(0.00219)	(0.00231)	(0.00228)	(0.00101)	(0.00192)	(0.00218)
Broadband=1 Provider per 1 Square Mile								
All (RMSE=1.39)								
Broadband	-0.16870	0.00359***	0.00226**	0.00299**	0.00209*	0.00056	-0.00151	-0.00063
	(0.22246)	(0.00118)	(0.00115)	(0.00140)	(0.00120)	(0.00054)	(0.00114)	(0.00114)
Urban (RMSE=1.45)	-0.29045	0.00255**	0.00214*	0.00330**	0.00189	0.0002	-0.00111	-0.00081
Broadband	(0.22518)	-0.00119	-0.00117	-0.00143	-0.00122	-0.00055	-0.00116	-0.00115
Rural (RMSE=2.06)								
Broadband	-0.56848	0.01195	0.00963	0.02461*	0.01992	-0.00385	-0.00064	0.02048
	(3.04643)	(0.01556)	(0.01351)	(0.01328)	(0.01346)	(0.00695)	(0.01141)	(0.01359)

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Estimates are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics. RMSE refers to the fit of the measure with separate urban and rural national trends in usage from the CPS.

Appendix Table 11: Effect of Broadband Availability in Junior Year on Application Outcomes, Controlling for SAT Score

	Applied 5+ College	Applied to a 4-Year College	Applied to Selective College	Applied to a Match College	Applied to a Top Liberal Arts College	Applied to the State Flagship	Applied to an Out-of-State College
Broadband	0.00286***	0.00325***	0.00313***	0.00292***	0.00154***	0.00090	0.00280***
	(0.00083)	(0.00081)	(0.00102)	(0.00086)	(0.00040)	(0.00080)	(0.00082)
Percent of initial coefficient	*80*	*99*	*87*	*100*	*90*	*83*	*87*
Observations	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302	7,452,302

Notes: Standard errors adjusted for clustering at the zip code-level are in parentheses (* p<.10 ** p<.05 *** p<.01). Sample includes students who took both the PSAT and SAT tests in the high school graduating cohorts of 2001 to 2008. Estimates are from separate regressions that include zip code and year fixed effects, as well as controls for gender, race, high school GPA, PSAT verbal and math sections, and time-varying zip code characteristics.