Entry Threat, Entry Delay, and Internet Speed: The Timing of the U.S. Broadband Rollout*

Kyle Wilson[1], Mo Xiao[2], and Peter F. Orazem[3]

[1]Pomona College

[2]University of Arizona

[3]Iowa State University

November 3, 2020

Abstract

In a rapidly growing industry, potential entrants strategically choose which local markets to enter. Facing the threat of additional entrants, a potential entrant may lower its expectation of future profits and delay entry into a local market, or it may accelerate entry due to preemptive motives. Using the evolution of local market structures of broadband Internet service providers from 1999 to 2007, we find that the former effect dominates the latter after allowing for spatial correlation across markets and accounting for endogenenous market structure. On average, it takes two years longer for threatened markets to receive their first broadband entrant. Moreover, this entry delay has long-run negative implications for the divergence of the U.S. broadband infrastructure: one year of entry delay translates into an 11% decrease in average present-day download speeds.

Keywords: Entry, Entry Threat, Endogenous Market Structure, Internet Service Providers, Internet Speed

*Kyle Wilson: Department of Economics, Pomona College, Claremont, CA 91711, kyle.wilson@pomona.edu; Mo Xiao, Department of Economics, University of Arizona, Tucson, AZ 85721, mxiao@eller.arizona.edu; Peter F. Orazem, Department of Economics, Iowa State University, Ames, IA 50011, pfo@iastate.edu. We thank Dan Ackerberg, Ashley Langer, Gautam Gowrisankaran, Miguel Villas-Boas, and participants of various seminars and conferences. Mo Xiao thanks the NET Institute (www.NETinst.org) and the Eller College small research grant for financial support. All errors are ours.

1 Introduction

The U.S. broadband industry has been plagued by the problem of "digital divide." The Federal Communications Commission (FCC) reported in 2015 that 53% of rural Americans but only 8% of urban Americans lack access to high-speed internet.[1] This inequality is often attributed to socio-economic differences in populations and cost differences across terrains, as broadband providers prefer highly-populated, more affluent markets with potential for growth.[2] Meanwhile, if multiple providers enter the same location, competition will erode profit. Broadband providers then face the task of balancing the underlying demand and cost shifters with competition intensity, leading to highly strategic local market entry decisions. Since the advent of the industry, internet service providers (ISPs) have sought to optimally locate and expand while responding to their rivals' attempts to do the same, and these interactions have given rise to the competitive landscape we face today.

We examine the early U.S. broadband industry and its aftermath under the lens of "path dependence," the dependence of economic outcomes on the path of previous outcomes, rather than simply on current conditions (Arthur, 1989; David, 1985). In the early 2000s, the internet industry is in its infancy. Potential entrants, whether they are telecommunications veterans or new start-ups, roll out their network gradually and strategically. Looking for the next market to enter, a potential entrant must anticipate the actions of potential rivals in the marketplace. Entry timing depends on the countervailing forces of expected competition, preemption incentives, and cost differences. A potential entrant may delay entry, anticipating that rival entry lowers the expected profitability of a local market. Or, it may accelerate entry to preempt competitors, especially if an early mover can make an irreversible investment in building capacity, if consumers have switching costs or inertia, or if the new entrants face financial constraints.[3] A firm may also simply spill over to a neighboring

[1]The FCC applied a benchmark of 25 megabits per second (Mbps) of download speed and 3 Mbps of upload speed in calculating these statistics.

[2]Greenstein (2020), for instance, writes "Today approximately 10 percent of the US population does not use the internet (Anderson et al. 2019). Some of that non-adoption is linked to demographic features of users, such as older age, low income, and less education. But an important factor is the location of a household, namely, in a rural or low-density location. While 97 percent of the land in the United States is rural, according to the Census Bureau, 19 percent of the population lives in rural locations — that is, areas with sparse residential housing. Cutting-edge internet infrastructure tends not to be available in low-density regions. In some of these locations, even internet infrastructure with older technology may not be available (for additional discussion, see Forman et al. 2018)."

[3]Theory has provided numerous rationales for preemptive incentives, including the capacity commitment story as in Dixit (1979), the strategic learning-by-doing as in Spence (1981), the cost-signaling as in Milgrom and Roberts (1982), and the switching costs story as in Klemperer (1987) and Farrell and Klemperer (2004). We think capacity and

locale if nearby network facilities give rise to a cost advantage. In this paper, we establish a setting that offers the opportunity to identify markets which face a credible threat of entry, and we estimate the impact of such a threat on firm actions. We discover a strong "entry delay" effect, suggesting that in the short-run, the expected competition effect outweighs the combined effects of preemption and cost advantages. More importantly, we find that this delayed entry into a local market early in the evolution of the industry has a direct effect on the market in the long run, as subscribers there face slower download speeds more than a decade later.

We use the FCC's bi-annual Form 477 data from 1999 to 2007 on the evolution of local market structures of facilities-based broadband providers at the zip code level. This data does not contain information on firm identities, which limits our ability to analyze firm-level heterogeneity along the lines of Mazzeo (2002). Nonetheless, the FCC's Form 477 data is the most complete data available for the early period of the U.S. broadband industry, and allows us to generate interesting insights about firms' entry strategies. In examining the data, we define a potential entrant to be *threatened* when a neighboring market houses at least one of its rivals. We then use the timing of entry into the market to understand how a firm's entry strategy is affected by the threat of future competition. Furthermore, we investigate whether the timing of entry has long-run implications for the current state of the industry. Specifically, we estimate the extent to which delayed initial entry into a market affects the download speeds available in 2013 using detailed speed data from the National Telecommunications and Information Administration's (NTIA) National Broadband Map.

The empirical strategy we adopt is as follows: we construct a latent variable representation of a market's profitability, which depends on observable market characteristics that affect demand and costs, and critically, whether or not a market is threatened by future entry. We then estimate the effect of entry threat on the probability of entry into a market, and on the length of time elapsed until a market is eventually entered. Lastly, we estimate the effect of the number of years of entry delay on the download speeds available in a market more than a decade later.

Within this framework, we recognize that whether or not a market is threatened is determined by previous entry into neighboring markets. Therefore, if characteristics of neighboring markets

consumers' switching costs are most relevant in the broadband industry. Beyond this, new entrants in this industry were often much smaller firms that needed funding: a bank or an investor needed to be persuaded of the profitability of the new entrants' action, and an early mover's preemptive actions can invalidate such proof of profitability.

which induced entry there are correlated with unobservable characteristics in the market of interest, then entry threat is endogenous. Additionally, if, as we claim, a prospective entrant into a market considers the market structure of neighboring markets when making its entry decision, then it must be true that incumbents in neighboring markets have engaged in a parallel exercise that incorporates their expectations about entry into the market of interest. This selection of entry threat further aggravates the endogeneity problem of our entry threat indicator.

We address these problems with two remedies. First, we allow markets close to one another to have spatially correlated error terms. Second, we instrument for entry threat using the market attributes of nearby markets. Specifically, these nearby markets are the second order neighbors (that is, neighbors of neighbors) of the market of interest.[4] These attributes directly affect entry into the second order neighbors themselves, which then allows them to more easily enter neighboring markets and therefore, by definition, affects the threat of entry into the market of interest. At the same time, these attributes can be considered exogenous to a potential entrant's decision to enter the market of interest, as long as attributes of markets which are two or more zip codes away do not directly affect entry into the market of interest.[5] In other words, we assume that firms do not enter a market *because*, at some future date, they plan to enter a market two zip codes away. During the years we study, the industry was in its infancy, leaving firms with tremendous uncertainty over firm turnover and future profitability. It is therefore reasonable to assume that firms do not have a perfectly forward-looking plan for broadband rollout.[6] Using these instruments, we estimate parameters of the model, including the extent of spatial correlation, in a generalized method of moments (GMM) framework developed by Pinkse and Slade (1998).

We find evidence that potential entrants place significant consideration on the possibility of future competition when making their entry decision. First, we demonstrate that our measure of entry threat is, in fact, credible: threatened markets are 8 percentage points more likely to be entered in the long run. However, in the short run, a market which is threatened by the entry of competitors is 20 percentage points less likely to be entered than its unthreatened counterpart. This is a substantial effect, as it represents the net of three separate effects: a threatened market

[4]We do not include second order neighbors which are also direct neighbors when constructing this set.

[5]These instruments are also in the spirit of Pearcy and Savage (2015), who use regional measures of costs and capacity to measure potential competition in bilateral international telecommunications markets.

[6]In principle, we only require that firms are sufficiently myopic. We could allow firms to consider any number of steps ahead and construct the appropriate set of instruments.

may be less likely to be entered because firms are unlikely to maintain market power; but on the other hand, it may be more likely to be entered due to preemptive motives; and finally, a threatened market, by definition, has firms nearby that can spill over due to economies of scale. Following this, we show that an open threatened market will, on average, wait about two years longer before being entered by its first broadband provider.

This delayed entry turns out to have important implications for the long-run development of broadband infrastructure. We find that for each additional year that initial entry is delayed, the download speeds available in 2013 fall by 11%. *A priori*, the expected direction of this effect is ambiguous. One might imagine that markets which experience delayed initial entry would receive the latest technology and therefore would have access to faster speeds today. However, nearly all markets in the U.S. had received their initial entrant by 2007, and the prevailing speeds of that time do not even meet the FCC's current definition of broadband. Instead, we argue and provide evidence that markets which experience delayed entry take longer to become competitive and therefore lack the sustained competitive pressure necessary to spur investment in quality improvements.

These findings fill in a relatively sparse empirical literature on the effects of entry threat on firm strategies. The theoretical literature is well developed, and has shown that firms facing the threat of a rival's entry have incentives to act preemptively. Notable studies have developed different mechanisms to explain why incumbents do not delay costly competitive actions until actual entry happens. For example, Spence (1981) shows that firms have incentives to enter early and invest to deter competition, and that early entrant advantages are magnified by learning. Milgrom and Roberts (1982) stress the importance of reputation and asymmetric information in deterring entry. Klemperer (1987) showed that firms adopt pricing strategies which take advantage of consumers' cost of switching to a new entrant.

Because identifying entry threat is difficult, there are only a handful of empirical studies on the effects of entry threat on incumbents' behavior. Ellison and Ellison (2011) propose that an incumbent firm's investment may be non-monotone in "market attractiveness" if investments are undertaken to preempt rivals from entering, because preemption is impossible in the most attractive markets. They then find evidence of this behavior in the pharmaceutical industry. Dafny (2005) similarly finds evidence of strategic investment behavior to deter potential entrants in the hospital industry. Goolsbee and Syverson (2008) find that incumbent airlines cut prices dramatically in response to the potential entry of Southwest, and confirm that this action was motivated by preemption rather than accommodation of their future competitor. Prince and Simon (2012) extends Goolsbee and Syverson (2008) to the non-price dimension. They find incumbents' on-time performance actually worsens in response to Southwest's entry threat and actual entry, and attribute this counterintuitive result to incumbents' incentives to differentiate from the high-performing potential entrant.

Most recently, Shapiro (2016) shows that pharmaceutical firms strategically delay the introduction of new versions of drugs until just before patent expiration of the original drug so their reformulated drugs compete directly with newly-entered generic drugs instead of the incumbents' own original version of the same drug;[7] Wen and Zhu (2019) find that when Google threatens to implement a new native app in its Android operating system, developers of similar existing apps shift innovation towards improving their unthreatened apps and developing new apps. These empirical findings point to a common theme: the incumbents' actions when facing entry threat depend on whether the entry threat can be deterred. If entry threat cannot be deterred, the incumbent faces a lowered expected profit stream as if entry threat is actual entry.

This insight carries over when we study potential entrants facing entry threat from other potential entrants. Potential entrants, like incumbents, may be incentivized to quickly enter a threatened market in an effort to preempt their rivals. But, they may also be motivated to delay entry, fearing that the market may become competitive and therefore less profitable in the future. Which of these effects dominates is therefore an empirical question, depending on the strength of the entry threat. Seamans (2012) shows that incumbent cable television providers, acting as potential entrants in this case, were more likely to begin offering internet service in areas where the local government might provide internet service in the future. Conversely, in our setting, we find that the effect of expected future competition dominates. We provide evidence that even after controlling for factors that influence demand and costs, internet service providers delay entry into markets that are threatened by future entry of rivals.

More broadly, our work relates to the literature on the effects of market structure on quality

[7]Huckfeldt and Knittel (2012) find similar strategic delay for pharmaceutical incumbents for more categories of drugs. The entry delay documented in both studies is different from the entry delay we discover: pharmaceuticals delay entry of reformulated drug to better compete with expected new entrants, while broadband providers delay entry into some local markets due to the anticipation of dissipated profits brought upon by new entrants.

provision. Theory predictions on the effect of market structure on quality provision is less clean cut than that on prices. Matsa (2011) shows that supermarkets facing more intense competition have better product availability. Mazzeo (2003) finds average flight delays are shorter in more competitive markets; Prince and Simon (2017), however, find airline mergers have negligible impacts on airlines' on-time performance measures. More relevant to broadband, Wallsten and Mallahan (2013) and Molnar and Savage (2017) find that competition in wireline ISPs increases wireline Internet speed. Our work adds a new angle to this line of work: we focus the *past* strategic actions of firms, which translate into meaningful differences in the quality of internet access available more than a decade later. Our results hold even after controlling for current market structure.[8] In other words, firms' past actions, which determine past market structure, have independent effects on both the market structure and firm performance we face today.

2 The Evolution of Broadband Internet and the Digital Divide

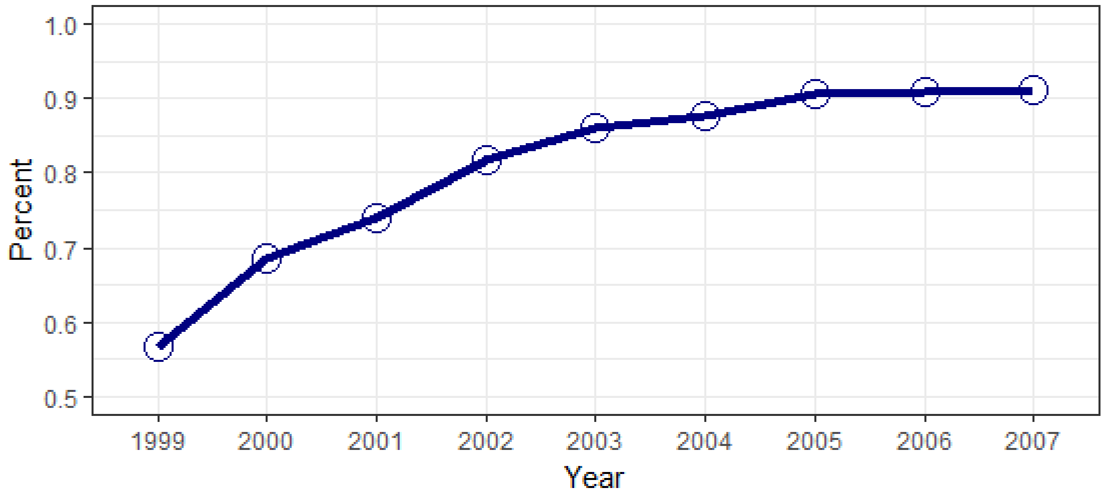
The 1996 Telecommunications Act was passed with the goal of encouraging competition in local telecommunications markets, largely by removing barriers to entry and by requiring incumbent firms to lease their lines to competitors. The act has been at least partially successful in achieving this goal, as several papers have investigated strategic interaction among competitors and the welfare effects of new entry into local telecommunications markets. Greenstein and Mazzeo (2006) found that telephone firms differentiate themselves strategically when entering markets; Economides, Seim, and Viard (2006) show that households in the state of New York benefit significantly from this resulting product differentiation.

In the late 1990s, household internet access underwent a transition from dial-up (or narrowband) access, capable of delivering speeds of up to 56 Kilobits per second (Kbps),[9] to broadband access capable of transmitting data at much faster speeds. At its most basic, broadband internet is characterized simply by the use of a wide band of frequencies to transmit data. However, as internet speeds have evolved, the FCC has adopted benchmarks to define broadband speeds. In 1996, the FCC first defined broadband speeds to be 200 Kbps of download and upload speed; in 2010, this was increased to 4 Mbps of download speed and 1 Mbps of upload speed; and in 2015 this

[8]Our results on current market structure corroborate the findings in Molnar and Savage (2017).

[9]1 Mbps = 1,000 Kbps

Figure 1: Fraction of Zip Codes with Broadband Internet



Source: FCC Form 477

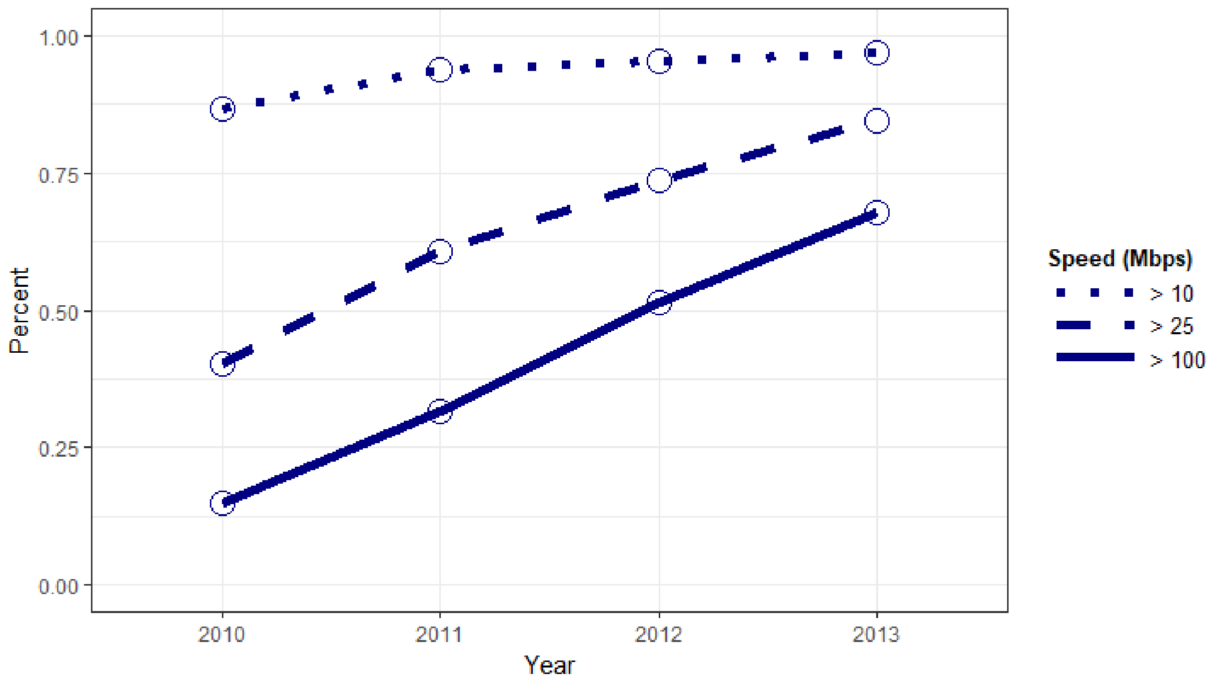
was further increased to 25 Mbps of download speed and 3 Mbps of upload speed. For the sake of clarity, and in keeping with our data, we will apply the most basic definition of broadband, labeling all firms utilizing a wide band of frequencies in their data transmission as broadband providers. As shown in figure 1, the fraction of zip codes with access to at least one broadband provider rose from 57% to 91% between 1999 and 2007, with most of this growth occurring by 2003. Furthermore, the 91% of zip codes which had broadband in 2007 accounted for over 99% of the U.S. population. As a result, the question of interest is not if, but rather when, individuals obtained access to broadband internet.

2.1 Firm Types and Quality Improvements

In the United States, internet service is provided predominantly by two types of firms, cable television and telephone companies. Cable firms provide broadband service using hybrid fiber-coaxial cable networks, and telephone companies provide service over digital subscriber lines (DSL). Both types of firms provide internet service primarily using the lines put in place for their preexisting cable television and telephone services, retrofitted to allow for the bilateral transfer of data required for internet usage.

Since its inception, the speeds of household broadband internet connections have continually improved. Cable firms have improved speeds by adopting common standards for data transmission,

Figure 2: Fraction of Zip Codes with Various Speeds



Source: National Broadband Map

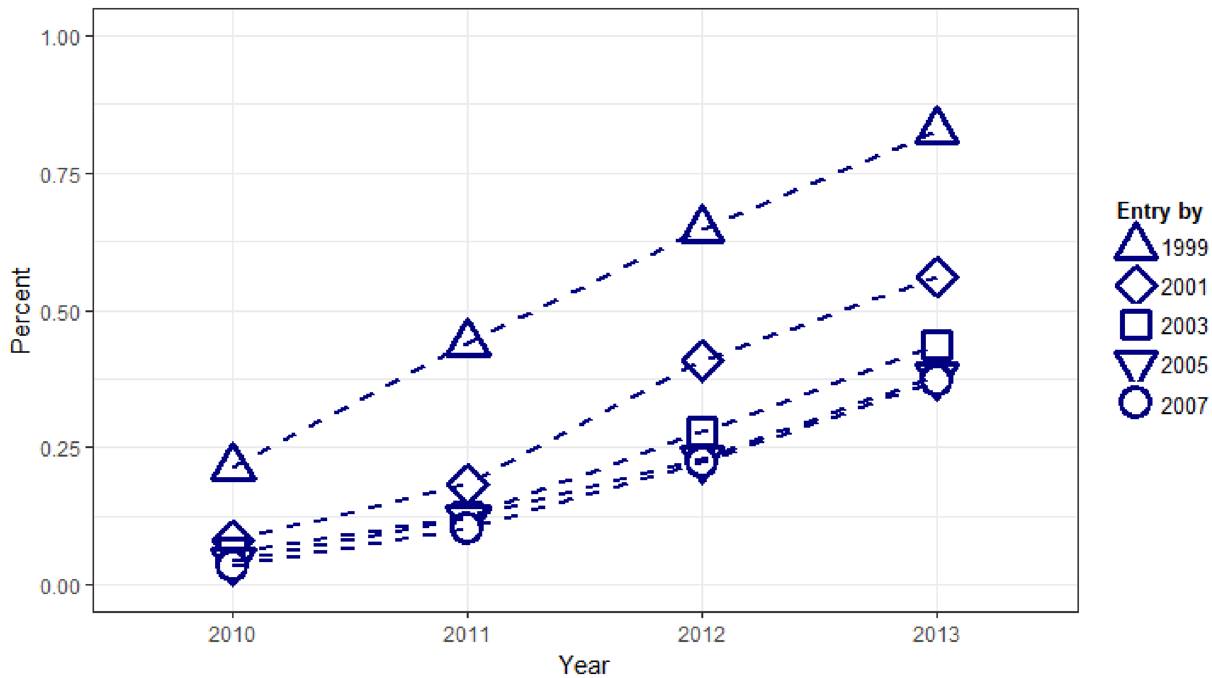
known as DOCSIS, the current version of which allows for many channels to be bonded together and used by a single subscriber. At the same time, cable firms have expanded their use of fiber-optic cables, which increase the available bandwidth and reduce congestion. Telephone firms have also deployed fiber-optic cables throughout their networks. This strategy is particularly important for them, as transmitting data at high speeds over long distances using their existing telephone wires is physically impossible. In some areas, cable and/or telephone firms have constructed networks which consist exclusively of fiber-optic cables, in what is known as fiber-to-the-premises.

As shown in figure 2, the fraction of zip codes with download speeds of at least 10 megabits per second (Mbps) increased from 87% to 97% from 2010 to 2013. Over the same period of time, the fraction of zip codes with download speeds of at least 25 Mbps rose from 40% to 84%. Finally, the share of zip codes with download speeds of at least 100 Mbps grew from just 15% to 68%.

2.2 Digital Divide and Government Policy

Despite the dramatic rise in broadband availability between 1999 and 2007, this deployment did not occur evenly across the country. In 2003, zip codes that did not yet have access to broadband

Figure 3: Fraction of Zip Codes with Access to 100 Mbps by Year of Initial Entry



Source: National Broadband Map

internet had median household incomes that were $7,340 less than those with broadband internet; they also had a rate of college graduation that was 9 percentage points lower and were far more likely to be in rural areas. The U.S. census records the percentage of each zip code which is considered rural, and markets without broadband in 2003 were, on average, 87% rural, while those with broadband were only 59% rural.[10]

Similarly, the improvements in speed between 2010 and 2013 were not uniform. In 2013, zip codes without access to download speeds of at least 25 Mbps, on average, had $10,505 lower household incomes, college graduation rates which were 7 percentage points lower, and were 37 percentage points more rural.

The timing of a market's initial availability of broadband internet and its present-day speeds are closely related. Figure 3 shows the trajectory of availability of download speeds of 100 Mbps, broken down by year of initial entry. Markets which were initially entered earlier were more likely to have access to 100 Mbps sooner; and, perhaps more interestingly, this gap has widened over time. Of

[10]Comparable statements can be made about the set of markets without access to broadband in any particular year. Year 2003 is in no way unique is this regard.

course, this relationship may simply be the result of the demographic correlations outlined earlier in this section; or, it may be the case that the timing of initial entry has a causal impact upon long-run broadband quality. If markets that are entered later do not become competitive until later still, and sustained competition leads to quality improvements, then this delayed entry will directly impact broadband quality in the long run. We investigate this relationship in our empirical analysis and find evidence to support this mechanism.

While demographic differences can explain a great deal of the disparity in internet availability and quality provision, the strategic entry decisions of firms may serve to exacerbate this issue. This inequality in access to and quality of broadband internet has been termed the "digital divide," and has consistently been a major policy concern in the United States. The Communications Act of 1934 established the goal of Universal Service in telecommunications, which meant that quality services should be made universally available at just and affordable rates without discrimination by income or ruralness. The 1996 Telecommunications Act codified these principles to apply to high-speed internet service. The FCC oversees a number of programs which aim to accomplish this goal, including the Connect America Fund, which subsidizes the expansion of ISPs' networks, and the Lifeline Program, which subsidizes prices paid by low-income households.

3 Data and Definitions

Our analysis is based primarily on two sources of data compiled by the FCC and in partnership with the National Telecommunication and Information Administration (NTIA). The first data set is the FCC's Form 477, collected bi-annually by the FCC beginning in 1999, and made available from 1999 to 2007. The FCC requires every facilities-based provider with at least 250 high-speed lines[11] to report its presence in all zip codes in which they have at least one customer. The FCC releases summary statistics to the public aggregated to the zip code level. From these snapshots of market structure, we can observe the timing of net entry and exit of broadband providers over six month intervals. In our study, we only use the December data, in order to allow sufficient time for changes in market structure to occur, and so that our net entry and exit is measured over one year intervals.

[11] High-speed lines are defined as those that provide speeds exceeding 200 kilobits per second (kbps) in at least one direction.

This data set, covering the entire United States and spanning multiple time periods, provides a rare opportunity for researchers to study market evolution in the early stages of a rapidly-growing service industry. However, we must acknowledge some drawbacks of the data. It lacks firm identities, so we can only observe net entry rather than actual entry and exit of firms. It also means that our inference of entry threat is derived from observations of the number, but not the identities, of incumbent providers across markets. We also cannot distinguish between different types of broadband services such as cable and DSL, and so we cannot test whether the effect of entry threat differs by provider type. Furthermore, very small providers (with less than 250 high-speed lines) are not required to report to the FCC, generating measurement errors in our econometric analysis.[12] Finally, for confidentiality reasons, the data indicates the presence of 1, 2, or 3 providers with a single indicator. As a result, we focus our analysis on the decision of the first entrant, rather than subsequent entries. Despite these limitations, the breadth and depth of the data generate interesting inferences about entry decisions which are not possible using other available data. Richer data on the identities of firms is only available through more contemporary data sets, such as the National Broadband Map, which began in 2010.

Accordingly, our second data set is the source data from the 2013 National Broadband Map, which was collected through the State Broadband Initiative, a program overseen by the NTIA. This data provides information about the current state of the U.S. broadband infrastructure, indicating the identities of each firm in each census block, along with their local technology and maximum advertised download and upload speeds. Maximum advertised speeds are reported as a categorical variable, whose values represent ranges of speeds; we replace these values with the median value of the relevant range. In order to pair this data with the early FCC data, we aggregate observations to the zip code level by taking a population-weighted average across the blocks within a zip code.

Finally, we use three auxiliary data sets. We use demographic characteristics from the 2000 Census and the 2010 American Community Survey (ACS), based upon zip code tabulation areas (ZCTAs).[13] The variables selected include population, average income, education, age, ethnicity, commuting distance, and population density, all of which affect local demand for and/or the cost of

[12] Fortunately, few providers fall into this category. Paradyne (2000) shows that entry is not profitable unless there are at least 200 lines in a DSL service area.

[13] ZCTAs, defined by the Census Bureau, are not identical to zip codes, which are defined by the U.S. Postal Service. However, all zip codes in the FCC data do have a match in the 2000 Census data.

providing internet service. The 2000 Zip Code Business Patterns provides the number of business establishments for each zip code, which serves as our measure of local business activity. Descriptions and summary statistics for these variables are provided in section 3.5.

3.1 Market Definition

In any service industry, consumer mobility determines the boundaries of a local market. This can be quite challenging, as researchers typically do not have good data on consumers' willingness to travel for desirable services. The broadband market, however, is fairly unique; consumers have no mobility at all, as they can only purchase a subscription from providers offering service at their residence. Therefore, we avoid the problem of blurred market boundaries which complicates many studies of market structure.[14]

Fortunately, the FCC data offer a natural definition for markets by indicating the number of firms offering service within each zip code. Since households cannot subscribe to a broadband provider who does not serve their zip code, this provides us with a clean market boundary. With that said, one might wonder whether broadband providers make entry decisions at such a fine geographic level. They may instead make decisions at the city, county, or even state level, though it would take years to roll out full coverage to these larger areas. This type of long-run strategy would not only compromise our market definition, but would also threaten the validity of the instruments we propose in section 5.2. Furthermore, the relevant market definition in the long-run likely varies considerably across firms, as some broadband providers have a national presence, while others serve only one city.

For these reasons, we focus on the short-run entry decision of firms, the gradual "rolling out" process of broadband providers. Specifically, we consider a broadband provider's marginal decision to expand service to one more local market. Since expanding service to a local market involves sunk costs, we can define the boundaries of the market by the nature of those costs. Following Xiao and Orazem (2011), we note that in the broadband industry, these sunk costs are the costs of deploying the so-called "last mile" of infrastructure. Firms must lay or renovate coaxial cables and telephone wires, as well as build or modify switching and distribution centers, cable television head ends, and

[14] Complete consumer immobility does, however, have the potential to create a problem of its own. If we define a local market to be too large, a provider within the market may not actually offer service to all households.

DSL access multiplexers. The distance between an end-user and a broadband provider is a primary factor in determining which neighborhoods can be served, particularly for telephone providers. This physical constraint limits the radius of a local market, as DSL can be provided reliably within a radius of 18,000 feet, or about 3.4 miles from the firm's central office. This again suggests that zip codes are the appropriate geographic approximation of a local broadband market, as the typical zip code has a radius of between 3 and 4 miles, according to the 2000 U.S. Census. While cable internet providers do not face such a distance constraint, as their data transmission speeds do not decay over distance, they do face a similar limitation. A cable provider runs fiber-optic cable from its headend to an optical node, then runs coaxial cable from the node to subscribers' premises. A coaxial cable there has a fixed amount of bandwidth and cable internet subscribers all share the coaxial cable nearest them, which means that data speeds are limited by the number of subscribers sharing a node. Thus, in order to provide consistent high-speed access, cable providers must make costly localized investments in installing nodes.

3.2 Neighboring Market Definition

Our data indicates the latitude and longitude of the centroid of each zip code in the United States. The distance between the centroids of two zip codes forms the basis of our definition of neighboring markets. However, this distance alone ignores the geographic sizes of the zip codes; in fact, two large zip codes could border one another but have centroids which are far apart. In order to address this issue, we assume that all zip codes are circular and calculate their radius as follows:

$$r_m = \sqrt{\frac{area_m}{3.14}} \tag{1}$$

where r_m is the radius of market m and $area_m$ is the geographic area of market m. We then define market m and market m' to be neighbors according to the following definition:

$$neighbors_{m,m'} = \begin{cases} 1 & \text{if } distance_{m,m'} \leq 3 + \max\{r_m, r_{m'}\} \\ 0 & \text{otherwise} \end{cases} \tag{2}$$

where $distance_{m,m'}$ is the distance between the centroids of markets m and m'. Put more simply, we define two markets to be neighbors if the distance from the centroid of one market to the boundary of the other is less than 3 miles. The choice of 3 miles is somewhat arbitrary, but supported by the physical limitations of DSL technology. A telephone provider with a central office located at the centroid of market m could feasibly serve market m' if the boundary of market m' was within about 3 miles of the centroid of market m, as noted in the previous section.[15] This technological constraint highlights the appeal of this definition of neighboring markets. Defining markets to be neighbors if they border one another, while seemingly logical, would be inappropriate in our setting. If adjacent markets are large, one may still be well out of range for providing DSL service using their existing central office in the other. And if markets are small, a single central office may be able to provide DSL service to nearby markets that it does not border.

3.3 Entry Threat Definition

In the broadband industry, there are enormous economies of scale in building out a network. As a result, an incumbent provider will find it much easier to spill over into an adjacent market than to enter a more distant market. Therefore, we define a market, m, to be *threatened* if at least one rival firm operates in some neighboring market, m', but not in market m. Formally, the entry threat status of market m at time t is

$$EntryThreat_{mt} = \begin{cases} 1 & \exists\ m' \text{ s.t. } neighbors_{m,m'} = 1 \text{ and } N_{m't} > 1 \\ 0 & \text{otherwise} \end{cases} \tag{3}$$

where $neighbors_{m,m'}$ is as defined in equation (2) and $N_{m't}$ is the number of firms serving market m' at time t.[16] Critically, this means that a potential entrant's presence in a neighboring market is not enough to label a market as threatened; at least one additional firm must be present as well.

[15] To test the sensitivity of our analysis to this choice, we replace the value of 3 in equation (2) with alternative values and obtain result which are qualitatively unchanged.

[16] Because our data is censored such that all markets with 1, 2, or 3 firms are recorded as a 1, we can only detect that a market has more than 1 firm if that market has at least 4 firms. $EntryThreat_{mt}$ is defined subject to this limitation.

Table 1: Neighbors Summary Statistics

Variable	Mean	Standard Deviation	Min	Median	Max
# Neighbors	2.611	1.533	1	2	28
Entry Threat	0.069	0.253	0	0	1
# Markets: 7,642					

3.4 Sample Selection

The FCC's Form 477 data contains deployment data on the universe of the 31,862 zip codes in the United States. From this set, we drop 1,784 markets which do not have a neighboring market. We also drop 64 markets which have more than 30 neighbors; these markets cover very little geographic area and therefore do not fit our market definition. Finally, there are 1,799 markets which are missing demographic data, and so we drop these observations from our sample, leaving us with a total of 28,207 markets.

Because the focus of our analysis is on the decision of the first entrant to enter an open market, and because our goal is to understand the effect of the threat of future competition on firms' entry decisions in the early rollout of broadband infrastructure, we restrict attention to markets which were unserved in the year 2000,[17] leaving us with 8,476 markets.

Finally, there are some zip codes which we do not observe in the 2013 National Broadband Map data. In the interest of maintaining a consistent set of observations across all specifications, we drop these from our sample, leaving us with 7,642 markets.

3.5 Summary Statistics

Summary statistics for neighbors and entry threat are shown in table 1. On average, a market has 2.61 neighbors, and this varies between 1 and 28 for all markets in our sample. Of the 7,642 markets in our sample, about 7%, or 535 markets are threatened.

Table 2 presents summary statistics for each of the market characteristics included in our specifications, broken down by entry threat status. Threatened and unthreatened markets are quite similar across most dimensions, though threatened markets typically have larger populations, are less rural, more densely populated, and have more businesses.

[17] Our data begins in 1999, and would therefore permit us to study one year earlier, but only a tiny fraction of open markets were threatened at that time, which affords our estimates very little power. For robustness, we carry out our full analysis for 1999 and find qualitatively similar results.

Table 2: Market Characteristic Summary Statistics

Variable	Mean | Threat = 0	Mean | Threat = 1	t Statistic
Population	1,159.095	1,784.641	3.018
% Black	0.050	0.036	-2.491
% Hispanic	0.038	0.059	3.272
% Am. Indian	0.023	0.007	-7.655
% Asian	0.003	0.008	4.411
log(Median Income)	10.361	10.629	14.784
% College	0.364	0.433	7.436
Household Size	2.571	2.601	1.797
% Female	0.498	0.503	2.806
% Senior	0.327	0.310	-3.647
% Work From Home	0.062	0.037	-11.074
% Long commute	0.199	0.188	-1.869
% Rent	0.196	0.226	4.650
% Phone	0.951	0.980	17.086
% Rural	0.946	0.599	-17.337
log(Population Density)	3.262	5.862	33.375
log(Business Density)	2.804	3.221	8.781
Δ Population	3.953	72.579	4.595
Δ log(Median Income)	0.022	0.019	-2.134
Δ % College	0.006	0.005	-0.904
Δ Household Size	-0.004	-0.004	0.128
Δ % Rural	-0.0005	-0.0053	-4.139
Δ log(Population Density)	-0.0050	-0.0049	0.034
Δ log(Business Density)	-0.041	-0.036	0.019
Number of ISPs, 2013	1.149	2.176	24.686
# Markets	7,115	527	

Table 3: Outcome Summary Statistics

Variable	Mean | Threat = 0	Mean | Threat = 1	t Statistic
Short-run Entry	0.341	0.182	-8.943
Long-run Entry	0.906	0.983	11.634
Entry Delay	3.468	6.476	19.770
Mean 2013 Download Speed (Mbps)	72.685	140.998	13.589
# Markets	7,115	527	

Despite these seemingly attractive features, in the short-run, threatened markets are entered much less often than their unthreatened counterparts. Table 3 shows that threatened markets were entered about half as often between 2000 and 2001. In addition, markets which were open in 2000 waited, on average, 3 years longer to be entered if they were threatened. In the long run, however, threatened markets were more likely to be entered, as 98% had access to at least 1 provider in 2013.

4 Empirical Framework

4.1 Entry in the Short Run

When modeling the decision of a firm to enter an open market, we consider a static profit function, the discounted value of the firm's expected future profit stream, in the spirit of Bresnahan and Reiss (1991). In this setting, firms make their entry decisions and then receive continuation values which depend on the actions of other firms. We choose this simpler profit function over a full-fledged Bellman equation representation of the expected value of entry because firms faced enormous uncertainty about their rivals' behavior. The industry was still in its infancy, which meant that industry norms had not yet formed and that the turnover rate was very high. Therefore, we do not believe that writing down an explicit value function, which would require firms to make predictions about the entire future evolution of the market, is appropriate for this setting. Since our data does not allow us to distinguish between markets with 1, 2, and 3 firms, we focus our analysis on open markets with no providers.

Consider the decision of a potential entrant to enter an open market, m, which contains no firms at the beginning. This potential entrant, p, observes the state of the market and assesses whether the expected discounted value of the future profit stream is sufficiently high to support entry. This assessment is based upon market demand, growth projections, cost of deployment and services, and, very importantly, its anticipated future market structure. As this is a new, highly uncertain industry with a mix of local and national providers, potential entrant p does not directly observe the number of other potential entrants to market m; instead, potential entrant p forms its expectation about the future market structure of market m based upon the market structures of neighboring markets. As defined in equation (3), firm p considers market m to be threatened if at least one rival firm operates in a neighboring market but not in market m. This set up differs from the traditional entry game frameworks of Bresnahan and Reiss (1990, 1991) or Berry (1992), as we are not trying to pin down the equilibrium *number* of firms entering market m, bur rather only *whether* market m is entered in the short run by one or more firms. This difference allows us to avoid making strong assumptions about the number or identity of the potential entrants.

The expected discounted value of future profits of potential entrant p from entering market m is

$$\mathbb{E}(\Pi_m^p) = \alpha_0 + X_m\alpha_1 + \alpha_2 EntryThreat_m + \nu_m^p \tag{4}$$

This reduced form representation of expected profits states that profits depend on a vector of market attributes (X_m), the threat of future competition, and a normally-distributed stochastic error term (ν_m^p) which includes factors influencing profits that are observed by firm p but not by the econometrician. Since we restrict our sample to markets with zero firms, the number of incumbents does not appear in the profit function. Firm p will therefore enter market m in the time period if and only if $\mathbb{E}(\Pi_m^p) \geq 0$. We represent this entry decision with a binary variable, D_m^p, where a value of 1 indicates entry. Therefore,

$$Pr(D_m^p = 1) = \Phi(\alpha_0 + X_m\alpha_1 + \alpha_2 EntryThreat_m) \tag{5}$$

where $\Phi(\cdot)$ is the cumulative distribution function of a normally distributed random variable, described in detail in section 5.1. In our primary specifications, we use data from the year 2000 to capture short-run entry behavior during the industry's infancy. For robustness, we repeat our analysis using the year 1999 and obtain comparable results, reported in section 6.4.1.

Market-specific variables which we expect to influence variable profits and fixed operating costs are represented by X_m. Market size, as measured by population, is a key determinant of profitability, as shown by Bresnahan and Reiss (1990, 1991). We also include local demographic variables such as ethnicity composition, age and gender profiles, education levels and household income as profit shifters. In addition, we include variables that measure potential consumers' time spent using the internet. For example, consumers with long commutes may have less time budgeted to internet use at home, while consumers who work from home may need the internet for effective work performance. We also consider factors affecting the cost of rolling out the physical network. Housing structure, telephone penetration rate, population density and business density[18] fall into this category. We suspect that the "time use" and "cost" variables are of secondary importance in the early days of the broadband industry though, as internet use had not become a norm. Lastly, we include the growth rates of market attributes, which capture firms' expectations about the evolution of demand over time.[19]

The intent of α_2 is to capture firm p's concern over a competitive future market structure. However, because we do not observe firm identities, when $EntryThreat_m = 1$, it is possible that firm p is itself one of the firms present in a neighboring market. In such a case, firm p can more easily enter market m, and α_{2t} will pick up this positive spillover effect. Therefore, α_2 represents the net of this spillover effect and any strategic effects of entry threat. Though the sign of the spillover effect is known, the direction of the strategic effect is *ex-ante* unclear. Firms may be likely to quickly enter a threatened market in order to preempt their competition; or, they may be less likely to enter because the likelihood of future competition lowers their expectation of the market's future profitability. Unfortunately, we cannot separately identify these effects; but, since the spillover effect is known to be positive, a negative $\hat{\alpha}_2$ would indicate that the presence of rivals in neighboring markets makes potential entrants less likely to enter a market, and that this effect dominates both the spillover and (potential) preemption effects.

We intend for our measure of entry threat to capture the likelihood of eventual entry into market m. Therefore, if this measure is credible, it must be the case that threatened markets are *more* likely to be entered in the long run, regardless of the sign of the short-run entry threat effect. To test this, we estimate equation 5, replacing D_m^p with an indicator for whether market m is entered

[18] Business density is defined as the total number of business establishments in a market divided by its population.
[19] Growth rates are computed as the average annual change in market attributes from 2000 to 2013.

in the long-run, by 2013.[20]

Importantly, we do not consider long-run entry to be a distinct decision made by firms. Rather, firms undertake an entry decision in each period, a decision driven by the discounted value of their future profit stream. This sequence of entry decisions then results in a long-run entry outcome. It follows that entry threat may lower the probability of entry in the short run but increase the probability of entry in the long run. In the short run, if the effect of future competition outweighs the spillover and preemption effects, entry threat will lower the probability of entry. In later periods, after observing rivals decline to enter market m, the spillover effect may dominate the threat of future competition and thereby increase the probability that firms with a presence in neighboring markets enter market m, thereby fulfilling the predictions of earlier potential entrants.[21]

4.2 Entry Delay

If the threat of future competition makes firms less likely to enter a market in the short run, then the natural follow up question becomes: how long do firms delay entry into a market as a result of this threat? Since potential entrants make entry decisions in each period, this short-run effect may persist, resulting in a lack of entry for several periods. To understand the impact of entry threat on the delay of entry, we estimate the following model:

$$EntryDelay_m = \beta_0 + X_m\beta_1 + \beta_2 EntryThreat_m + \omega_m \tag{6}$$

where $EntryDelay_m$ is equal to the number of years elapsed from 2000 until market m receives its first entrant.[22] X_m is the same set of observed market characteristics affecting market m's profitability, and ω_m are normally-distributed stochastic error terms affecting entry into market m. β_2 represents the degree to which firms delay entry into market m in response to the threat of future competition.

[20] Again, for robustness, we estimate a specification using 1999 as the base year and obtain qualitatively similar results, reported in section 6.4.1. We also estimate a specification where the outcome is a binary variable equal to 1 if and only if market m contains at least 4 firms by 2013, in order to test whether threatened markets are more likely to become competitive and obtain nearly identical results.

[21] Appendix B develops a simple model to illustrate the entry threat and geographic spillover tradeoffs. We show that entry threat may lower the probability of entry in an open market in the short run but increase the probability of entry in the long run.

[22] For robustness, we estimate a specification using 1999 as the base year and obtain comparable results, reported in section 6.4.1.

As in the previous model, we cannot give a structural interpretation to β_2, as the parameter captures the net of all reasons why entry threat would effect the delay of entry into a market. However, even in the reduced form, the parameter is very telling: a positive β_2 informs us that on average, firms delay entry into a market which is threatened by competitors from neighboring markets.

4.3 Broadband Speeds in the Long Run

Finally, we investigate whether the early evolution of the broadband industry has had long-term effects. Nearly all Americans have access to broadband internet today, so it is important to ask whether the delay in gaining access experienced by many still matters. To this end, we estimate the effect of entry delay on maximal modern-day download speeds. We choose download speeds as our outcome of interest over alternatives such as upload speed or latency, as this is the characteristic that is most salient to consumers. This is the single statistic used by virtually every national internet service provider to differentiate each of their plans in advertisements. With this in mind, we estimate the following model:

$$\begin{aligned} \log(Speed_m) = \delta_0 + X_m\delta_1 + \delta_2 EntryDelay_m + \\ + \delta_3 \log(\#ISPs)_m + \eta_m \end{aligned} \tag{7}$$

where $Speed_m$ is the maximum advertised download speed (in Mbps) available in market m in 2013, X_m includes the 2013 values of the same attributes of the previous specifications and the growth rates in , $EntryDelay_m$ is the number of years elapsed since 2000[23] until market m is entered, $\log(\#ISPs)_m$ is the log of the number of ISPs present in the market in 2013, and η_m are unobservable shocks affecting broadband speeds.

It is important that we control for the number of ISPs present in this specification. Markets which experienced entry delay are, on average, less competitive even today, which we expect decreases their quality of service. Therefore, without this control, our estimated effect of entry delay could be fully attributable to the current market structure. While this is interesting in its own

[23] For robustness, we estimate a specification using 1999 as the base year and obtain nearly identical results, reported in section 6.4.1.

right, we are instead interested in whether the early evolution of the industry has a *direct* impact on present-day outcomes.

5 Identification and Estimation

5.1 Modeling Spatial Correlation in the Errors

In this subsection, we discuss the spatial correlation of the error terms ν_m, ω_m, and η_m in equations (4), (6) and (7). In the following, we use u to represent them, the random shock affecting firm decisions made in market m at time t. These shocks may capture regional spikes in demand, local economic fluctuations, regulatory hurdles, and any other factors which are not controlled for through observable market characteristics. As such, it is likely that these shocks are not isolated to a single zip code, but rather are correlated with the shocks experienced by other nearby markets. To allow for this possibility, we impose the following structure on the error terms:

$$u = \psi W u + \varepsilon \tag{8}$$

In equation (8), u is an $M \times 1$ vector of error terms, where M is the total number of markets; ψ is a scalar measuring the degree of spatial correlation; W is an $M \times M$ symmetric matrix with elements w_{ij} such that w_{ij} is a binary variable equal to 1 if and only if $i \neq j$ and $neighbors_{ij} = 1$; ε is an $M \times 1$ vector of identically and independently distributed random variables such that

$$\varepsilon \sim N(0, I_M) \tag{9}$$

where I_M is the identity matrix with dimension M. Note that m and t subscripts have been suppressed in equations (8) and (9) for ease of notation.

The ψ term captures spatial correlation in the errors, u. If $\psi = 0$, then there is no spatial correlation and each u_i is simply drawn from the normal distribution in equation (9). If $\psi \neq 0$, then spatial correlation exists, and there are unobservable factors which influence firms' decisions that are correlated across neighboring markets.

Given equation (8), the variance-covariance matrix of u is

$$V(u) = [(I_M - \psi W)'(I_M - \psi W)]^{-1} \tag{10}$$

and is heteroskedastic if $\psi \neq 0$. Therefore, if there is spatial correlation in the errors, the maximum likelihood estimates of the probit model will be inconsistent. To deal with this, we use the generalized method of moments (GMM) approach developed by Pinkse and Slade (1998), which yields consistent estimates of the parameters under spatial correlation.

In order to construct the residuals necessary to estimate equation (5), we follow Pinkse and Slade (1998) to define the generalized error term as

$$u_m(\theta) = [D_m - \Phi(G_m(\theta))]\frac{\phi(G_m(\theta))}{\Phi(G_m(\theta))[1 - \Phi(G_m(\theta))]} \tag{11}$$

where θ represents α or γ, D_m represents entry in the short or long run, $G_m(\theta) = \frac{\tilde{X}_m\theta}{v_m(\psi)}$, $\tilde{X}_m\theta$ represents the term inside $\Phi(\cdot)$ in equation (5), $v_m(\psi)$ is the square root of the m^{th} diagonal element of $V(u)$, and $\phi(\cdot)$ and $\Phi(\cdot)$ are the density and cumulative distribution functions of a standard normal random variable. Since equations (6) and (7) have continuous outcomes, after accounting for spatial correlation, their error term is

$$u_m(\theta) = \frac{y_m - \tilde{X}_m\theta}{v_m(\psi)} \tag{12}$$

where y_m represents the relevant outcome variable.

5.2 Endogeneity of Entry Threat

The interconnected nature of firms' decisions across markets gives rise to two sources of potential endogeneity in our entry threat variable. First, entry threat is not exogenous if error terms are spatially correlated across neighboring markets, which gives rise to a bias through the *spatial correlation channel*. To see this, consider two isolated neighboring markets: market of interest m and neighboring market m'. Given equation (4), unobservable factors in the error term of market m' influence firms' entry decisions there, which in turn determine the entry threat of market m. Therefore, if the the error terms of market m and m' are correlated, then the entry threat status and error term of market m are correlated. In the entry model of equation (4), this generates a positive bias through this channel, as the error term of market m' is positively correlated with both the error term and entry threat of market m, thereby generating a positive relationship between the two. By the same logic, in the entry delay model, the error term of market m' is positively correlated with the error term of market m but negatively correlated with entry threat,[24] leading to a negative bias through this channel. In the long-run speed model, entry delay is likely endogenous, but the direction of the bias is less clear, as firms' entry and speed decisions are made at different points in time.

Second, if firms do indeed consider the threat of future competition when making entry decisions, then our entry threat variable cannot be exogenous, as firms in neighboring markets each make entry decisions while considering the threat of spillover from the other. This gives rise to a bias through the *simultaneity channel.* To see this in the entry model of equation (4), first note that an increase in the error term of market m increases the probability of entry into market m, and therefore increases the entry threat of market m'. Next, if entry threat reduces (increases) the probability of short-run entry, then this increase in the entry threat of market m' will reduce (increase) the probability of entry into market m', which therefore reduces (increases) the entry threat of market m. Put together, the error term and entry threat of market m are negatively (positively) correlated through this channel, creating a negative (positive) bias in the entry model.[25] By the same logic, this channel generates a positive (negative) bias in the entry delay model, as the probability of short-run entry and entry delay are inversely related to one another. Again, the direction of the bias in the long-run speed model is less clear, as entry and speed decisions are made at different points in time.

Notably, this second source of endoegeneity exists regardless of the presence of spatial correlation, as long as entry threat has a direct effect on entry decisions. Since the two sources of endogeneity differ in their implied biases, the net direction of the bias is ultimately ambiguous. Nonetheless, they affect each of our empirical specifications and must be dealt with in order to give our estimates a causal interpretation.

[24]This is because an increase in the error term of market m' increases entry delay in market m', which therefore lowers entry threat in market m.

[25]This logic holds identically for our estimation of equation (4) using the outcome of long-run entry as well, because the endogenous variable, entry threat in year 2000, is determined in the short-run.

One might be concerned about another potential source of endogeneity, that threatened open markets may systematically have unattractive unobservable characteristics, relative to open markets without neighbors that have been entered. However, we do not believe this to be a serious concern. Table 2 shows that threatened markets have more favorable observable characteristics, such as higher population, income, education, and household size, and it is likely that unobservable characteristics conform to this pattern as well. This would bias us away from our hypothesis that entry threat lowers the probability of short run entry. More importantly, we argue that the instrumental variables we propose in the following paragraph are orthogonal to any such market-level unobservables, an assumption critical for resolving the endogeneity issues previously discussed.

In the spirit of Pinkse and Slade (1998), we use the average market attributes of all of a market's *neighbors'* neighbors as instruments for entry threat. These characteristics will affect entry into a market's neighbors, but should have no direct effect on the decision to enter the market itself. To be concrete, suppose that market m' has a neighbor, m'', but that market m and market m'' are not themselves neighbors. The attributes of market m'' clearly affect whether market m'' contains any firms. Additionally, a firm's presence in market m'' increases its likelihood of spilling over into market m', which then creates an entry threat for market m. In order for these instruments to be exogenous, firms may consider the attributes of neighboring markets when making their entry decision, but they must not base their decision upon the attributes of other, further away markets. In other words, firms can be forward-looking but must be sufficiently myopic. In principle, we could allow firms to consider any number of steps ahead and construct the appropriate set of instruments. It is important to note, however, that it is not valid to construct an instrument using own market characteristics, as they are not excluded. Any own-market characteristic that might predict a market's entry threat will almost certainly affect its profits directly.

5.3 Estimation

In order to deal with both the presence of spatial correlation and the endogeneity of our key variable, we adopt the generalized method of moments estimation framework of Pinkse and Slade (1998). We form the moment conditions using a set of L instruments, Z, such that $\mathbb{E}[Z'u(\alpha)] = 0$. Z is therefore an $M \times L$ matrix with L greater than or equal to the length of the parameter vector, θ.[26] The sample analogue of the moment is

$$S(\theta) = \frac{1}{M} Z' \hat{u}(\theta) \tag{13}$$

where $\hat{u}(\theta)$ is the generalized residual, the generalized error of equation 11), evaluated at $\hat{\psi}$.

When the model is just-identified, we solve for the $\hat{\theta}$ such that $S(\hat{\theta}) = 0$. When the model is over-identified,

$$\hat{\theta} = \arg\min_{\theta} S'(\theta) \Omega S(\theta) \tag{14}$$

where Ω is an $L \times L$ positive definite weighting matrix.[27]

Under this procedure, even when $\psi \neq 0$, $\hat{\theta}$ is consistent and asymptotically normal. We then estimate the variance-covariance matrix of $\hat{\theta}$ by using the following property of $\hat{\theta}$:

$$\sqrt{M}(\hat{\theta} - \theta) \sim N(0, [B_2(\theta)]^{-1} \frac{\partial S'(\theta)}{\partial \theta} \Omega B_1(\theta) \Omega \frac{\partial S(\theta)}{\partial \theta'} [B_2(\theta)]^{-1}) \tag{15}$$

where $B_1(\theta) = M\mathbb{E}[S(\theta)S'(\theta)]$ and $B_2(\theta) = \frac{\partial S'(\theta)}{\partial \theta} \Omega \frac{\partial S(\theta)}{\partial \theta'}$.

6 Results

6.1 Is Entry Threat Credible?

We have argued that the broadband industry has a natural indicator for the threat of future entry into a market, the presence of other firms in a neighboring market. The enormous localized fixed costs of broadband infrastructure ensure that entering nearby markets is far more efficient than entering more distant markets. Therefore, we should see that markets which are threatened in the early stages of the rollout are more likely to be entered in the long run. In support of this claim, table 4 presents the results of estimating equation (5), where $D^p_m = 1$ if market m is entered by 2013. The

[26] Note that even when all right hand size variables are exogenous, we still need one extra instrument in order to identify ψ.

[27] We use the optimal weighting matrix for Ω, which we construct according to the following steps: first, we get a consistent GMM estimate, $\hat{\theta}_1$ by using $\Omega = I_M$ in equation (14); second, we construct $\hat{\Omega} = ME[S(\hat{\theta}_1)S'(\hat{\theta}_1)]$.

first two columns report the results of estimating a linear probability model; the first column was estimated with ordinary least squares, while the second is estimated using neighbors' neighbors' attributes as instruments. The third column reports the average partial effects calculated from the results of the GMM estimation procedure described in section 5.3, which controls for spatial correlation in the error terms.[28]

We find that after accounting for the endogeneity of entry threat, markets which are threatened in the year 2000 are more likely to be entered by 2013. Under our preferred specification in column 3, we find that entry threat increases the long run probability of entry by 8.4 percentage points. This finding is consistent with the idea that firms can more easily spill over into areas where the firm already has a foothold, and critically, demonstrates that this threat of entry is "real".

Our results also illustrate the importance of addressing the endogeneity concerns over the entry threat variable. First, we find evidence of substantial spatial correlation in the error terms. We estimate ψ in equation (8) to be 1.000 with strong statistical significance.[29] Comparing columns 2 and 3 shows that failing to account for spatial correlation leads to a positive bias in our estimate of the entry threat effect, exactly as we predict in section 5.2. Second, the OLS estimates reported in column 1 suggest that there is no statistically significant effect of entry threat on long-run entry. The non-existence of the entry threat effect may be the balancing of biases in different directions through the *spatial correlation* and *simultaneity channels*, as discussed in section 5.2. In this set of results, it seems that the OLS regression mostly suffers from the endogeneity bias resulting from the *simultaneity channel*, given the net negative bias in our estimate of the entry threat effect.

6.2 Does the Threat of Future Competition Delay Entry?

Equipped with evidence that the threat of entry does indeed lead to entry in the long run, we turn to the question of how potential entrants revise their entry strategies in light of the threat of future competition. The results of estimating equation (4) are reported in table 5 and demonstrate that firms were less likely to enter a threatened market, other things equal. This finding is robust across all specifications. As before, we report results of estimating a linear probability model using OLS

[28] For this model and each of the models that follow, we a report the first stage estimates for the endogenous variable(s) in Appendix A. We also provide results tables which report the point estimates from the OLS, IV, and GMM specifications.

[29] We performed a test to reject the null hypothesis of no spatial correlation, or $\psi = 0$.

Table 4: Probability of Long-run Entry

	(1)	(2)	(3)
Entry Threat	-0.008	0.120***	0.084***
	(0.014)	(0.044)	(0.005)
Population (1,000)	0.009***	0.009***	0.021***
	(0.002)	(0.002)	(0.005)
Percent Black	-0.076***	-0.082***	-0.104***
	(0.024)	(0.024)	(0.023)
Percent Hispanic	-0.144***	-0.154***	-0.147***
	(0.032)	(0.032)	(0.032)
Percent American Indian	-0.319***	-0.329***	-0.197***
	(0.040)	(0.041)	(0.039)
Percent Asian	0.102	0.129	1.933**
	(0.149)	(0.150)	(0.948)
log(Median Household Income)	0.084***	0.068***	0.060***
	(0.016)	(0.017)	(0.018)
Percent Graduated College	-0.067***	-0.078**	-0.056
	(0.032)	(0.032)	(0.037)
Average Household Size	0.005	0.005	-0.006
	(0.017)	(0.018)	(0.024)
Percent Female	0.238**	0.254**	0.257**
	(0.101)	(0.101)	(0.102)
Percent Senior	0.019	0.003	0.029
	(0.061)	(0.061)	(0.069)
Percent Work from Home	0.142**	0.134**	0.170***
	(0.056)	(0.056)	(0.063)
Percent Long Commute	-0.311***	-0.319***	-0.212***
	(0.025)	(0.026)	(0.025)
Percent Rent	0.065*	0.052	0.059
	(0.038)	(0.039)	(0.048)
Percent with Phone	0.156***	0.151**	0.016
	(0.066)	(0.067)	(0.058)
Percent Rural	-0.021	0.013	-0.360
	(0.017)	(0.020)	(0.383)
log(Population Density)	0.018***	0.013***	0.017***
	(0.003)	(0.003)	(0.003)
log(Business Density)	0.001	0.001	-0.001
	(0.004)	(0.004)	(0.004)
Δ Population (1,000)	0.031	0.053	0.057
	(0.033)	(0.033)	(0.163)
Δ log(Median Household Income)	0.471***	0.374**	0.432***
	(0.150)	(0.154)	(0.153)
Δ Percent Graduated College	0.425	0.449	0.139
	(0.367)	(0.368)	(0.398)
Δ Average Household Size	-0.316***	-0.274**	-0.184*
	(0.107)	(0.108)	(0.102)
Δ Percent Rural	-0.127	0.122	-2.301
	(0.252)	(0.265)	(4.13)
Δ log(Population Density)	0.111*	0.084	0.145
	(0.062)	(0.063)	(0.077)
Δ log(Business Density)	0.067***	0.067***	0.013*
	(0.018)	(0.044)	(0.016)
ψ	-	-	1.000***
	-	-	(0.104)
Instruments: Neighbors' Neighbors' Attributes	N	Y	Y
Allow for Spatial Correlation	N	N	Y
# Markets	7,642	7,642	7,642

and IV in columns 1 and 2, and the average partial effects of our GMM estimation which controls for spatial correlation in column 3. We find that, on average, entry threat decreases the probability that a market is entered in the short run by 20.2 percentage points. We observe a positive bias in our OLS estimates, suggesting that spatial correlation is the dominant source of endogeneity. Our GMM estimates provide a further bias correction, as our IV estimate of the entry threat effect remains positively biased by spatial correlation, as discussed in section 5.2.

Since firms appear reluctant to enter threatened markets, we next investigate the persistence of this effect. Rather than looking solely at the entry decisions of a single time period, we estimate the effect of entry threat on the length of time elapsed until a market is entered, in order to take advantage of our panel of data. We report the results of estimating equation (6) in table 6.

We find that an open market which is threatened in the year 2000 is entered, on average, about 2 years later than its unthreatened counterpart. For perspective, the average unthreatened market which was open in 2000 waited about 3.5 years until being entered, so this estimated effect represents a significant delay.[30] There is a substantial negative bias in our OLS estimates, once again suggesting that spatial correlation is the dominant source of endogeneity.

6.3 Does Delayed Entry Affect Broadband Speeds in the Long Run?

Despite the delay created by internet service providers' reluctance to compete against their rivals, over 95% of zip codes had at least one internet service provider as of 2013. It is then natural to ask whether the delay we find really mattered, or whether "time heals all wounds." We therefore estimate the impact of delayed initial entry on the download speeds available in 2013, more than 10 years after the start of our sample. We report the results of estimating equation (7) in table 7. Since entry delay and the number of firms operating in a market may be endogenous, we instrument for them using the instruments described in section 5.2. We report our OLS and IV estimates in columns 1 and 2, and our GMM average partial effects in column 3.

We find evidence that delayed entry early in the rollout of the U.S. broadband infrastructure had a significant impact on download speeds available in 2013, specifically that each additional year that a market remained open translates into an 10.5% decrease in future download speeds. Remarkably,

[30] Our model also fits the data quite well. In the data, the mean (standard deviation) time until entry was 3.68 (3.09) years. The mean (standard deviation) in the fitted values computed using the estimates from column 3 of table 6 is 3.77 (2.02).

Table 5: Probability of Short-run Entry

	(1)	(2)	(3)
Entry Threat	-0.100***	-0.131*	-0.202***
	(0.023)	(0.071)	(0.057)
Population (1,000)	0.071***	0.073***	0.063***
	(0.003)	(0.003)	(0.013)
Percent Black	-0.032	-0.030	0.048
	(0.040)	(0.040)	(0.043)
Percent Hispanic	-0.113**	-0.110**	0.014
	(0.052)	(0.052)	(0.060)
Percent American Indian	-0.190***	-0.188***	-0.204**
	(0.066)	(0.066)	(0.073)
Percent Asian	-0.754***	-0.760***	-2.947
	(0.245)	(0.245)	(1.857)
log(Median Household Income)	0.121***	0.125***	0.200***
	(0.026)	(0.027)	(0.031)
Percent Graduated College	-0.001	0.001	0.009
	(0.052)	(0.053)	(0.056)
Average Household Size	0.004	0.004	-0.023
	(0.029)	(0.029)	(0.036)
Percent Female	0.168	0.164	0.101
	(0.165)	(0.165)	(0.211)
Percent Senior	-0.054	-0.050	-0.098
	(0.100)	(0.100)	(0.122)
Percent Work from Home	-0.316***	-0.314***	-0.341***
	(0.092)	(0.092)	(0.097)
Percent Long Commute	0.002	0.004	0.001
	(0.041)	(0.042)	(0.051)
Percent Rent	0.015	0.018	0.054
	(0.063)	(0.063)	(0.079)
Percent with Phone	-0.125	-0.124	-0.088
	(0.109)	(0.109)	(0.124)
Percent Rural	0.098***	0.090***	0.077*
	(0.028)	(0.033)	(0.046)
log(Population Density)	-0.033***	-0.031***	-0.032***
	(0.004)	(0.005)	(0.004)
log(Business Density)	-0.007	-0.007	-0.007
	(0.007)	(0.007)	(0.007)
Δ Population (1,000)	-0.245***	-0.240***	-0.146
	(0.053)	(0.054)	(0.163)
Δ log(Median Household Income)	0.025	0.048	0.230
	(0.245)	(0.250)	(0.153)
Δ Percent Graduated College	0.947	0.941	1.480**
	(0.601)	(0.600)	(0.398)
Δ Average Household Size	0.200	0.190	0.149
	(0.175)	(0.176)	(0.102)
Δ Percent Rural	0.145	0.085	-0.326
	(0.412)	(0.432)	(4.13)
Δ log(Population Density)	-0.022	-0.016	0.160
	(0.102)	(0.102)	(0.077)
Δ log(Business Density)	0.140***	0.140***	0.242***
	(0.029)	(0.029)	(0.016)
ψ	-	-	2.270
	-	-	(0.261)***
Instruments: Neighbors' Neighbors' Attributes	N	Y	Y
Allow for Spatial Correlation	N	N	Y
# Markets	7,642	7,642	7,642

Table 6: Entry Delay

	(1)	(2)	(3)
Entry Threat	1.055***	2.428***	2.043***
	(0.127)	(0.394)	(0.526)
Population (1,000)	-0.535***	-0.526***	-0.594***
	(0.017)	(0.017)	(0.053)
Percent Black	0.324	0.262	0.115
	(0.218)	(0.220)	(0.316)
Percent Hispanic	3.023***	2.916***	2.840***
	(0.283)	(0.286)	(0.445)
Percent American Indian	2.572***	2.464***	2.989***
	(0.361)	(0.364)	(0.521)
Percent Asian	-0.896	-0.601	-2.323
	(1.341)	(1.351)	(1.677)
log(Median Household Income)	-1.099***	-1.276***	-1.329***
	(0.142)	(0.151)	(0.237)
Percent Graduated College	1.413***	1.289***	1.218***
	(0.287)	(0.290)	(0.408)
Average Household Size	0.095	0.096	-0.051
	(0.156)	(0.157)	(0.228)
Percent Female	-2.433***	-2.253**	0.196
	(0.906)	(0.913)	(1.488)
Percent Senior	1.973***	1.796***	1.690**
	(0.547)	(0.553)	(0.771)
Percent Work from Home	0.889*	0.805	-2.267
	(0.502)	(0.505)	(0.772)
Percent Long Commute	0.254	0.165	-0.003
	(0.227)	(0.229)	(0.336)
Percent Rent	1.106***	0.974**	1.191**
	(0.344)	(0.348)	(0.507)
Percent with Phone	-1.101*	-1.153*	0.377
	(0.596)	(0.600)	(0.848)
Percent Rural	-0.522***	-0.158	0.0.096
	(0.151)	(0.181)	(0.232)
log(Population Density)	0.832***	0.782***	0.805***
	(0.023)	(0.026)	(0.037)
log(Business Density)	0.341***	0.336***	0.398***
	(0.036)	(0.036)	(0.051)
Δ Population (1,000)	0.533*	0.297	0.013
	(0.292)	(0.301)	(0.622)
Δ log(Median Household Income)	-3.047**	-4.091***	-4.928**
	(1.345)	(1.383)	(2.035)
Δ Percent Graduated College	-3.191	-2.929	0.367
	(3.295)	(3.316)	(4.792)
Δ Average Household Size	-2.518***	-2.058**	-0.242
	(0.960)	(0.974)	(1.323)
Δ Percent Rural	-1.530	1.141	1.194
	(2.259)	(2.386)	(3.099)
Δ log(Population Density)	3.833***	3.542***	3.630***
	(0.557)	(0.566)	(0.899)
Δ log(Business Density)	-1.215***	-1.216***	-0.779***
	(0.160)	(0.161)	(0.219)
ψ	-	-	3.885***
	-	-	(0.039)
Instruments: Neighbors' Neighbors' Attributes	N	Y	Y
Allow for Spatial Correlation	N	N	Y
# Markets	7,642	7,642	7,642

Table 7: 2013 log Maximum Available Download Speed (Mbps)

Variable	(1)	(2)	(3)
Entry Delay	-0.004	-0.088***	-0.105***
	(0.005)	(0.027)	(0.030)
log(Number of ISPs)	0.370***	1.046***	1.124***
	(0.023)	(0.106)	(0.154)
Population (1,000)	0.014*	-0.044***	-0.055***
	(0.008)	(0.017)	(0.020)
Percent Black	-0.080	-0.005	0.017
	(0.092)	(0.105)	(0.118)
Percent Hispanic	-0.210*	0.308**	0.443***
	(0.112)	(0.143)	(0.160)
Percent American Indian	-0.731***	0.009	0.573*
	(0.153)	(0.194)	(0.315)
Percent Asian	0.377	0.222	-0.230
	(0.526)	(0.565)	(0.324)
log(Median Household Income)	0.531***	0.208**	0.127
	(0.063)	(0.081)	(0.108)
Percent Graduated College	0.254*	0.416***	0.349*
	(0.130)	(0.144)	(0.184)
Average Household Size	0.048	0.094	0.055
	(0.059)	(0.064)	(0.078)
Percent Female	0.041	-0.188	0.078
	(0.200)	(0.216)	(0.278)
Percent Senior	-0.224*	0.135	0.286
	(0.131)	(0.154)	(0.192)
Percent Work from Home	0.375**	0.162	0.116
	(0.178)	(0.198)	(0.238)
Percent Long Commute	0.001	0.255***	0.315**
	(0.084)	(0.098)	(0.123)
Percent Rent	0.010	0.047	0.015
	(0.107)	(0.117)	(0.147)
Percent with Phone	0.723***	0.440	0.504
	(0.266)	(0.288)	(0.414)
Percent Rural	-0.008	0.164*	0.237**
	(0.068)	(0.084)	(0.102)
log(Population Density)	0.053***	0.059**	0.078**
	(0.011)	(0.027)	(0.034)
log(Business Density)	0.019	0.031	0.051**
	(0.016)	(0.020)	(0.025)
Δ Population (1,000)	-0.414**	0.412	0.603*
	(0.190)	(0.286)	(0.332)
Δ log(Median Household Income)	-4.870***	-2.036**	-1.006
	(0.732)	(0.891)	(1.188)
Δ Percent Graduated College	-2.329	-4.171**	-2.813
	(1.751)	(1.932)	(2.455)
Δ Average Household Size	-1.981***	-1.547**	-0.448
	(0.722)	(0.783)	(0.918)
Δ Percent Rural	1.223	0.764	1.032
	(1.031)	(1.118)	(1.338)
Δ log(Population Density)	-0.836***	-1.034***	-1.080***
	(0.253)	(0.332)	(0.407)
Δ log(Business Density)	-0.210	-0.534**	-0.597**
	(0.181)	(0.238)	(0.302)
ψ	-	-	2.9278***
	-	-	(0.015)
Instruments: Neighbors' Neighbors' Attributes	N	Y	Y
Allow for Spatial Correlation	N	N	Y
# Markets	7,642	7,642	7,642

this result is true even when controlling for the current number of firms serving a market. This means that the mechanism for this effect is not simply that markets which are entered later are still less competitive in 2013 and therefore have faster available speeds. We do also find that the number of firms serving a market has a significant positive impact on local download speeds; we estimate that doubling the number of ISPs in a market translates into a 112.4% increase in download speeds.[31] This finding is consistent with Molnar and Savage (2017). We estimate a stronger effect, though our results are not directly comparable due to specification differences. We build on their result, as our results suggest that if two identical markets have the same number of service providers, the one which was initially entered first will have access to faster speeds.

We hypothesize that this is because when facing a competitive market structure, firms are under constant pressure to upgrade the quality of their service; absent this pressure, firms are more likely to remain stagnant. Mazzeo (2003), Matsa (2011), Wallsten and Mallahan (2013), Prince and Simon (2017), and Molnar and Savage (2017) find evidence of this in a static sense, that product quality is increasing in the number of current firms. We posit that current quality also depends on the duration of sustained competition, and that as a result, download speeds in markets which did not exhibit this competitive pressure until recently lag behind speeds in those which developed early. We propose a theoretic illustration in Appendix C to formalize this intuition. One might be tempted to predict the opposite result, that markets which are initially entered later are equipped with better technology and therefore would have faster download speeds today. However, 91% of zip codes had been entered by 2007, presumably with the cutting edge technology of the time. But, the prevailing download speeds of 2007 are wholly obsolete by today's standards; in fact, the average download speeds of 2007 do not even meet the FCC's current definition of broadband speeds. Therefore, regardless of the initial technology installed, it is only through continual improvements that firms can provide the download speeds we enjoy today.

To empirically test our hypothesis, we replace $EntryDelay_m$ in equation (7) with $CompetitiveDelay_m$, a variable which represents the number of years from 2000 it takes for the market to become competitive.[32] We present the results of this estimation in table 8. Indeed, we find that the longer a market takes to become competitive, the slower its 2013 download speed. In fact, the effect of

[31]Because we estimate a log-log specification, our results imply that a 100% increase in the number of ISPs in a market leads to a 112.4% increase in download speeds.

[32]We define a market to be competitive when it has at least 4 firms.

delayed competition is stronger than the effect of delayed entry, with each year of competitive delay translating into a 21% reduction in download speeds.

6.4 Robustness

6.4.1 Time Period

Our framework requires that we choose an initial time period, as entry delay and long-run entry must be defined relative to some base year. As our goal is to understand firms' strategies in the formative years of the industry, it was important that our base year be very early in the sample. 1999 is the earliest year in our data, but at that time, only 2% of open markets were threatened, so we chose to use 2000 as our base year.

For robustness, we also estimated all of our models using 1999 as our base year and obtained nearly identical results. Table 9 reports the average partial effects of our variables of interest, estimated using the GMM specification which allows for spatial correlation and uses the characteristics of neighbors' neighbors as instruments.

With a base year of 1999, we find that entry threat decreases the probability of short-run entry by 0.16, and that this translates into an entry delay of 3.59 years. This entry delay is one year longer than what we find when using 2000 as our base year, but this is to be expected, as the number of years elapsed from the base year until entry will change mechanically with the base year. We again estimate that markets which are threatened are more likely to be entered in the long run. Our estimate is comparable in magnitude to our primary specification, though not statistically significant. Finally, we estimate that one year of delayed initial entry leads to a 13.4% decrease in 2013 download speeds, and that each year that a competitive market structure is delayed leads to a 19.8% decrease in 2013 download speeds, estimates which are nearly identical to those under our primary specifications.

6.4.2 Definition of Neighbors

When defining whether two markets are neighbors, we necessarily made a choice of the maximum distance separating them. Our decision to use 3 miles was motivated by the technological constraints of the industry, but was, nonetheless, somewhat arbitrary. Therefore, we repeat our analysis using

Table 8: 2013 log Maximum Available Download Speed (Mbps)

Variable	(1)	(2)	(3)
Competitive Delay	-0.013*	-0.266***	-0.210***
	(0.008)	(0.052)	(0.056)
log(Number of ISPs)	0.369***	0.980***	1.042***
	(0.023)	(0.111)	(0.177)
Population (1,000)	0.010	-0.123**	-0.094***
	(0.008)	(0.026)	(0.029)
Percent Black	-0.091	-0.233**	-0.154
	(0.097)	(0.115)	(0.126)
Percent Hispanic	-0.213*	0.207	0.308*
	(0.111)	(0.136)	(0.160)
Percent American Indian	-0.727***	0.025	0.453
	(0.152)	(0.195)	(0.331)
Percent Asian	0.398	0.676	0.144
	(0.526)	(0.590)	(0.335)
log(Median Household Income)	0.521***	0.040	0.055
	(0.063)	(0.094)	(0.109)
Percent Graduated College	0.249*	0.292**	0.202
	(0.130)	(0.145)	(0.173)
Average Household Size	0.047	0.080	0.032
	(0.059)	(0.066)	(0.079)
Percent Female	0.032	-0.358	-0.022
	(0.200)	(0.229)	(0.291)
Percent Senior	-0.216*	0.256	0.278
	(0.131)	(0.161)	(0.194)
Percent Work from Home	0.384**	0.335	0.236
	(0.178)	(0.211)	(0.244)
Percent Long Commute	0.001	0.255**	0.290**
	(0.084)	(0.102)	(0.121)
Percent Rent	0.009	0.018	0.002
	(0.107)	(0.120)	(0.143)
Percent with Phone	0.719***	0.368	0.447
	(0.266)	(0.300)	(0.405)
Percent Rural	-0.226***	0.109	0.193*
	(0.068)	(0.088)	(0.113)
log(Population Density)	0.053***	0.067***	0.058**
	(0.010)	(0.022)	(0.028)
log(Business Density)	0.020	0.029	0.039*
	(0.016)	(0.019)	(0.023)
Δ Population (1,000)	-0.355*	1.470***	1.026**
	(0.192)	(0.396)	(0.433)
Δ log(Median Household Income)	-4.763***	-0.170	-0.161
	(0.734)	(1.036)	(1.227)
Δ Percent Graduated College	-2.314	-3.560*	-2.131
	(1.749)	(1.962)	(2.385)
Δ Average Household Size	-1.977***	-1.448*	-0.331
	(0.721)	(0.813)	(0.935)
Δ Percent Rural	1.261	1.513	1.549
	(1.031)	(1.179)	(1.370)
Δ log(Population Density)	-0.845***	-1.122***	-0.903**
	(0.251)	(0.317)	(0.373)
Δ log(Business Density)	-0.215	-0.569**	-0.485*
	(0.179)	(0.223)	(0.271)
ψ	-	-	2.927***
	-	-	(0.033)
Instruments: Neighbors' Neighbors' Attributes	N	Y	Y
Allow for Spatial Correlation	N	N	Y
# Markets	7,642	7,642	7,642

Table 9: Key Parameter Estimates, Base Year 1999

Variable	Outcome			
	Short-run Entry	Entry Delay	Long-run Entry	Long-run Speeds
Entry Threat	-0.164**	3.585***	0.060	-
	(0.071)	(0.812)	(0.074)	-
Entry Delay	-	-	-	-0.134***
	-	-	-	(0.027)
Competitive Delay				-0.198***
				(0.047)
# Markets	10,990	10,990	10,990	10,990

Table 10: Key Parameter Estimates, Alternate Neighbor Definitions

Neighbor Radius	Variable	Outcome			
		Short-run Entry	Entry Delay	Long-run Entry	Long-run Speeds (Mbps)
2 Miles	Entry Threat	-0.146*	3.049***	0.071**	-
		(0.085)	(0.686)	(0.035)	-
	Entry Delay	-	-	-	-0.022
		-	-	-	(0.035)
	Competitive Delay	-	-	-	-0.175**
		-	-	-	(0.073)
	# Markets	6,110	6,110	6,110	6,110
4 Miles	Entry Threat	-0.221***	1.026*	0.079**	-
		(0.037)	(0.584)	(0.031)	-
	Entry Delay	-	-	-	-0.083***
		-	-	-	(0.030)
	Competitive Delay	-	-	-	-0.198***
		-	-	-	(0.060)
	# Markets	8,348	8,348	8,348	8,348

a radius of 2 and 4 miles in order to test the sensitivity of our estimates to this choice. Table 10 reports the average partial effects of our key variables, estimated using the GMM specification which allows for spatial correlation and uses the characteristics of neighbors' neighbors as instruments.

Again, our results are quite robust to the neighboring markets definition and our qualitative conclusions remain unchanged. We now estimate that threatened markets receive their first entrant between 1 and 3 years later than their unthreatened counterparts, a finding which remains both statistically and economically significant. Our estimates of the effect of entry threat on short-run and long-run entry are statistically significant and comparable in magnitude to those under our primary specifications. When using a 2 mile radius in our definition of neighbors, we estimate a negative effect of entry delay on long-run download speeds, but this result is not statistically significant. This is likely driven by the reduction in sample size. When requiring that markets be within 2 miles of one another to be considered neighbors, far fewer markets have neighbors' neighbors and our sample falls to just 6,110. Nonetheless, we still find that each additional year of competitive delay decreases 2013 download speeds by 17.5% in this specification.

Table 11: Key Parameter Estimates, Excluded Neighbors' Neighbors

Required Distance	Variable	Outcome			
		Short-run Entry	Entry Delay	Long-run Entry	Long-run Speeds (Mbps)
5 Miles	Entry Threat	-0.156**	3.012***	0.084***	-
		(0.067)	(0.572)	(0.010)	-
	Entry Delay	-	-	-	-0.103***
		-	-	-	(0.030)
	Competitive Delay	-	-	-	-0.208***
		-	-	-	(0.057)
	# Markets	7,640	7,640	7,640	7,640
10 Miles	Entry Threat	-0.191***	2.159***	0.087***	-
		(0.069)	(0.602)	(0.011)	-
	Entry Delay	-	-	-	-0.112***
		-	-	-	(0.033)
	Competitive Delay	-	-	-	-0.239***
		-	-	-	(0.067)
	# Markets	6,982	6,982	6,982	6,982

6.4.3 Excluded Markets

Finally, the validity of our instruments rests on the assumption that firms do not plan entry decisions two or more steps ahead. That is, when deciding to enter a market, they may be influenced by their desire to enter neighboring markets in the future; however, they may *not* be influenced by markets which neighbor those neighboring markets. We believe that this assumption is credible, as the industry was in its infancy and firms faced enormous uncertainty about both the industry itself and their own viability as an ISP. However, in some cases, a market and its neighbor's neighbor may be in such close proximity that our assumption is unrealistic. Therefore, we repeat the estimation while constructing our instruments using only neighbors' neighbors which are sufficiently far away from the focal market. We report results in table 11 for specifications which exclude neighbors' neighbors which are within 5 and 10 miles of the focal market, and report the average partial effects estimated using the GMM specification which allows for spatial correlation and uses the characteristics of neighbors' neighbors as instruments. Our results under each restriction are nearly identical to those under our primary specifications, which lends credibility to the exogeneity of our instruments.[33]

To conclude our presentation of results, we discuss the role of predetermined market attributes such as population, household income, and education levels. From Table 4 to Table 11, we can see highly-populated, more affluent markets experienced more entry, while highly-educated, densely-populated markets enjoyed high internet speed. These are factors that determine the demand and

[33]We also test for sensitivity by estimating specifications in which we drop observations altogether if the market is too close to either its nearest or average neighbor and obtain nearly identical results.

cost side of internet service provision, and, ultimately, the competition landscape of the broadband industry. The results we see are often highly intuitive. The purpose of our study, however, is to highlight that firm strategy also plays an important role in shaping the uneven deployment of internet across markets. The effects of firm strategy can be anti-intuitive, as we show the contrast in the short-run and long-run entry probability in markets with neighboring firms. The robustness of such effects, again, shows that the phenomenon we find is real and persistent, leading to a long-lasting effect in this industry.

7 Conclusion

There is an established literature on how incumbent firms respond to the threat of rivals' future entry. It is then natural to back up one step to ask: before entering a market, do potential entrants consider the possibility of future entry of competitors and adjust their entry strategies accordingly? If so, do potential entrants delay entry due to lowered expectation of future profits, or do they accelerate entry due to preemptive incentives? To our knowledge, Seamans (2012) is the only predecessor to our work that has explored this angle.

In this paper, we focus on the early years of the broadband industry, when the industry was far away from its long-run steady-state equilibrium. In these first few years, the market environment was highly uncertain (chaotic even, with the 2000 tech bubble crush), and firms gradually entered local markets anticipating competition, fluctuations, and changes to their best ability. Accounting for spatial correlation and selective entry, we find evidence that firms delayed their entry, by an average of two years, into markets that were more likely to experience rival entry. Although this early deployment stage is long past, this delayed entry appears to have had effects which persist even today, as markets which experienced their initial entry later have access to considerably slower download speeds.

Broadband is pivotal infrastructure to a country. Equal access to such infrastructure has been a fundamental telecommunication policy goal in the United States since the 1996 Telecommunications Act. For example, the FCC's "Connect America Fund" provides substantial subsidies to entrants into rural, insular, and high-cost areas. Our findings suggest that public policies intended to encourage entry and competition should not restrict attention to preferences, cost, and technology considerations. We find that threatened markets experienced entry delay and inferior quality, while table 2 shows that these threatened markets actually appear to be more attractive, in many aspects, than unthreatened markets. The threatened markets have much larger populations, higher income, and better education; they are also more urban and more densely populated in both population and businesses. As Ellison and Ellison (2001) point out, strategic investments to deter entry matter most for medium-sized markets because they are unnecessary in small markets, and impossible in large ones. In our study, the negative effects of strategic entry delay have larger incidence on markets which also have more favorable characteristics than those at the bottom. In these left-behind markets, consumers suffered a lack of access, choices, and quality. In this sense, our findings echo Shapiro (2016), which finds that strategic entry delay limits consumer welfare significantly because consumers experience less product variety and valuable product characteristics for seven years.

Our study highlights the conflicts between short-run and long-run firm strategies. A long-run firm strategy consists of many short-run ones, but the long-run one is by no means the simple addition of short-run ones. As years go by, firms enter, profit, invest, expand, learn, contract, and exit. Gradually, uncertainty is resolved, entry and exit rates equalize, and the industry reaches a long-run steady state equilibrium (Jovanovic, 1982; Pakes and Ericson, 1990; Hopenhayn, 1992). During this dynamic process, some sources of firm- and market- level heterogeneity are transitory and others are persistent. Our empirical framework, albeit not a full-fledged dynamic structural model, captures a transition process in which transitory heterogeneity develops into persistent heterogeneity. We show that choices made on the basis of transitory conditions can persist long after those conditions have changed. In other words, history matters. In this way, our study is an open call to industrial organization researchers to develop equilibrium, dynamic, and spatial approaches in order to move beyond the current conditions of technology, preferences, and other factors that determine outcomes.

References

Anderson, M., A. Perrin, J. Jiang, & M. Kumar. (2019), "10% of Americans Don't Use the Internet. Who are They?" Pew Research Center, April 22. https://www.pewresearch.org/fact-tank/2019/04/22/some-americans-dont-use-the-internet-who-are-they/.

Anselin, L. (1988), Spatial Econometrics: Methods and Models, Kluwer Academic Publishers, Dordrecht.

Arthur, W. B. (1989), "Competing Technologies, Increasing Returns, and Lock-In by Historical Events," The Economic Journal, 99(394), 116-131.

Bresnahan, T. F. & P. C. Reiss (1990), "Entry in Monopoly Markets," The Review of Economic Studies, 57(4), 531-551.

—- (1991), "Entry and Competition in Concentrated Markets," Journal of Political Economy, 99(5), 977- 1009.

Dafny, L. (2005), "Games Hospitals Play: Entry Deterrence in Hospital Procedure Markets," Journal of Economics and Management Strategy, 14(3), 513-542.

David, P. A. (1985), "Clio and the Economics of QWERTY," The American Economic Review, 75(2), 332-337.

Dixit, A. (1979), "A Model of Duopoly Suggesting a Theory of Entry Barriers," Bell Journal of Economics, 10(1), 20-32.

Economides, N., K. Seim, & B. Viard (2008), "Quantifying the Benefits of Entry into Local Telephone Service," RAND Journal of Economics, 39(3), 699-730.

Ellison, G. & S. F. Ellison (2011), "Strategic Entry Deterrence and the Behavior of Pharmaceutical Incumbents Prior to Patent Expiration," American Economic Journal: Microeconomics, 3(1), 1-36.

Farrell, J. and P. Klemperer (2007), "Coordination and Lock-In: Competition with Switching Costs and Network Effects," Michael Armstrong and Robert Porter, eds., Handbook of Industrial Organization, Vol. 3, Amsterdam: North-Holland.

Federal Communications Commission (2015), "2015 Broadband Progress Report" https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2015-broadband-progress-report.

Gilbert, R. and D. Newbery (1982), “Preemptive Patenting and the Persistence of Monopoly,” The American Economic Review, 72(2), 514-526.

Goolsbee, A. & C. Syverson (2008), “How Do Incumbents Respond to the Threat of Entry? Evidence from the Major Airlines,” Quarterly Journal of Economics, 123(4), 1611-1633.

Greenstein, S. (2020), “The Basic Economics of Internet Infrastructure,” Journal of Economic Perspectives, 34(2) 192-214.

Greenstein, S., C. Forman, and A. Goldfarb (2018), “How Geography Shapes — and is Shaped by the Internet,” In The New Oxford Handbook of Economic Geography, edited by G. L. Clark, M. P. Feldman, M. S. Gertler, and D. Wojcik, 269–85, published by Oxford University Press, Oxford, UK.

Greenstein, S. & M. Mazzeo (2006), “The Role of Differentiation Strategy in Local Telecommunication Entry and Market Evolution: 1999-2002”, Journal of Industrial Economics, 54(3), 323-349.

Hopenhayn, H. A. (1992), “Entry, Exit, and Firm Dynamics in Long Run Equilibrium,” Econometrica, 60(5), 1127-1150.

Huckfeldt, P., & C. Knittel. (2012), “Pharmaceutical Use Following Generic Entry: Paying Less and Buying Less,” NBER Working Paper No. 17046.

Jackson, C. L. (2002), “Wired High-Speed Access,” Broadband —- Should We Regulate High-Speed Internet Access, edited by R.W. Crandall & J. H. Alleman, published by Brookings Institution Press, Washington D.C.

Jovanovic, B. (1982), “Selection and the Evolution of Industry,” Econometrica, 50, 649-670.

Klemperer, P. (1987), “Entry Deterrence in Markets with Consumer Switching Costs,” The Economic Journal, 97 (Supplement: Conference Papers), 9799-117.

Matsa, D. (2011), “Competition and Product Quality in the Supermarket Industry,” Quarterly Journal of Economics, 126(3), 1539-1591.

Mazzeo, M., 2002. “Competitive Outcomes in Product-Differentiated Oligopoly,” The Review of Economics and Statistics, 84(4), pp. 716-728.

Mazzeo, M. (2003), “Competition and Service Quality in the US Airline Industry,” Review of Industrial Organization, 22(4), 275-296.

Milgrom, P. & J. Roberts (1982), “Limit Pricing and Entry Under Incomplete Information: An Equilibrium Analysis,” Econometrica, 50(2), 443-460.

Molnar, G. & S. J. Savage (2017), "Market Structure and Broadband Internet Quality," Journal of Industrial Economics, 65(1), 73-104.

Pakes, A. & R. Ericson (1990), "Empirical Implications of Alternative Models of Firm Dynamics," Working Paper, Yale University, 1990.

Paradyne Corp. (2000), The DSL Sourcebook, http:www.paradyne.comsourcebook_offersb_1file.pdf.

Pinkse, J. & M. Slade (1998), "Contracting in Space: an Application of Spatial Statistics to the Discrete Choice Model," Journal of Econometrics, 85, 125-154.

Pinkse, J., M. Slade, & C. Brett (2002), "Spatial Price Competition: a Semiparametric Approach," Econometrica, 70(3), 1111-1153.

Prince, J. & D. Simon (2014), "Do Incumbents Improve Service Quality in Response to Entry? Evidence from Airlines' On-Time Performance," Management Science, 61(2), 372-390.

—- (2017), "The Impact of Mergers on Quality Provision: Evidence from the Airline Industry," The Journal of Industrial Economics, 65(2), 336-362.

Seamans, R. (2012), "Fighting City Hall: Entry Deterrence and New Technology Deployment in Cable TV Markets," Management Science, 58(3), 461-475.

Shapiro, B. T. (2016), "Estimating the Cost of Strategic Entry Delay in Pharmaceuticals: The Case of Ambien CR," Quantitative Marketing and Economics, 14(3), 201-231.

Spence, M. (1981), "The Learning Curve and Competition," Bell Journal of Economics, 12(1), 1981, 49- 70.

Wallsten, S., and Mallahan, C. (2013), "Residential Broadband Competition in the United States," in Shane Greenstein, Avi Goldfarb, and Catherine Tucker (Eds.), The Economics of Digitalization (Elgar Publishing: Northampton, Massachusetts, U.S.A.).

Wen, W. & F. Zhu (2019), "Threat of Platform-Owner Entry and Complementor Responses: Evidence from the Mobile App Market," Strategic Management Journal, 40(9), 1336-1367.

Appendix A

Instrumental Variables First Stage Results

Table 12: Entry Threat First Stage

Average Neighbors' Neighbors' Attributes		Market Attributes	
Population (1,000)	0.002***	Population (1,000)	-0.005***
	(0.001)		(0.001)
Percent Black	0.099***	Percent Black	0.030
	(0.034)		(0.030)
Percent Hispanic	0.049	Percent Hispanic	0.019
	(0.050)		(0.045)
Percent American Indian	0.139***	Percent American Indian	0.057
	(0.050)		(0.040)
Percent Asian	0.707***	Percent Asian	-0.924***
	(0.210)		(0.187)
log(Median Household Income)	0.105***	log(Median Household Income)	0.065***
	(0.016)		(0.014)
Percent Graduated College	0.029	Percent Graduated College	0.048*
	(0.034)		(0.028)
Average Household Size	0.029	Average Household Size	-0.015
	(0.019)		(0.014)
Percent Female	-0.159	Percent Female	-0.133*
	(0.123)		(0.078)
Percent Senior	0.242***	Percent Senior	0.094*
	(0.071)		(0.050)
Percent Work From Home	0.228***	Percent Work From Home	0.060
	(0.062)		(0.044)
Percent Long Commute	0.067**	Percent Long Commute	0.010
	(0.031)		(0.022)
Percent Rent	-0.073*	Percent Rent	0.089***
	(0.044)		(0.031)
Percent With Phone	0.040	Percent With Phone	0.006
	(0.075)		(0.055)
Percent Rural	-0.109***	Percent Rural	-0.187***
	(0.017)		(0.013)
log(Population Density)	0.013***	log(Population Density)	0.012***
	(0.003)		(0.002)
log(Business Density)	-0.006	log(Business Density)	0.007**
	(0.005)		(0.003)
# Neighbors' Neighbors	0.016***	Δ Population (1,000)	0.094***
	(0.001)		(0.025)
		Δ log(Median Household Income)	0.404***
			(0.117)
		Δ Percent College	-0.178
			(0.289)
		Δ Average Household Size	-0.253***
			(0.083)
		Δ Percent Rural	-1.527***
			(0.194)
		Δ log(Population Density)	0.057
			(0.049)
		Δ log(Business Density)	0.012
			(0.014)
# Markets	7,642		
F Statistic	49.19		

Table 13: log(# ISPs) First Stage

Average Neighbors' Neighbors' Attributes		Market Attributes	
Population (1,000)	-0.004***	Population (1,000)	0.022***
	(0.002)		(0.004)
Percent Black	-0.055	Percent Black	0.053
	(0.081)		(0.071)
Percent Hispanic	-0.494***	Percent Hispanic	-0.145*
	(0.109)		(0.084)
Percent American Indian	-0.284**	Percent American Indian	-0.494***
	(0.127)		(0.097)
Percent Asian	-0.575	Percent Asian	0.147
	(0.430)		(0.329)
log(Median Household Income)	0.214***	log(Median Household Income)	0.210***
	(0.042)		(0.034)
Percent Graduated College	-0.225**	Percent Graduated College	0.006
	(0.088)		(0.072)
Average Household Size	0.140***	Average Household Size	-0.047
	(0.048)		(0.031)
Percent Female	0.605*	Percent Female	0.233**
	(0.318)		(0.097)
Percent Senior	0.323*	Percent Senior	-0.193***
	(0.176)		(0.065)
Percent Work From Home	0.703***	Percent Work From Home	0.363***
	(0.158)		(0.088)
Percent Long Commute	-0.774***	Percent Long Commute	-0.205***
	(0.076)		(0.044)
Percent Rent	0.020	Percent Rent	0.048
	(0.110)		(0.053)
Percent With Phone	0.253	Percent With Phone	0.242*
	(0.185)		(0.130)
Percent Rural	-0.125***	Percent Rural	-0.293***
	(0.045)		(0.034)
log(Population Density)	0.050***	log(Population Density)	0.061***
	(0.009)		(0.006)
log(Business Density)	0.006	log(Business Density)	0.021***
	(0.012)		(0.008)
# Neighbors' Neighbors	0.015***	Δ Population (1,000)	-0.348***
	(0.002)		(0.091)
		Δ log(Median Household Income)	-1.576***
			(0.384)
		Δ Percent College	0.385
			(0.895)
		Δ Average Household Size	-0.474
			(0.376)
		Δ Percent Rural	0.857*
			(0.503)
		Δ log(Population Density)	-0.369***
			(0.124)
		Δ log(Business Density)	-0.152*
			(0.088)
# Markets	7,642		
F Statistic	24.21		

Table 14: Entry Delay First Stage

Average Neighbors' Neighbors' Attributes		Market Attributes	
Population (1,000)	-0.024***	Population (1,000)	-0.502***
	(0.008)		(0.017)
Percent Black	0.592	Percent Black	0.277
	(0.361)		(0.316)
Percent Hispanic	2.416***	Percent Hispanic	0.206
	(0.483)		(0.373)
Percent American Indian	1.477***	Percent American Indian	1.188***
	(0.561)		(0.432)
Percent Asian	1.879	Percent Asian	-2.782*
	(1.907)		(1.460)
log(Median Household Income)	-0.687***	log(Median Household Income)	-0.754***
	(0.188)		(0.151)
Percent Graduated College	2.633***	Percent Graduated College	-0.155
	(0.389)		(0.320)
Average Household Size	1.099***	Average Household Size	0.004
	(0.213)		(0.137)
Percent Female	-1.029	Percent Female	-0.280
	(1.411)		(0.431)
Percent Senior	1.195	Percent Senior	2.041***
	(0.782)		(0.287)
Percent Work From Home	-1.641**	Percent Work From Home	0.738*
	(0.699)		(0.389)
Percent Long Commute	-0.114	Percent Long Commute	0.342*
	(0.335)		(0.196)
Percent Rent	-1.308***	Percent Rent	0.913***
	(0.490)		(0.235)
Percent With Phone	0.883	Percent With Phone	-0.076
	(0.820)		(0.577)
Percent Rural	-1.609***	Percent Rural	-0.311**
	(0.199)		(0.150)
log(Population Density)	-0.357***	log(Population Density)	0.969***
	(0.038)		(0.026)
log(Business Density)	0.297***	log(Business Density)	0.315***
	(0.053)		(0.035)
# Neighbors' Neighbors	0.080***	Δ Population (1,000)	7.082***
	(0.010)		(0.402)
		Δ log(Median Household Income)	7.721***
			(1.703)
		Δ Percent College	-4.386
			(3.968)
		Δ Average Household Size	0.754
			(1.666)
		Δ Percent Rural	1.154
			(2.231)
		Δ log(Population Density)	-7.694***
			(0.550)
		Δ log(Business Density)	-4.912***
			(0.391)
# Markets	7,642		
F Statistic	19.67		

Table 15: Competitive Delay First Stage

Average Neighbors' Neighbors' Attributes		Market Attributes	
Population (1,000)	-0.007	Population (1,000)	-0.462***
	(0.005)		(0.012)
Percent Black	-0.420*	Percent Black	-0.184
	(0.251)		(0.219)
Percent Hispanic	0.419	Percent Hispanic	0.034
	(0.335)		(0.259)
Percent American Indian	0.271	Percent American Indian	0.762**
	(0.390)		(0.300)
Percent Asian	4.919***	Percent Asian	-1.170
	(1.325)		(1.014)
log(Median Household Income)	-0.593***	log(Median Household Income)	-0.757***
	(0.131)		(0.105)
Percent Graduated College	0.748***	Percent Graduated College	-0.467**
	(0.270)		(0.222)
Average Household Size	0.572***	Average Household Size	-0.055
	(0.148)		(0.095)
Percent Female	-1.744*	Percent Female	-0.760**
	(0.980)		(0.299)
Percent Senior	1.145**	Percent Senior	1.142***
	(0.543)		(0.199)
Percent Work From Home	-0.499	Percent Work From Home	0.737***
	(0.485)		(0.270)
Percent Long Commute	0.316	Percent Long Commute	0.188
	(0.233)		(0.136)
Percent Rent	-0.719**	Percent Rent	0.162
	(0.340)		(0.163)
Percent With Phone	0.387	Percent With Phone	-0.388
	(0.569)		(0.401)
Percent Rural	-0.515***	Percent Rural	-0.169
	(0.138)		(0.104)
log(Population Density)	-0.191***	log(Population Density)	0.348***
	(0.026)		(0.018)
log(Business Density)	0.155***	log(Business Density)	0.096***
	(0.037)		(0.024)
# Neighbors' Neighbors	0.055***	Δ Population (1,000)	6.263***
	(0.007)		(0.279)
		Δ log(Median Household Income)	7.897***
			(1.183)
		Δ Percent College	1.059
			(2.756)
		Δ Average Household Size	0.918
			(1.157)
		Δ Percent Rural	2.172
			(1.549)
		Δ log(Population Density)	-2.857***
			(0.382)
		Δ log(Business Density)	-1.732***
			(0.272)
# Markets	7,642		
F Statistic	11.79		

Point Estimates

Table 16: Probability of Short Run Entry

	(1)	(2)	(3)
Entry Threat	-0.100***	-0.131*	-1.276**
	(0.023)	(0.071)	(0.616)
Population (1,000)	0.071***	0.073***	0.302***
	(0.003)	(0.003)	(0.066)
Percent Black	-0.032	-0.030	0.231
	(0.040)	(0.040)	(0.22)
Percent Hispanic	-0.113**	-0.110**	0.066
	(0.052)	(0.052)	(0.261)
Percent American Indian	-0.190***	-0.188***	-0.975**
	(0.066)	(0.066)	(0.460)
Percent Asian	-0.754***	-0.760***	-14.098
	(0.245)	(0.245)	(8.927)
log(Median Household Income)	0.121***	0.125***	0.958***
	(0.026)	(0.027)	(0.174)
Percent Graduated College	-0.001	0.001	0.044
	(0.052)	(0.053)	(0.317)
Average Household Size	0.004	0.004	-0.111
	(0.029)	(0.029)	(0.168)
Percent Female	0.168	0.164	0.485
	(0.165)	(0.165)	(1.058)
Percent Senior	-0.054	-0.050	-0.471
	(0.100)	(0.100)	(0.584)
Percent Work from Home	-0.316***	-0.314***	-1.629***
	(0.092)	(0.092)	(0.523)
Percent Long Commute	0.002	0.004	0.006
	(0.041)	(0.042)	(0.221)
Percent Rent	0.015	0.018	0.260
	(0.063)	(0.063)	(0.387)
Percent with Phone	-0.125	-0.124	-0.422
	(0.109)	(0.109)	(0.703)
Percent Rural	0.098***	0.090***	0.370*
	(0.028)	(0.033)	(0.217)
log(Population Density)	-0.033***	-0.031***	-0.153***
	(0.004)	(0.005)	(0.022)
log(Business Density)	-0.007	-0.007	-0.031
	(0.007)	(0.007)	(0.041)
Δ Population (1,000)	-0.245***	-0.240***	-0.699
	(0.053)	(0.054)	(0.602)
Δ log(Median Household Income)	0.025	0.048	1.100
	(0.245)	(0.250)	(1.353)
Δ Percent Graduated College	0.947	0.941	7.077**
	(0.601)	(0.600)	(3.219)
Δ Average Household Size	0.200	0.190	0.714*
	(0.175)	(0.176)	(1.043)
Δ Percent Rural	0.145	0.085	-1.558
	(0.412)	(0.432)	(2.555)
Δ log(Population Density)	-0.022	-0.016	0.764
	(0.102)	(0.102)	(0.645)
Δ log(Business Density)	0.140***	0.140***	1.158***
	(0.029)	(0.029)	(0.259)
ψ			2.269***
			(0.261)
# Markets	7,642	7,642	7,642

Table 17: Probability of Long Run Entry

	(1)	(2)	(3)
Entry Threat	-0.008	0.120***	4.348
	(0.014)	(0.044)	(3.393)
Population (1,000)	0.009***	0.009***	0.177***
	(0.002)	(0.002)	(0.045)
Percent Black	-0.076***	-0.082***	-0.887***
	(0.024)	(0.024)	(0198)
Percent Hispanic	-0.144***	-0.154***	-1.260***
	(0.032)	(0.032)	(0.268)
Percent American Indian	-0.319***	-0.329***	-1.684***
	(0.040)	(0.041)	(0.327)
Percent Asian	0.102	0.129	16.556**
	(0.149)	(0.150)	(8.160)
log(Median Household Income)	0.084***	0.068***	0.517***
	(0.016)	(0.017)	(0.154)
Percent Graduated College	-0.067***	-0.078**	-0.482
	(0.032)	(0.032)	(0.314)
Average Household Size	0.005	0.005	-0.050
	(0.017)	(0.018)	(0.203)
Percent Female	0.238**	0.254**	2.199**
	(0.101)	(0.101)	(0.866)
Percent Senior	0.019	0.003	0.247
	(0.061)	(0.061)	(0.592)
Percent Work from Home	0.142**	0.134**	1.454***
	(0.056)	(0.056)	(0.540)
Percent Long Commute	-0.311***	-0.319***	-1.813***
	(0.025)	(0.026)	(0.218)
Percent Rent	0.065*	0.052	0.501
	(0.038)	(0.039)	(0.409)
Percent with Phone	0.156***	0.151**	0.135
	(0.066)	(0.067)	(0.499)
Percent Rural	-0.021	0.013	-3.087
	(0.017)	(0.020)	(3.250)
log(Population Density)	0.018***	0.013***	0.142***
	(0.003)	(0.003)	(0.028)
log(Business Density)	0.001	0.001	-0.005
	(0.004)	(0.004)	(0.034)
Δ Population (1,000)	0.031	0.053	0.488
	(0.033)	(0.033)	(1.394)
Δ log(Median Household Income)	0.471***	0.374**	3.696***
	(0.150)	(0.154)	(1.302)
Δ Percent Graduated College	0.425	0.449	1.189
	(0.367)	(0.368)	(3.412)
Δ Average Household Size	-0.316***	-0.274**	-1.573*
	(0.107)	(0.108)	(0.875)
Δ Percent Rural	-0.127	0.122	-19.712
	(0.252)	(0.265)	(35.189)
Δ log(Population Density)	0.111*	0.084	1.238*
	(0.062)	(0.063)	(0.670)
Δ log(Business Density)	0.067***	0.067***	0.114
	(0.018)	(0.044)	(0.135)
ψ			1.000***
			(0.104)
# Markets	7,642	7,642	7,642

Appendix B

In this appendix, we provide a simple theoretic example to illustrate the entry threat and geographic spillover effects trade-offs. We show that one potential entrant's entry incentive weakens when a market of interest faces entry threat by a competitor, and this entry threat effect has opposite implications for market entry in the short run and long run.

Two potential entrants, firm A and firm B, consider entry into an open market. They are given the opportunity to enter sequentially in two periods. Firm A considers entry in period 1 and firm B considers entry in period 2. Firms A and B face sunk costs of entry of $K_A + \varepsilon_A$ and $(1-\alpha)K_B + \varepsilon_B$, respectively, where $\alpha \in [0, 1)$. K_A and K_B represent expected sunk costs; ε_A and ε_B represent random shocks to expected sunk costs, and are assumed to be drawn *i.i.d.* from a standard normal distribution. Parameter α represents firm B's spillover effect, its cost reduction resulting from its (potential) presence in a neighboring market.

If firm B does not have a presence in any neighboring market, firm B experiences no spillover effect ($\alpha = 0$) and this market is therefore not *threatened*. If firm B is present in some neighboring market, firm B experiences a spillover effect ($\alpha > 0$) and this market is therefore *threatened*. The magnitude of α then corresponds to the strength of the spillover effect. Importantly, we do not take a stance on whether firm A has a presence in a neighboring market or the size of its costs. We allow for K_A to be less than or greater than $(1-\alpha)K_B$, and show that regardless, firm B's presence in a neighboring market lowers firm A's probability of entry.

If only one firm is present in the market during a given period, this firm will earn monopoly profits, π^m; if two firms are present in the market during a given period, each firm will earn duopoly profits, π^d, where $\pi^m > \pi^d$.

Error terms ε_A and ε_B are private information of firm A and firm B, respectively; all other parameters are common knowledge to both firms. Firms will earn zero profits if they do not enter, so a firm will enter the market if and only if its expected profit is greater than zero. After entering the market, the firm does not exit.

We solve the model using backward induction.

Period 2

Firm B is the only potential entrant. If firm A has entered in period 1, firm B will enter in period 2 if and only if

$$\pi^d - ((1-\alpha)K_B + \varepsilon_B) \geq 0 \tag{16}$$

Thus the conditional probability that firm B enters is

$$\begin{aligned} Pr(\text{Firm B Entry}|\text{Firm A Entry}) &= Pr(\pi_B \geq 0|\text{Firm A Entry}) \\ &= \Phi(\pi^d - ((1-\alpha)K_B)) \end{aligned} \tag{17}$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution.

If firm A did not enter in period 1, firm B will enter in period 2 if and only if

$$\pi^m - ((1-\alpha)K_B + \varepsilon_B) \geq 0 \tag{18}$$

The conditional probability that firm B enters is then

$$\begin{aligned} Pr(\text{Firm B Entry}|\text{No Firm A Entry}) &= Pr(\pi^B \geq 0)|\text{No Firm A Entry}) \\ &= \Phi(\pi^m - ((1-\alpha)K_B)) \end{aligned} \tag{19}$$

Period 1

Firm A is the only potential entrant. Firm A's expected profit from entering the market in period 1 is:

$$\begin{aligned} \pi^A = \pi_1^A + \pi_2^A &= [\pi^m + \pi^d - (K_A + \varepsilon_A)]\Phi[\pi^d - (1-\alpha)K_B] + [2\pi^m - (K_A + \varepsilon_A)]\{1 - \Phi[\pi^d - (1-\alpha)K_B)]\} \\ &= \pi^m - (K_A + \varepsilon_A) + \pi^d\Phi(\pi^d - (1-\alpha)K_B) + \pi^m[1 - \Phi(\pi^d - (1-\alpha)K_B)] \end{aligned} \tag{20}$$

Since firm A cannot enter in period 2, its expected profit of not entering the market in period

1 is 0. Therefore, firm A will enter in period 1 iff

$$\pi^A \geq 0 \tag{21}$$

Therefore, firm A enters the market with probability

$$Pr(\pi^A \geq 0) = \Phi\{\pi^m - K_A + \pi^d \Phi(\pi^d - (1-\alpha)K_B) + \pi^m[1 - \Phi(\pi^d - (1-\alpha)K_B)]\} \tag{22}$$

Proposition 1. *$Pr(\pi^A \geq 0)$ is a decreasing function of α.*

Proof.

$$\frac{\partial Pr(\pi^A \geq 0)}{\partial \alpha} = \tag{23}$$

$$\phi\{\pi^m - K_A + \pi^d \Phi(\pi^d - (1-\alpha)K_B) + \pi^m[1 - \Phi(\pi^d - (1-\alpha)K_B)]\} K_B \phi(\pi^d - (1-\alpha)K_B)(\pi^d - \pi^m)$$

$$< 0$$

■

where $\phi(\cdot)$ is the PDF of the standard normal distribution. The inequality holds because $\pi^m > \pi^d$ by assumption. Intuitively, this result is obtained because in period 1, firm A places greater weight on the possibility of earning duopoly profits in period 2 as α increases, because the probability that firm B enters in period 2 increases with α. Firm A therefore faces lower expected profits and has a lower probability of entry in the presence of entry threat from firm B. This proposition establishes that entry threat lowers the probability of entry in the short-run.

Next, we establish that firm B's probability of entry increases when it is present in a neighboring market. Conditional on firm A's entry decision, it is straightforward to show that firm B's entry probablity increases with α using equations (17) and (19). But in order to characterize firm B's entry decision at the start of period 1, we must calculate its unconditional entry probability. Unconditional on firm A's entry decision, firm B's expected profits if it chooses to enter in period 2 is:

$$\pi_B = (\pi^d - [(1-\alpha)K_B + \varepsilon_B])Pr(\pi^A \geq 0) + (\pi^m - [(1-\alpha)K_B + \varepsilon_B])[1 - Pr(\pi^A \geq 0)]$$

$$= (\pi^d - (1-\alpha)K_B)Pr(\pi^A \geq 0) + (\pi^m - (1-\alpha)K_B)[1 - Pr(\pi^A \geq 0)] - \varepsilon_A \tag{24}$$

Firm B's unconditional entry probability is therefore

$$Pr(\pi_B \geq 0) = \Phi\{(\pi^d - (1-\alpha)K_B)Pr(\pi^A \geq 0) + (\pi^m - (1-\alpha)K_B)[1 - Pr(\pi^A \geq 0)]\} \tag{25}$$

Proposition 2. *$Pr(\pi_B \geq 0)$ is an increasing function of α.*

Proof.

$$\begin{aligned}
\frac{\partial Pr(\pi_B \geq 0)}{\partial \alpha} &= \phi\{(\pi^d - (1-\alpha)K_B)Pr(\pi^A \geq 0) + (\pi^m - (1-\alpha)K_B)[1 - Pr(\pi^A \geq 0)]\} \\
&\quad \{K_B Pr(\pi^A \geq 0) + [(\pi^d - (1-\alpha)K_B]\frac{\partial Pr(\pi^A \geq 0)}{\partial \alpha} \\
&\quad + K_B[1 - Pr(\pi^A \geq 0)] - [(\pi^m - (1-\alpha)K_B]\frac{\partial Pr(\pi^A \geq 0)}{\partial \alpha}\} \\
&= \phi\{(\pi^d - (1-\alpha)K_B)Pr(\pi^A \geq 0) + (\pi^m - (1-\alpha)K_B)[1 - Pr(\pi^A \geq 0)]\} \\
&\quad [K_B + (\pi^d - \pi^m)\frac{\partial Pr(\pi^A \geq 0)}{\partial \alpha}] \\
&> 0
\end{aligned} \tag{26}$$

■

The inequality holds because we have already proved that $\frac{\partial Pr(\pi^A \geq 0)}{\partial \alpha} < 0$ and, by assumption, $\pi^m > \pi^d$. Intuitively, as α increases, firm B's entry cost decreases. At the same time, firm A's entry probability decreases, leaving firm B more likely to enjoy monopoly profit in period 2. Firm B therefore has a higher probability of entry as α increases.

Appendix C

In this appendix, we provide a simple theoretic example to show that the late arrival of a competitive market structure leads to lower internet speed. which is a proxy for product quality in broadband industry.

Consider an incumbent firm's download speed decision over an infinite time horizon. At time $t = 0$, market m contains one incumbent firm, firm i. In each subsequent period, exogenous entry occurs with probability p. Once in the market, firms do not exit. The variable profits earned by firm i in period t are given by

$$\pi_{it} = g(N_t, S_{it}) \tag{27}$$

where N_t denotes the number of firms in market m at time t and S_{it} denotes the download speed offered by firm i at time t. $N_t \in \{1, 2+\}$, representing the market as either monopolized or competitive; $S_{it} \in \{L, H\}$ where $L < H$, representing low or high speeds. In this equation, $g(\cdot)$ is a deterministic function mapping competition and download speed into variable profits and is assumed to be decreasing in N_{it}, increasing in S_{it}, and supermodular.[34] This assumption implies that a firm's profits increase with an increase in speed, and more so when the firm faces competition than when it is a monopolist. This is a reasonable assumption, since it can only poach consumers from competitors through offering higher speeds when the market is not monopolized, while the cost associated with offering higher speeds is independent of the number of competitors. Or put another way, it is reasonable to expect that an increase in competition will decrease the incumbent firm's profits by a greater amount when the incumbent offers low speeds than when it offers high speeds.

For a firm currently offering slow speeds, the one-time sunk cost of upgrading to offering high speeds is $K + \varepsilon$, where K represents the expected sunk cost and ε is a random shock assumed to be drawn *i.i.d.* from a standard normal distribution.

At time $t = 0$, firm i offers slow speeds. In each period, it can choose whether or not to upgrade

[34]Formally, a function $f : \mathbb{R}^k \to \mathbb{R}$ is supermodular if $f(x \wedge x') + f(x \vee x') \geq f(x) + f(x')\ \forall x, x' \in \mathbb{R}^k$, where $x \wedge x'$ and $x \vee x'$ denote the componentwise maximum and minimum of x and x'. If f were twice continuously differentiable, supermodularity is equivalent to the condition that $\frac{\partial^2 f}{\partial x_i \partial x_j} \geq 0\ \forall i \neq j$. In this application, supermodularity implies that $g(2+, H) + g(1, L) \geq g(2+, L) + g(1, H)$.

its speed from low to high. Upon upgrading to high speed, the firm offers high speeds in all periods thereafter. A firm offering low speeds in period t therefore chooses to upgrade to offering high speeds if and only if its increase in expected future profits exceeds the sunk cost.[35] Firms discount future profits at a discount factor, δ.

Firm i's Upgrade Probability in a Competitive Market

For a market with 2+ firms at time t, if firm i upgrades to high speed at time t, its expected flow of variable profits is

$$\begin{aligned} V(2+, H) &= \sum_{s=0}^{\infty} \delta^s g(2+, H) \\ &= \frac{g(2+, H)}{1-\delta} \end{aligned} \tag{28}$$

Similarly, firm i's expected flow of variable profits from maintaining low speed at time t is

$$\begin{aligned} V(2+, L) &= \sum_{s=0}^{\infty} \delta^s g(2+, L) \\ &= \frac{g(2+, L)}{1-\delta} \end{aligned} \tag{29}$$

Thus, firm i will upgrade to high speed if and only if

$$V(2+, H) - K - \varepsilon \geq V(2+, L) \tag{30}$$

$$\iff \varepsilon \leq V(2+, H) - V(2+, L) - K \tag{31}$$

Therefore, the probability that firm i upgrades to high speed in a competitive market is

$$Pr(H|N = 2+) = \Phi[V(2+, H) - V(2+, L) - K] \tag{32}$$

where $\Phi(\cdot)$ is the CDF of the standard normal distribution and

$$V(2+, H) - V(2+, L) = \frac{g(2+, H) - g(2+, L)}{1-\delta} \tag{33}$$

[35] For tractability, we assume that firms do not consider the option value of waiting for a better cost shock when making their decisions.

Firm i's Upgrade Probability in a Monopolized Market

For a market with 1 firm at time t, firm i must consider the likelihood of future entry when making its upgrade decision. At any point s periods into the future, the probability that market m remains monopolized is $(1-p)^{(s+1)}$. Therefore, if firm i upgrades to high speed at time t, its expected flow of variable profits is

$$\begin{aligned} V(1,H) &= \sum_{s=0}^{\infty} \delta^s \{(1-p)^{(s+1)} g(1,H) + [1-(1-p)^{(s+1)}] g(2+,H)\} \\ &= \sum_{s=0}^{\infty} \delta^s g(2+,H) + (1-p) \sum_{s=0}^{\infty} [\delta(1-p)]^s [g(1,H) - g(2+,H)] \\ &= \frac{g(2+,H)}{1-\delta} + \frac{(1-p)[g(1,H)-g(2+,H)]}{1-\delta(1-p)} \end{aligned} \tag{34}$$

Similarly, firm i's expected flow of variable profits from maintaining low speed at time t is

$$V(1,L) = \frac{g(2+,L)}{1-\delta} + \frac{(1-p)[g(1,L)-g(2+,L)]}{1-\delta(1-p)} \tag{35}$$

Thus, firm i will upgrade to high speed if and only if

$$V(1,H) - K - \varepsilon \geq V(1,L) \tag{36}$$

$$\iff \varepsilon \leq V(1,H) - V(1,L) - K \tag{37}$$

Therefore, the probability that firm i upgrades to high speed in a monopolized market is

$$Pr(H|N=1) = \Phi[V(1,H) - V(1,L) - K] \tag{38}$$

where

$$V(1,H) - V(1,L) = \frac{g(2+,H)}{1-\delta} + \frac{(1-p)[g(1,H)-g(2+,H)]}{1-\delta(1-p)} - \{\frac{g(2+,L)}{1-\delta} + \frac{(1-p)[g(1,L)-g(2+,L)]}{1-\delta(1-p)}\} \tag{39}$$

$$= \frac{g(2+,H)-g(2+,L)}{1-\delta} + \frac{(1-p)\{[g(1,H)-g(2+,H)]-[g(1,L)-g(2+,L)]\}}{1-\delta(1-p)} \tag{40}$$

$$= [V(2+, H) - V(2+, L)] + \frac{(1-p)\{[g(1,H) - g(2+,H)] - [g(1,L) - g(2+,L)]\}}{1 - \delta(1-p)} \tag{41}$$

where the final equality is obtained by substitution of equation (33).

Delayed Competition and the Probability of High Speed Provision

Now, let W denote the number of periods elapsed from time 0 until entry occurs and the market becomes competitive. The probability that firm i offers high speeds by some arbitrary future period, T, is then given by

$$\begin{aligned} Pr(H \text{ by } T) &= 1 - \{Pr(L|N=1)^W Pr(L|N=2+)^{(T-W)}\} \quad (42) \\ &= 1 - \{[1 - Pr(H|N=1)]^W [1 - Pr(H|N=2+)]^{(T-W)}\} \end{aligned}$$

Proposition 3. *$Pr(H$ by $T)$ is a decreasing function of W.*

Proof. For ease of notation, let $a \equiv 1 - Pr(H|N=1)$ and let $b \equiv 1 - Pr(H|N=2+)$. Then

$$Pr(H \text{ by } T) = 1 - a^W b^{(T-W)} \tag{43}$$

and therefore

$$\begin{aligned} \frac{\partial Pr(H \text{ by } T)}{\partial W} &= -\{a^W \log(a) b^{(T-W)} - a^W b^{(T-W)} \log(b)\} \quad (44) \\ &= -a^W b^{(T-W)}[\log(a) - \log(b)] \\ &= -a^W b^{(T-W)} \log(\frac{a}{b}) \end{aligned}$$

Thus, $Pr(H$ by $T)$ is decreasing in W if and only if

$$\log(\frac{a}{b}) \geq 0 \tag{45}$$

$$\iff a \geq b \tag{46}$$

$$\iff 1 - Pr(H|N=1) \geq 1 - Pr(H|N=2+) \tag{47}$$

$$\iff Pr(H|N=2+) \geq Pr(H|N=1) \tag{48}$$

$$\iff \Phi[V(2+,H) - V(2+,L) - K] \geq \Phi[V(1,H) - V(1,L) - K] \tag{49}$$

Since $\Phi(\cdot)$ is strictly increasing, this holds if and only if

$$V(2+,H) - V(2+,L) \geq V(1,H) - V(1,L) \tag{50}$$

$$\iff V(2+,H) - V(2+,L) \geq [V(2+,H) - V(2+,L)] + \frac{(1-p)\{[g(1,H) - g(2+,H)] - [g(1,L) - g(2+,L)]\}}{1-\delta(1-p)} \tag{51}$$

$$\iff g(2+,H) - g(1,H) \geq g(2+,L) - g(1,L) \tag{52}$$

■

which is true by definition of supermodularity, as defined in footnote 34. Note that the second line is obtained by substitution of equation (39). The intuition for this result is that since upgrading to high speeds increases profits by a greater amount in competitive markets than monopolized markets, the probability of an upgrade taking place increases with the number of competitive periods. In other words, a delay in the arrival of a competitive market structure will result in lower long-run internet speeds.