THREE EMPIRICAL ESSAYS ON UNITED STATES BROADBAND INTERNET MARKETS: QUALITY, COMPETITION, AND WELFARE

by

GABOR MOLNAR

M.S., Budapest University of Technology and Economics, l992

M.S., University of Colorado, 1995

B.S., Corvinus University of Budapest, 1997

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy
Interdisciplinary Telecommunications Program

2014

This thesis entitled:
Three Empirical Essays on United States Broadband Internet Markets:
Quality, Competition, and Welfare
written by Gabor Molnar
has been approved for the Interdisciplinary Telecommunications Program

____
Professor Douglas C. Sicker, Co-Chair

____
Professor Scott J. Savage, Co-Chair

____
Professor Lijun Chen

____
Professor Richard R. Green

____
Professor Thomas G. Thibodeau

Date______________

The final copy of this thesis has been examined by the signatories, and we
Find that both the content and the form meet acceptable presentation standards
Of scholarly work in the above mentioned discipline.

Molnar, Gabor (Ph.D., Telecommunications)

Three Empirical Essays on United States Broadband Markets: Quality, Competition, and Welfare

Thesis co-directed by Professor Scott J. Savage and Professor Douglas C. Sicker

This dissertation includes three multidisciplinary and empirical essays each focusing on a different aspect of broadband Internet access markets in the United States. The first essay investigates the effects of the number of Internet Service Providers (ISPs) and their product-type on broadband Internet quality. The second essay conducts an empirical analysis of quality competition and examines how incumbent telcos respond to competition from cable ISPs and other telco ISP market entrants. The third essay investigates whether people are willing to pay more money for real estate located in areas where high-speed broadband is available than for a property that does not offer this amenity. The analyses use data on broadband availability from the National Broadband Map, performance measurement information from the Measuring Broadband America program, and additional public and non-public data sets.

ACKNOWLEDGEMENTS

The Interdisciplinary Telecommunications Program (ITP) at the University of Colorado Boulder has provided me with a tremendous Ph.D. experience. An adventure that was intellectually stimulating, academically enriching, and genuinely enjoyable. When I got my masters degree from the ITP in 1995, I could have not imagined that I would return to the program after spending sixteen years in the telecom industry and have the opportunity to earn the first ITP Ph.D.

Faculty, friends, and family members have helped me to complete this work, and I wish to acknowledge their support and assistance. First and foremost, I want to express my heartfelt gratitude to my co-advisors, Professor Scott Savage and Professor Douglas Sicker for their excellent guidance, mentorship, and all their contributions of their time. It has been my privilege to work closely with them on challenging research projects, and they have boosted my confidence by giving me an equal voice in our work together. My gratitude for their contributions to my Ph.D. research and this dissertation is immeasurable.

I would like to thank my other dissertation committee members, Professor Lijun Chen, Dr. Richard Green, and Professor Thomas Thibodeau for their support over the past years as I moved from my initial research proposal to the final format of this dissertation. I could not have asked for a better committee; each member was supportive, diligent, and inspirational. I have genuinely enjoyed the opportunity to learn from their knowledge and experience.

I would also like to thank Professor James Prieger, Associate Professor of Public Policy at Pepperdine University. Though Jim was not a member of my Ph.D. committee, we collaborated on the second essay in this dissertation. He pushed me to raise my standards for econometric modeling and academic writing.

I also thank Shane Greenstein, William Lehr, Greg Rosston, Brad Wimmer, seminar participants at CAIDA's 3rd Workshop on Internet Economics 2012, attendees of the WEAI Annual Conference 2012 and audience at the Telecommunications Policy Research Conference 2013 & 2014 for their feedback on certain parts in this dissertation. I wish to express my gratitude to Tom Lookabaugh for his guidance and suggestions and to Marcell Pigniczki for his help with the data.

Over the past three years, I have benefited enormously from the support of the faculty and non-teaching staff of the Interdisciplinary Telecommunications Program as well as the faculty of the Silicon Flatirons Center. I extend gracious thanks to Dr. David Reed and Dr. Timothy Brown for their feedbacks at critical stages in the development of my research, and to Dean Phil Weiser for making me to feel that I am also part of the family at the Silicon Flatirons Center. I am also grateful to Professor Géza Gordos and Professor János Solymosi, who have been steady hands to steer me through my undergraduate and graduate education in Budapest, many-many years ago.

I am deeply appreciative of the ITP's Teaching Fellowship for funding the first two years of my Ph.D. studies, and the Google Faculty Research Award that

provided a source of research funding for the third year. I gratefully acknowledge the use of data from Telogical Systems and thank them for their excellent support.

I would not have contemplated this road if not for my two special friends, Pat and Dale, who were always supporting me and encouraging me with their love. This dissertation would also not be possible without their faith and inspiration. They have given me a Colorado-based family and a home away from home. I doubt that I will ever be able to convey my appreciation fully, but I owe them my eternal gratitude.

Finally, I would like to thank everyone in my family back in Hungary. Above all, I would like to thank my Father and Mother for the love they have given me and the constant encouragement they have provided me for all of my life. Without them, I would not be the person I am today. I am indebted to Viktória, my beloved wife, for her love and encouragement. Without her sacrifice, help, and understanding I would not have completed this work. I also credit our children, Eszter, Ádám and Sára. They inspire me and bring great joy to my life every day.

CONTENTS

CHAPTER I.

INTRODUCTION

This dissertation comprises three multidisciplinary and empirical essays in telecom economics. The first essay (Chapter II) investigates the relationship between market structure and broadband Internet quality. The research uses a model estimate that relates the actual speeds delivered in census block groups to the number of wireline and wireless ISPs, cost and demand conditions, and correction terms for the endogeneity of market structure. The empirical model employs a two-step control function approach. The first-step considers factors that determine profitability and market entry by wireline and wireless firms. The second-step shows how the number of wireline and wireless ISPs in the market affects the actual wireline speeds supplied to households. Chapter II is an abridged and edited version of a journal paper submission I prepared with the help of Professor Scott Savage during the years of 2013-2014.

How ISPs respond to the anticipated entry and product offerings from their rivals is an intriguing empirical question. To explore this, the second essay in Chapter III goes further and analyzes the quality competition among broadband ISPs, using National Broadband Map data for 2011-2013 for about thousand local markets in California. The research goal is to empirically analyze how incumbent ADSL (Asymmetric Digital Subscriber Line) service providers respond to competition from CLECs (Competitive Local Exchange Carriers) and cable Internet service providers. The model follows a static game theoretic approach to the profit maximization decision of a broadband provider that leads to a simple two-stage method of estimation of the structural parameters of the ISPs' profit functions. This methodology accounts for both the strategic aspect of each firm's quality decision, as well as the endogeneity problems inherent in the estimation problem. Chapter III is based on a conference paper that I wrote jointly with Professor James Prieger and Professor Scott Savage for the 42nd Research Conference on Communication, Information and Internet Policy (TPRC42) in 2014.

The third and final study in Chapter IV conducts an empirical analysis of the impact of access to the high-speed Internet on real estate values. This research explores why constant-quality house prices vary, where constant-quality house is defined as a single family residential property where structural, land, and community attributes are all held constant. This study empirically observes the differences in value across regional markets using a hedonic model. Chapter IV is an enhanced version of a conference paper that I wrote with the help of Professor Scott Savage and Professor Douglas Sicker in 2012. I presented this conference paper first at the 41st Research Conference on Communication, Information and Internet Policy (TPRC41) in 2013.

The dissertation is structured as follows. After this introductory chapter, Chapter II, Chapter III, and Chapter IV present the three empirical studies focusing on a different aspect of broadband Internet access markets in the United States. The last chapter, Chapter V, summarizes and details planned future research.

CHAPTER II.

MARKET STRUCTURE AND BROADBAND INTERNET QUALITY

1. Introduction

How do firms compete? Many studies have answered this fundamental question in industrial organization by examining the relationship between market structure and prices. Because of increased product proliferation in modern markets, deregulation of finance, transport and telecommunications services, and access to more detailed consumer- and firm-level data, a growing empirical literature has emerged on the role of differentiation among rivals. This essay in Chapter II adds to the literature by investigating empirically the effects of the number of firms and their product-type on broadband Internet quality.

Broadband quality is typically measured with surveys that ask consumers to indicate the time it takes to upload and download a file between their home network and the Internet and with the speed listed by Internet service providers (ISPs) on their promotions or regulatory submissions (Wallsten and Mallahan, 2013; Federal Communications Commission (FCC), 2011a; Greenstein and McDevitt, 2011). This approach is imperfect because self-reported survey answers are often subjective, and some ISPs may overstate the speed they can actually deliver to households.

More recent studies use web-based tests, such as Ookla, to measure actual speed with the time it takes to upload and receive a file from the home network to the Ookla server (Sundaresan et. al., 2011). While an improvement, web-based tests produce confounding results because they do not isolate the impact of home network traffic and because home computers and routers are often sub-optimally configured.

In 2011, the FCC launched field tests, which installed automated performance-measuring devices in over 7,000 US households. Located between the consumer's home network and the Internet connection, these devices provide regular speed measurements during different times of the day and year. We use these data to measure broadband quality by the amount of data that can be actually downloaded and uploaded to and from the home computer in a given time period.

Our measures distinguish between burst and sustained speeds to ensure that markets with transient speed-enhancing technologies can be objectively compared with those that do not use these technologies. We estimate a model that relates the actual speeds delivered in census block groups to the number of wireline and wireless ISPs, cost and demand conditions, and correction terms for the endogeneity of market structure. Model estimates show four main findings. Wireline speeds are often higher in markets with two or more wireline ISPs than with a single wireline ISP. Excluding the correction terms from the analysis understates this effect. Increases in wireline speeds are larger in the upstream direction, and there is no relationship between wireline speeds and the number of wireless ISPs in the market.

Theory shows that competition can result in lower or higher quality. Gaynor (2006) uses a variant of the Dorfman-Steiner (1954) condition, which relates quality expenditures to the ratio of the quality elasticity of demand to the price elasticity of demand, to explain the intuition from many of these papers. When prices are fixed, firms increase quality in an attempt to gain market share (e.g., Douglas and Miller, 1974; Schmalensee, 1977). When competition lowers prices, firms' incentives to invest in quality will also be lower. When competition increases both elasticities, quality may increase or decrease depending on the relative strengths of the two effects. For example, Kranton (2003) studied the effect of competition on quality when consumers are imperfectly informed about quality. She shows that when firms compete in price for market share, both price and quality can be lower, which is analogous to the price elasticity of demand exceeding the quality elasticity of demand.[1]

Several empirical studies have recently quantified the relationship between market structure and quality. For example, Mazzeo (2003) shows that average flight delays are shorter in less concentrated markets. Goolsbee and Petrin (2004) show that the number of cable television (TV) channels increased in response to satellite entry. Matsa (2011) shows that supermarkets facing more intense competition have more product availability. In contrast, Domberger and Sherr (1989) find no correlation between the threat of new entry and customer's satisfaction with their real estate attorney, and Prince and Simon (2013) find that flight delays for incumbents worsen in response to entry threats by Southwest Airlines. Chen and Gayle (2013) find that the non-stop flight distance for airlines

---

[1] This ambiguity has a long history in industrial organization. Chamberlin (1933) and Abbott (1955) show that firms with market power may reduce product quality to save costs and maximize their profits. Swan (1970, 1971) defines conditions under which a competitive and monopoly market introduce a product with the same level of quality but the monopoly will charge a higher price. See Katz (2013) for a recent summary of the ambiguous effects between competition and healthcare quality.

increased in markets where merging airlines did not compete *ex-ante*, and decreased in markets where they did.

This research is also related to studies of telecommunications competition. Macher et. al. (2012) estimate household demand for telecommunications and find that wireline and wireless telephony are substitutes. They do not include broadband in the household's telecommunications portfolio. Wallsten and Mallahan (2010) find that the number of wireline ISPs in the US census tracts is negatively correlated with prices, and positively correlated with the highest advertised downstream speeds. This study differs from Wallsten and Mallahan because we measure the actual speeds delivered by ISPs, analyze markets with more than three ISPs, and differentiate between wireline and wireless ISPs. Using the FCC's Form 477 data, Xiao and Orazam (2011) find that sunk costs are an important determinant of wireline entry in US zip codes. However, they are unable to distinguish between one, two or three providers, and do not estimate the direct effects of entry on market outcomes such as price and quality. Nardotto et. al. (2012) show a positive relationship between lower barriers to entry, measured by the presence of local loop unbundling, and average broadband download speeds in the UK.[2] Given there is no unbundling, their results may not directly translate to the US experience. Moreover, they do not examine upstream speeds and measure downstream speeds with an online broadband speed checker that does not isolate

---

[2] Unbundling requires the incumbent telephone company to lease their "last-mile" connection to the household to new entrants so they can compete in the final product market for broadband Internet.

the potential impact of the consumer's computer or their home network on their measured data speed.

The research presented in Chapter II contributes to this literature by offering evidence from a new measure of quality in broadband Internet markets. The key findings are that wireline providers respond to entry from wireline competitors and their resulting quality improvements emphasize upstream performance. Given wireless broadband in our sample was delivered with relatively low-bandwidth, third-generation (3G) technology, the finding of no relationship between wireline speeds and the number of wireless ISPs is not too surprising. At this stage of the product cycle, consumers likely view these product-types as more complementary than substitutable with wireline providing bandwidth and wireless providing mobility. In the short-run, this finding suggests there may be benefits from a vertical merger between a wireline and wireless provider in terms of improved quality. Of course, such benefits will be mitigated when wireline and wireless are more substitutable, and in smaller, less competitive markets where there are opportunities for bundling and complimentary foreclosures. Given its welfare potential, anti-trust officials and policy makers may also want to consider upstream quality when assessing the competitive state of broadband markets.

The next section overviews the broadband industry. Section 3 describes the empirical model, and Section 4 details the data. Results are presented in Section 5 and Section 6 concludes.

2. Industry Overview

2.1 Technology

Several different technologies provide broadband Internet access. Wireline ISPs typically provide access with cable modem, digital subscriber line (DSL) or "fiber-to-the-home" (FTTH) network infrastructure. Cable Internet is supplied by the local cable-TV operator using hybrid fiber-coax (HFC) architecture. While cable subscribers share the network with other users in their local area network, they have a higher bandwidth threshold relative to DSL. DSL is provided by the local telephone company using copper wires and a DSL access multiplexer. DSL subscribers have a dedicated connection with the telephone company's central office, but the maximum bandwidth threshold is lower than cable and the quality of the connection degrades with distance from the central office. FTTH is typically supplied by the telephone company by placing direct or shared fiber all the way to the customer premises. Fiber has more capacity than DSL or cable, offering virtually unlimited bandwidth.

Until recently, cable, DSL and FTTH could have been considered different products. They employed different technologies, which affected their costs, pricing, and performance. They also differed in the way they were regulated, how they billed their services, and the customers they targeted. Cable operators began providing high-speed Internet in 1995 without any regulatory obligation to share their network infrastructure to rivals. In contrast, incumbent telephone companies faced unbundling regulations intended to encourage entry by DSL service providers and increase subscribership. The provision of DSL has at times also involved up to three separate entities, the incumbent telephone company, the DSL retailer and the web portal. Cable Internet is a “one-stop shop” service. While DSL targeted small businesses with a higher-price offering that included multiple IP addresses, cable Internet had been positioned as a residential product (Gillett and Lehr, 1999). However, by 2011, offerings from these wireline ISPs had many similarities. Specifically, all technologies were largely symmetrically regulated and provided a high-bandwidth, one-stop shop service for residential and small-business customers with preferences for speed.

Wireless ISPs use satellite, fixed terrestrial microwave, and mobile technologies to offer high-speed Internet, and they all assume over-the-air data transmission at some point between the ISP and the end-user. Wireless technologies have not traditionally offered data speeds approaching those of wireline, but they do provide mobility and, as such, have been marketed to consumers who have wireline service. In contrast, satellite and terrestrial wireless broadband services have been marketed mainly to rural consumers who did not have access to wireline service, and they could be considered to be different products.

An ISP’s network capacity is smaller than the sum of the individual speeds sold to their customers. Over-subscription can result in network congestion, lower speeds and customer complaints. Under-subscription will improve quality but excess capacity is costly. ISPs typically build the minimum capacity required to handle their subscribers’ forecasted traffic needs. When it becomes congested, managers can “tweak” the network or perform a minor or major upgrade. Tweaks are small, low-cost modifications, such as using more strict traffic management practices or redefining capacity ratios, that typically result in a five to 15 percent improvement in capacity and, as such, provide ISPs with some short-run flexibility to affect the actual speeds delivered to customers. A minor upgrade would involve, for example, the allocation of more channels in the cable network for data service so there are less cable modems per channel. This could be completed in a few months.

Alternatively, the ISP may consider a major upgrade such as adding fiber capacity to the Internet backbone, more efficient use of the spectral capacity within the existing network, or re-architecting the last-mile connection to the home. For example, an upgrade from the data over cable service interface specification (DOCSIS) 3.0 to the DOCSIS 3.1 standard can dramatically increase downstream capacity by using more efficient modulation and by increasing the total available spectrum for service delivery, but may take up to two years to complete.[3]

2.2 Competition

While wireline ISPs compete along several non-price dimensions, such as product bundling, security and support services, speed is arguably the most important feature to consumers (Rosston, 2009; Pew, 2010; Greenstein and McDevitt, 2011; Greenstein and Prince, 2013; Nevo, et. al.; 2013). *Figure 1* shows the price and speed from the median service plan provided by the 17 largest ISPs in 27 US cities. Plans

---

[3] The last-mile to the home from cable and DSL are different. The last mile for cable is shared between multiple homes and the last mile for DSL is not. Splitting cable nodes can decrease the number of homes per node and increase per-subscriber capacity. DSL providers can deploy more switches or use local loop extenders to move switch intelligence closer to customers to alleviate the degradation of speed due to distance.

with promotional discounts and plans that bundle Internet with telephone and/or subscription TV are omitted.[4]

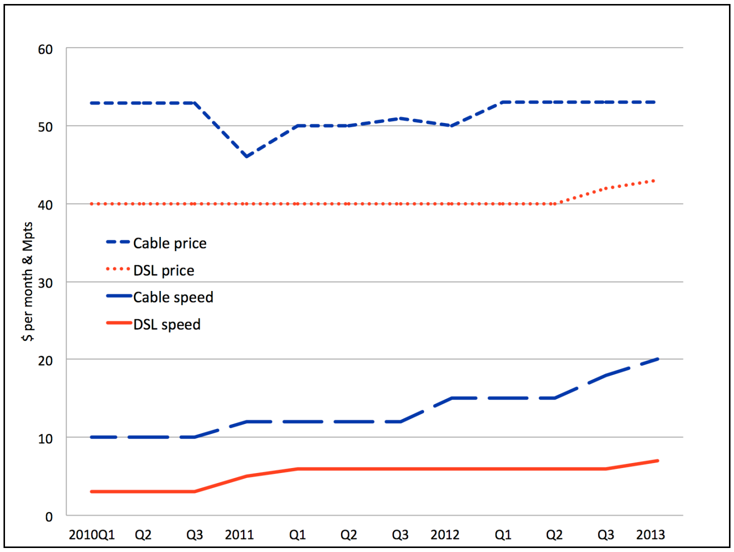


Figure 1: Median price and download speed 2010-2013

The price of the typical wireline broadband Internet service remained relatively flat during this period, but advertised upload speeds doubled. The median price for cable-modem service was $52.95 per month for an upload speed of ten megabits per second (Mbps) in January, 2010 and $52.99 for 20 Mbps in December, 2012. The median price for DSL service was $39.95 for three Mbps in 2010 and $43 for seven Mbps in 2012. Moreover, just like on-time arrival in the airline industry, ISPs will try to attract customers with marketing that claims they deliver faster actual speeds than their rivals. For example, Comcast have advertised that their XFINITY service “is the most blazingly fast thing around”, whereas Verizon badmouthed Cablevision with

---

[4] The Leichtman Research Group (2014) estimated that 60 percent of households received a bundle of TV, Internet, and/or telephone services from one company in 2014. It is difficult to analyze the prices of Internet services that are part of a bundle as we may compare apples with oranges. Typically, ISPs offer multiple double-play and triple-play bundles per market, bundle prices and bundle components vary, and package discounts are often not broken down to individual service components.

advertisements that stated that "consumers deserve to get the broadband speeds they are promised, but they don't with Cablevision."[5] The FCC (2013) benchmarks ISPs by publishing SamKnows data on actual speeds. During September, 2012, they reported that, on average, ISPs were delivering between 81 to 137 percent of the speeds promised in their advertising.

Table 1 provides a simple example of the potential relationship between wireline competition and broadband Internet quality. Besides their name, Park County, Colorado and Park County, Montana are similar in terms of geography and population. They are both in the mountain west region of the US and contain about 16,000 persons dispersed over 2,211 and 2,813 square miles, respectively. Using Ookla data to measure broadband quality, the median actual download speed in Park Count, Montana at June 2011, where there are four providers, Bresnan Communications, Bridgeband Communications, Century Link and Triangle Telephone, was 2.76 Mbps. In Park County, Colorado, where there are only two providers, Century Link and Integra Telecom Holding, the median actual download speed was 1.02 Mbps.

	Park County, CO	**Park County, MT**
Population	16,122	15,810
Area (square miles)	2,211	2,813
Median download speed (Mbps)	1.02	2.76
Number of wireline ISPs	2	4

Table 1: Comparison of two counties

---

[5] Cablevision sued Verizon for misrepresenting their speeds. FCC (2012) data from October, 2011 show that subscribers to Cablevision's 15 Mbps service received average actual download speeds during peaks hours that were about 90 percent of the advertised speed, compared to 59 percent in March, 2011. This increase in performance over a relatively short period likely resulted from tweaking and minor upgrades to the network.

FCC (2011a) data shows that the number of wireless connections quadrupled from about 26 million in December, 2008 to 119 million in June, 2011, while the number of wireline connections grew modestly from about 76 to 87 million. Wireless technologies have not traditionally offered data speeds approaching wireline broadband. At June, 2011, fewer than five percent of wireless connections had downstream speeds above 1.5 Mbps compared to 35 percent for wireline. Moreover, although about 77 percent of the population lives in an area served by three or more 3G wireless ISPs, service quality can often be poor because of relatively low signal strength or data speed, or in-building coverage. These data suggest consumers likely view these product-types as more complementary than substitutable, with wireline providing speed and wireless providing mobility. However, recent technical enhancements in fourth-generation (4G) cellular technology permit wireless ISPs to directly compete with wireline ISPs. For example, at February, 2013 CLEAR provided a wireless Internet service for $49.99 per month with download speeds ranging from three to six Mbps. How wireline ISPs respond to the anticipated entry and product offerings from their rivals is an interesting empirical question and the focus of the remainder of this chapter.

3. Empirical Model

The empirical model employs a two-step control function approach to investigate the relationship between market structure and broadband Internet quality. The first step estimates factors that determine profitability and market entry by wireline and wireless ISPs. Estimated parameters are then used to construct correction terms similar to the inverse Mills ratio in two-step selection models by Heckman (1979), Mazzeo (2002) and Manuszak and Moul (2008). The second-step estimates show how the number of wireline and wireless ISPs in the market affect the actual wireline speeds supplied to households. We can then see whether, conditional on cost and demand conditions and the correction terms for the endogeneity of market structure, there is a systematic relationship between broadband quality and the number of firms in the market, and whether this relationship varies by product-type.

We assume two product-types, wireline and wireless ISPs, that tradeoff speed. We treat ISPs within each product-type as homogenous even though they may use different technologies to provide service.[6] The wireline broadband quality equation for market $j = 1, 2, \ldots, J$ is:

$$Q_j^* = X_j\beta + g(W_j, M_j; \beta_W, \beta_M) + u_j \qquad \text{(II-1)}$$

where Q_j^* is the latent (unobserved) average actual speed delivered by wireline ISPs in the market, X is a vector of market-specific variable cost and demand factors, $g(.)$ is a function that describes the effects on quality from competition, $W_j = w_j$ ($w_j = 1, 2, 3, \ldots$) indicates the number of ISPs in the market that provide residential wireline broadband Internet connections, $M_j = m_j$ ($m_j = 1, 2, 3, \ldots$) indicates the number of ISPs in the market that provide residential wireless (or "mobile") broadband Internet connections, β measures the relationship between quality and cost and demand factors, β_W measures the relationship between quality and wireline competition, β_M measures the

---

[6] We control for technology when estimating the quality equation. See *Section 4.2* for discussion.

relationship between quality and wireless competition, and u_j is a random error term that reflects unobserved factors in market *j* that affect quality.

We address two econometric issues during estimation of equation (II-1). The first concerns unobservables. ISPs consider the effects on expected profits from the anticipated entry of competitors when making decisions about entry, geographical and product location, and quality. It is likely that unobserved cost and demand factors that affect market structure also affect quality decisions over the short and medium run so that $E[u_j \mid W_j, M_j] \neq 0$. For example, a market with lower unobserved demand due to high quality public access at libraries, museums, post offices and/or work places is likely to attract fewer entrants. However, because ISPs compete against these public options, it is also possible that they provide higher quality and OLS estimates of and β_W and β_M will have negative bias. Alternatively, a market with higher unobserved costs will attract fewer entrants who may also provide lower quality since it is expensive. Here, OLS estimates of estimates of β_W and β_M will have positive bias.

We account for the potential endogeneity of market structure with a two-step, control function approach. The first-step describes the number of firms in the market with payoff functions that relate the expected profits for wireline and wireless ISPs, respectively, to market size, cost and demand conditions, and anticipated competition. The expected profits in market $j = 1, 2, \ldots, J$ for "high-speed" (*H*) wireline ISPs and for "low-speed" (*L*) mobile ISPs are specified as:

$$\Pi^*_H(Z_j, W_j, M_j, e_{Hj}; \tau_H) = \pi_H(Z_j, W_j, M_j; \tau_H) + e_{Hj}$$

$$\Pi^*_L(Z_j, W_j, M_j, e_{Lj}; \tau_L) = \pi_L(Z_j, W_j, M_j; \tau_L) + e_{Lj} \qquad \text{(II-2)}$$

where Π^*_H is the unobserved continuous index of market profits for wireline ISPs, Π^*_L is the unobserved continuous index of market profits for mobile ISPs, $Z_j = [S_j, X_j, Y_j]$ is a vector of market-specific factors that affect the profitability of the market, S_j is market size, Y_j is a vector of fixed costs, e_{Hj} is a random error term that reflects unobserved factors in market j that affect the profits for wireline ISPs, and e_{Lj} is a random error term that reflects unobserved factors in market j that affect the profits for wireless ISPs.

Realistically, a wireline and wireless ISP cannot enter the other's business. As such, the appropriate choice for a wireline ISP is to enter (stay) or not enter (exit) as a predetermined wireline firm, and the choice for a wireless ISP is to enter (stay) or not enter (exit) as a predetermined wireless firm. Each firm makes its entry decision market by market and in the Nash equilibrium firms will continue to enter a market until it is no longer profitable. Because no ISP can switch product-types, the equilibrium number of firms in market j is characterized by standard profit inequalities:

$$W_j = W_j^* : \pi_H(Z_j, W_j^*, M_j; \tau_H) + e_{Hj} > 0 \text{ and } \pi_H(Z_j, W_j^* + 1, M_j; \tau_H) + e_{Hj} < 0$$

$$M_j = M_j^* : \pi_L(Z_j, W_j, M_j^*; \tau_L) + e_{Lj} > 0 \text{ and } \pi_L(Z_j, W_j, M_j^* + 1; \tau_L) + e_{Lj} < 0$$

The conditions are an equilibrium because each firm is playing its best response given the choices of other firms and market characteristics. The best response for ISPs inside the market is to provide service, and the best response for ISPs outside the market is not to.

We assume that the error terms u_j, e_{Hj}, and e_{Hj} are jointly normally distributed with mean vector zero and covariance matrix:

$$\Sigma = \begin{bmatrix} \sigma_u^2 & \sigma_{ue_H} & \sigma_{ue_L} \\ & 1 & 0 \\ & & 1 \end{bmatrix} \qquad \text{(II-3)}$$

where the variances of e_{Hj}, and e_{Hj} are normalized to one.[7] Given these assumptions, the conditional expectation of quality can be written as:

$$E[Q_j^* \mid X_j, W_j, M_j, Z_j] = X_j\beta + g(W_j, M_j; \beta_W, \beta_M) + \sigma_{ue_H}\lambda_{Hj} + \sigma_{ue_L}\lambda_{Lj} \qquad \text{(II-4)}$$

where

$$\sigma_{ue_H}\lambda_{Hj} + \sigma_{ue_L}\lambda_{Lj} = \sigma_{ue_H}\frac{\varphi(\bar{\pi}_{Hj}^{w}) - \varphi(\bar{\pi}_{Hj}^{w+1})}{\Phi(\bar{\pi}_{Hj}^{w+1}) - \Phi(\bar{\pi}_{Hj}^{w})} + \sigma_{ue_L}\frac{\varphi(\bar{\pi}_{Lj}^{m}) - \varphi(\bar{\pi}_{Lj}^{m+1})}{\Phi(\bar{\pi}_{Lj}^{m+1}) - \Phi(\bar{\pi}_{Lj}^{m})}$$

$\phi(.)$ is the standard normal density and $\Phi(.)$ is the standard normal distribution. We use an ordered probit model to estimate the profit functions for wireline and wireless ISPs, respectively, and use the linear predictions ($\bar{\pi}$) from this first-step to get consistent estimates of λ_{Hj} and λ_{Lj}. We then add the correction terms to the quality equation (II-1) and specify the modified quality equation (II-5) in step-two:

$$Q_j^* = X_j\beta + g(W_j, M_j; \beta_W, \beta_M) + \sigma_{ue_H}\lambda_{Hj} + \sigma_{ue_H}\lambda_{Lj} + v_j \qquad \text{(II-5)}$$

---

[7] The assumption of zero correlation between the errors from wireline and wireless profits helps the convergence of the ordered probit models in Section 5. This assumption is problematic if any unobserved correlation is large and is captured by the estimated parameters on *W* and *M*. We were able to estimate a more parsimonious model with correlation that did not separate the effects from own product-type entry on variable profits and fixed costs. The correlation from this specification of 0.118 was reasonably close to zero and the estimated profit parameters were qualitatively similar to those reported in *Table 7*.

Since $v_j \equiv u_j - \sigma_{ue_H}\lambda_{Hj} - \sigma_{ue_L}\lambda_{Lj}$ is mean zero conditional on X_j, M_j, W_j and Z_j, unobserved factors correlated with market entry and quality are controlled for and OLS estimates of β_W and β_M should be unbiased and consistent.

A second econometric issue arises because we do not observe the true value of average quality for the market. Instead, we observe quality from an individual household draw (q_j) within each market. Let $q_j = Q_j^* + \xi_j$, where ξ is random measurement error.[8] Substituting this expression into equation (II-5) gives:

$$q_j = X_j\beta + g(W_j, M_j; \beta_W, \beta_M) + \sigma_{ue_H}\lambda_{Hj} + \sigma_{ue_L}\lambda_{Lj} + \varepsilon_j \qquad \text{(II-6)}$$

where $\varepsilon_j = v_j + \xi_j$. Under the assumption that the measurement errors are random and are not correlated with W and M, the consequence from OLS estimation of equation (II-6) is that the standard errors of the estimated coefficients will be inflated. If the measurement errors are correlated with W or M, OLS estimates of β_W and β_M will also be biased. However, since the number of wireline and wireless ISPs is already treated endogenous in our analysis, potential bias from measurement error should not be a serious problem. In Section 4, we also describe additional controls for measurement error when estimating the quality equation.

---

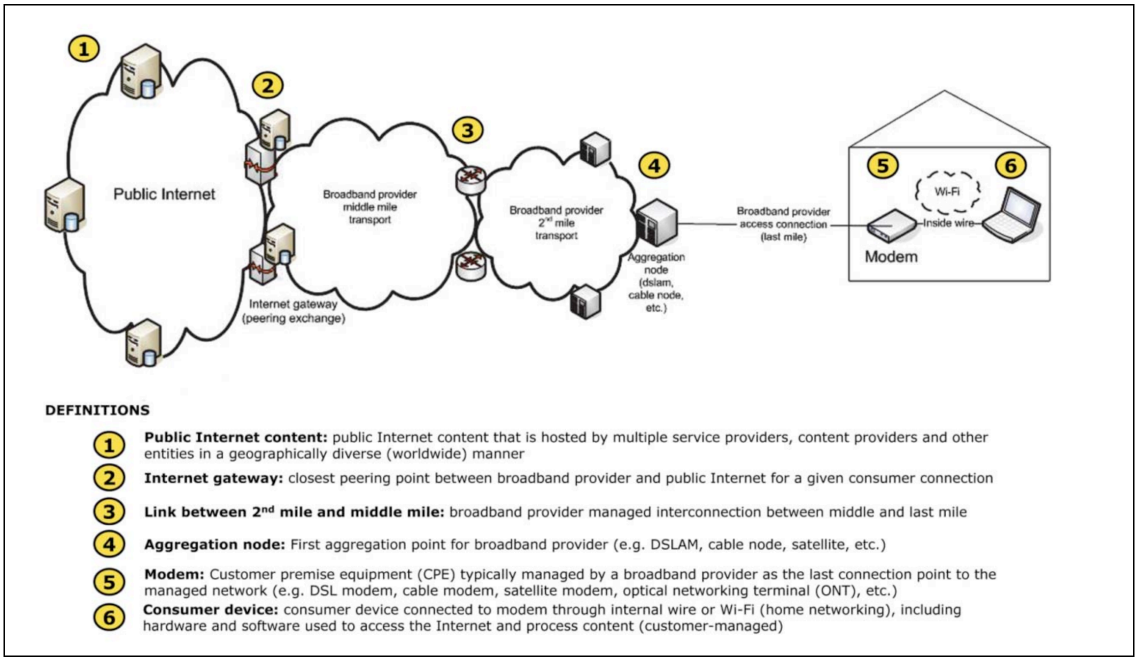
[8] When collecting data on actual speed, the SamKnows sampling method emphasizes breadth across US census block groups rather than depth within a census block group. Ideally, we would like 30 or more household observations within each census block group to calculate average quality (Q). More household observations can be obtained when the quality and market structure measures are aggregated to a higher level, say county or state, but the market definition would be too broad so that the number of firms is imprecisely measured. For example, there may be four ISPs operating in the county overall but perhaps only one ISP in many census block groups that comprise the county.

4. Data

4.1 Sample

Household data from the first US field test of wireline broadband performance were obtained from the FCC (2011b) for March, 2011. The main objective was to measure the actual downstream and upstream broadband speeds delivered by ISPs to households while excluding the effects of household network computers and traffic. The field tests are operated by SamKnows, a broadband quality test form, based in the United Kingdom. More than 78,000 households initially volunteered to participate in the test and in the first phase 7,377 were selected as participants in the final sample. The FCC's Measuring Broadband America program selected the sample with the goal of covering the major ISPs, geographical locations, and broadband technologies across the U.S. Households were sampled from each census subdivision, Pacific, Mountain, West North Central, East North Central, West South Central, East South Central, South Atlantic, Mid Atlantic and New England, to represent broadband performance in three speed ranges: fewer than three Mbps; between three and ten Mbps; and greater than ten Mbps. The participating ISPs from which wireline speeds were measured are AT&T, Brighthouse, Cablevision, CenturyLink, Charter, Comcast, Cox, Frontier, Mediacom, Insight, Qwest, TimeWarner, Verizon and Windstream. At March, 2011, these ISPs served about 90 percent of the national wireline Internet market (Leichtman Research Group, 2011). Participating households installed a speed measurement device, which came pre-loaded with proprietary testing software. Due to the small number of volunteers for satellite and fixed-wireless services, limited data was collected on these technologies, and they are not considered in this analysis. Households with mobile broadband services were also excluded. As such, the empirical analysis of residential broadband Internet quality in Chapter II will only consider actual speeds provided by wireline ISPs, either cable, DSL or FTTH.

Figure 2 shows one of the main strengths from the field test data. Because the data measure the connection parameters from the content source to the consumer, any bandwidth limitations or delays incurred in the consumers' home or in segments of the Internet outside an ISP's network are not reflected in our analysis. Moreover, the performance-measuring program is open and transparent.



DEFINITIONS

① **Public Internet content:** public Internet content that is hosted by multiple service providers, content providers and other entities in a geographically diverse (worldwide) manner

② **Internet gateway:** closest peering point between broadband provider and public Internet for a given consumer connection

③ **Link between 2nd mile and middle mile:** broadband provider managed interconnection between middle and last mile

④ **Aggregation node:** First aggregation point for broadband provider (e.g. DSLAM, cable node, satellite, etc.)

⑤ **Modem:** Customer premise equipment (CPE) typically managed by a broadband provider as the last connection point to the managed network (e.g. DSL modem, cable modem, satellite modem, optical networking terminal (ONT), etc.)

⑥ **Consumer device:** consumer device connected to modem through internal wire or Wi-Fi (home networking), including hardware and software used to access the Internet and process content (customer-managed)

Figure 2: Network diagram (FCC 2011b, p. 9)

The FCC's Measuring Broadband America program publishes the method used to collect the data and releases all data within one year of collection (FCC 2011b).

Speed data are collected using three, 30-second simultaneous transmission control protocol (TCP) connections to a test node. The total throughput of the connection, measured in Mbps, was defined as the sum of the data transmitted in bytes divided by the time elapsed in seconds. TCP slow start, a warm up period due to the nature of the communication protocol, and congestion were taken into account and all three connections were required to have completed the warm up period before the real testing began (SamKnows, 2012). A potential disadvantage of the speed test data is that it only samples from the most popular service tiers within an ISP's offerings. While these data may best represent the average US household, they contain no information on the actual speeds from some of the less popular offerings. Furthermore, because the FCC did not collect demographics, it is not possible to test the representativeness of the actual speeds sampled from households. However, because we know each household's geographical identifier, it is possible to examine the representativeness of the sample census block groups in Section 4.3.

4.2 Variables

Previous studies have used census tract, county, cable franchise, local telephone exchange and zip-code boundaries to define the geographical market for broadband Internet (Gillett and Lehr, 1999; Prieger, 2003; Wallsten and Mallahan, 2010; Chen and Savage, 2011; Xiao and Orazem, 2011; Nardotto et. al., 2012). Because ISP decisions to enter and roll out new services are usually made for smaller geographical footprints, and because our firm data can be measured for each CBG, we define the market for broadband Internet to be a CBG, which generally contains between 600 and 2,400 people.[9] This definition is also empirically useful as it includes markets with one or two firms. In contrast, larger geographical markets will almost always have at least two or three firms.[10] We use SamKnows (2012) speed test data from March, 2011 to construct four different measures of Internet quality and match these with information on the number of wireline and wireless ISPs in each of the sampled CBGs. As noted in previous studies, a nice feature of these data is that the markets are well defined. Within-market customers cannot move to a neighboring market to buy a better plan as this would imply moving house.

Quality. Computer scientists define service quality as the ability to provide different priority to different applications, users, or data flows, or to guarantee a certain level of performance to a data flow. Because the Internet is a best-effort network, broadband quality is also often associated with advertised speeds on ISP's promotional materials. However, advertised speeds are an imperfect measure of service quality because they do not represent the actual speed delivered to the customer. It is also appropriate to distinguish between burst and sustained speeds to ensure that markets with transient speed-enhancing technologies, such as

---

[9] Discussions with former telco and cable executives indicate that the decision to provide and upgrade high-speed Internet service by neighborhood is the common industry practice. This is consistent with CenturyLink's recent deployment of its gigabit service to selected neighborhoods within the metropolitan area of Denver (Vuong, 2014). Similarly, Google has defined "fiberhoods" as consisting of 250 to 1,500 households and deploy their fiber service when the demand threshold reaches five to 25 percent of these households (Lardinois, 2012). Customer reviews also indicate that certain neighborhoods often get deployed or upgraded to higher speed before others. For example, see http://www.dslreports.com/comments/2170, http://www.broadbandexpert.com/high-speed-internet-reviews/att/, and http://tech.slashdot.org/story/14/03/01/0322206/how-i-cut-my-time-warner-cable-bill-by-33.

[10] Because cable operators provide service in neighboring CBGs, entry and quality decisions may be correlated across larger geographical areas. We re-estimated all quality equations with a robust estimate of the variance-covariance matrix that accounted for within-county correlation between errors. Model results, not presented here, are qualitatively similar to those reported in Tables 9 through 11.

Powerboost®, can be objectively compared with those that do not. Burst speed, the maximum speed an ISP is capable of delivering over a short time frame, is important for web surfers or gamers who need to transmit relatively small amounts of information in bursts. Sustained speed, the maximum speed an ISP is capable of over a longer time frame, is important to video streamers because they require a consistent transmission rate over a longer period of time. Historically, downstream speed has been the most important Internet feature for residential users, reflecting preferences for video, web surfing and shopping. ISPs have catered for these preferences with technologies and plans that deliver higher downstream speeds relative to the upstream. However, upstream performance is becoming increasingly important as consumers upload files to the cloud, run their own web servers from home, become active in social networking, and use business applications to work at home while maintaining remote access to their place of employment.

Accordingly, the four measures of quality used in this study are burst and sustained speeds in the upstream and downstream directions. Following the FCC (2011b) burst speed is measured by the data throughput during the first five seconds of the test and sustained speed as the throughput at the 25 to 30 second interval of the test. Households with missing data on performance and/or other aspects of their Internet plan are omitted from the sample and monthly trimmed averages are calculated for burst and sustained speeds for March, 2011 for the entire day, peak period and off-peak period. The peak period is seven to eleven *p.m.* local time, and the off-peak period is three to seven *a.m.* local time. The average number of tests per household for an entire day was 262. The average number for the peak and off-peak periods of the day were 42 and 46, respectively.

Market structure. Data from the National Broadband Map (NBM; 2011) were used to count the number of ISPs in each census block group at June, 2011. Created from a collaboration between the National Telecommunications and Information Administration, the FCC, and all states, territories and Districts of the US, the NBM is an online tool that provides semi-annual information on the ISPs, their product-type, technology and their maximum advertised upload and download speeds in each US census block group. We identify the census block group that each sample household is located in and then count the number of wireline and wireless ISPs in that census block group. Because wireline ISPs use a fixed wired network to send and receive information to and from the Internet, W is the number of cable, DSL and FTTH ISPs in the market. Because wireless ISPs use mobile non-wire technologies somewhere in their network to send and receive information, M is the number of fixed terrestrial microwave and mobile wireless ISPs in the market.[11]

Market size, cost and demand. Data from the 2010 population census and the 2006-2010 American Community Survey were used to construct market-level controls for the size of the market, costs and demand conditions. Market size (S) is population. The vector of variable cost and demand variables (X) that characterize per capita-variable profits includes: mean household income (*INCOME*); mean age in years

---

[11] Greenstein and Mazzeo (2006) and Xiao and Orazem (2011) note that in telecommunication markets with many firms, the marginal effect of additional competitors is likely to be small. We measure the number of wireline ISPs with $Wj = wj$ ($wj = 1, 2, \dots , 7$ or more) and the number of mobile wireless ISPs with $Mj = mj$ ($mj = 1, 2, \dots , 7$ or more). About a percent of the markets in our sample have eight or more wireline ISPs and about seven percent have eight or more wireless ISPs. This measure of firms also permits easier estimation of the first-step latent profit functions by maximum likelihood.

(*AGE*); mean number of years of schooling for the population over 25 years of age (*EDUC*); standard deviation of the number of years of schooling for the population over 25 years of years of age (*SD EDUC*); and square kilometers of geographical area (*AREA*). *INCOME*, *AGE*, *EDUC* and *SD EDUC* measure preferences and *AREA* measures variable costs. The vector *X* also includes an electricity regulation index (*REG*) that approximates the price of electricity in each state, which is an important broadband production input. The index equals zero when the state has no competition in electricity markets, 0.5 when the state has wholesale competition in electricity markets, and one when the state has both wholesale and retail competition.

Bresnahan and Reiss (1991) model entry by assuming firms require their fixed entry expenses to be covered by variable profits to generate a sufficient rate of return. Previous studies have found that density variables explain significant variation in the fixed costs of deploying and maintaining the equipment used to produce telecommunications services (Shin and Ying, 1992; FCC, 2000; Greenstein and Mazzeo, 2006, Zager, 2011). We use *ROADS* (the number of road miles) to control for the size of the outside plant used to connect the backbone network to each subscriber's household. *BEDROCK* (the coefficient of variation of the average depth at which bedrock is first encountered across the CBG), *WETLANDS* (percentage of wetlands), and *INTERSECTIONS* (the number of road intersections) control for the physical constraints that make the deployment and maintenance of equipment more difficult.[12] Because managing a large telecommunications deployment is generally more cost

---

[12] Areas with more road intersections may also have higher costs because it is more likely that the ISP will have to coordinate and negotiate the use of existing, congested rights of way with other public utilities.

effective than managing a smaller deployment, *HOUSES* (number of houses) controls for potential economies of scale.

Other controls. Because the dependent variable (q) is measured with error, we employ several capacity, technology, brand and regional variables as additional controls for any systematic variation in ξ in the quality equation. The capacity variables are: the mean maximum advertised downstream speed (Mbps) of all ISPs in the CBG that provide residential wireline broadband Internet connections (*BANDDOWN*); the mean maximum advertised upstream speed (Mbps) of all ISPs in the CBG that provide residential wireline broadband Internet connections (*BANDUP*); the household's maximum advertised downstream speed (Mbps) for their service plan (*ISPDOWN*); and the household's maximum advertised upstream speed for their service plan (*ISPUP*).[13] The technology variables are: an indicator that equals one when the household's Internet connection is cable, and zero otherwise (*CABLE*); and an indicator that equals one when the household's connection is FTTH, and zero otherwise (*FTTH*). The brand controls are firm-specific dummy variables corresponding to the b = 2, 3, … , 17 ISPs that provide wireline Internet connections to the household in our sample ($BRAND_b$). The regional controls are region-specific dummy variables corresponding to the r = 2, 3, … , 9 census subdivisions covered by the SamKnows sampling approach ($REGION_r$).[14] By including these additional variables in equation (II-6) we "homogenize" the Internet service from with speed is recorded by controlling

---

[13] *ISPDOWN* and *ISPUP* are maximum downstream and upstream speeds for the household's service tier when they signed on to the SamKnows panel. Because they are less reliably measured than actual speed, we do not use them to construct a quality variable that is the difference between actual and advertised speeds.

[14] For robustness, we also estimated equation (II-6) with state-specific dummy variables instead of regional-specific dummies. The results, not reported, are similar to those reported in Table 7.

for service-specific effects they may be correlated with actual speed measurements, e.g., household distance from the central office.

4.3. Summary statistics

Because some of the households and census block groups in our gross sample have missing or incomplete data, our final sample is comprised of 5,281 census block groups. *Table 2* describes the variables used in the empirical analysis. Firm counts, presented in *Table 3*, show that all markets in our sample are served by at least one wireline ISP and at least one wireless ISP. This is a function of the SamKnows sampling approach which only records speeds in markets where broadband is available.

Variable	**Description and data source**
W	Number of ISPs in the census block group that provide residential wireline broadband Internet connections at June 2011. Source: NBM (2011).
M	Number of ISPs in the census block group that provide residential wireless ("mobile") broadband Internet connections at June 2011. Source: NBM (2011).
q (burst)	Data throughput (Mbps) during the first five seconds of the household's speed test. Source: FCC (2011b).
q (sustained)	Data throughput (Mbps) during the 25 to 30 second interval of the household's speed test. Source: FCC (2011b).
S	Number of persons in the census block group. Source: GeoLytics (2012).
INCOME	Mean household income ($1,000) for all households in the CBG. Source: GeoLytics (2012).
AGE	Mean age in years of the population in the census block group. Source: GeoLytics (2012).
EDUC	Mean number of years of schooling of the population over 25 years of age in the census block group. Source: GeoLytics (2012).
SD EDUC	Standard deviation of the number of years of schooling of the population over 25 years of age in the census block group. Source: GeoLytics (2012).
AREA	Geographical area (km^2) of the census block group. Source: GeoLytics (2012).
REG	One when the census block group is located in a state with wholesale and retail competition in electricity markets; 0.5 when the census block group is located in a state with wholesale competition in electricity markets; and zero otherwise. Source: Craig and Savage (2013).

CPDEN	Number of persons per km^2 in the county the census block group is located.. Source GeoLytics (2012).
BANDDOWN	Mean maximum advertised downstream speed (Mbps) of all ISPs in the census block group that provide residential wireline broadband connections at June 2011. Source: FCC (2011b).
BANDUP	Mean maximum advertised upstream speed (Mbps) of all ISPs in the census block group that provide residential wireline broadband Internet connections at June 2011. Source: FCC (2011b).
ISPDOWN	Advertised downstream speed (Mbps) of the household's Internet connection. Source: FCC (2011b).
ISPUP	Advertised upstream speed (Mbps) of the household's Internet connection. Source: FCC (2011b).
CABLE	One when the household has a cable modem Internet connection; and zero otherwise. Source: FCC (2011b).
FTTH	One when the household has a fiber to the home Internet connection; and zero otherwise. Source: FCC (2011b).
BRAND$_b$	One when the household's Internet connection is supplied by ISP b = 2, 3, … , 17; and zero otherwise. Source: FCC (2011b).
REGION$_r$	One when the census block group is located in census subdivision r = 2, 3, … , 9; and zero otherwise. Source: US Census Bureau (2012).
ROADS	The total length in miles of S1100 (primary), S1200 (secondary), and S1400 (local) roads in the census block group. Source: U.S. Census Bureau (2013c).
INTERSECTIONS	The number of S1100 (primary), S1200 (secondary), and S1400 (local) road intersections in the census block group. Source: U.S. Census Bureau (2013c).
BEDROCK	The average distance from the soil surface to the top of a bedrock layer in the census block group (in cm). Source: USDA (2013).
WETLANDS	The total area of wetlands in the census block group (in % of total land area). Source: GeoLytics (2012).
HOUSES	The number of houses in the census block group. Source: GeoLytics (2012).

Table 2: Variable descriptions

Wireline			**Wireless**	
Firms	Freq.	Cum.	Firms	Freq.
1	94	1.78	1	94
2	1,713	32.44	2	1,713
3	1,826	68.82	3	1,826
4	1,052	88.70	4	1,052
5	380	95.89	5	380
6	111	97.98	6	111
7 or more	105	100.00	7 or more	105
Total	5,281		Total	5,281

Table 3: Firms per observed market (NBM, 2011)

Table 4 presents summary statistics for all the variables used in our empirical analysis. On average, the typical census block group market has 1,805 persons in a geographical area of 10.04 square kilometers. Average age per market is about 37.9 years, average number of years of schooling is just over 13 and average household income is $68,509.

Variable	Obs.	Mean	s.d.	Min	Max
W	5,281	3.121	1.171	1	7
M	5,281	5.041	1.243	1	7
S	5,281	1,805	1,233	1	28,537
INCOME	5,281	68.51	27.99	5.630	192.2
AGE	5,281	37.90	5.440	16.65	82
EDUC	5,281	13.27	1.486	5.43	17.8
SD EDUC	5,281	4.610	1.774	0	19.04
AREA	5,281	10.04	39.81	0.008	2,097
REG	5,281	0.599	0.384	0	1
HDEN	5,281	942.6	2,354	0.013	40,398
CPDEN	5,281	2,075	6,379	3.167	69,468
BANDDOWN	5,281	26.11	30.04	0.768	960.4
BANDUP	5,281	7.352	27.32	0.200	960.1
q (burst down: all day)	5,649	15.37	10.31	0.608	94.21
q (burst up: all day)	5,646	3.286	6.377	0.110	43.71
q (sustained down: all day)	5,649	11.71	8.594	0.709	80.44
q (sustained up: all day)	5,646	3.088	6.419	0.126	43.77
q (burst down: peak)	5,633	13.82	9.862	0.618	84.88
q (burst up: peak)	5,620	3.269	6.356	0.114	43.73
q (sustained down: peak)	5,631	11.06	8.469	0.651	75.31
q (sustained up: peak)	5,617	3.086	6.422	0.120	43.75
q (burst down: off peak)	5,610	16.50	10.95	0.574	101.8
q (burst up: off peak)	5,597	3.307	6.393	0.107	43.73
q (sustained down: off peak)	5,608	12.09	8.750	0.700	71.17
q (sustained up: off peak)	5,591	3.109	6.440	0.130	43.77
ISPDOWN	5,649	11.80	8.077	0.768	100
ISPUP	5,645	2.831	5.834	0.128	35
CABLE	5,281	0.585	0.493	0	1
FTTH	5,281	0.082	0.274	0	1
ROADS	5,281	14.85	27.64	0	1,067
INTERSECTIONS	5,281	63.20	64.06	0	1,739
BEDROCK	5,281	0.036	0.120	0	1.944
WETLANDS	5,281	0.024	0.080	0	0.893
HOUSES	5,281	764.3	484.3	1	10,286

NOTES. Obs. is number of observations. Household observations for q, ISPDOWN, ISPUP, CABLE and FTTH. Census block group observations for all other variables. S.d. is standard deviation.

Table 4: Sample summary statistics

The typical market is in a state that permits wholesale electricity competition. The average maximum downstream bandwidth in the market is 26.11 Mbps and the average maximum upstream bandwidth is 7.35 Mbps. The average all day burst speed across all households in our sample was 15.37 Mbps in the downstream direction and 3.29 Mbps upstream, while the average sustained speed is 11.71 Mbps downstream and 3.09 Mbps upstream. Off-peak period speeds are greater than peak period speeds. The average Internet plan across all households is 11.8 Mbps of downstream speed and 2.83 Mbps of upstream speed. Most households have a cable-modem Internet connection (59 percent), followed by DSL (32 percent) and FTTH (9 percent). These data are reasonably close to FCC (2011a) data on market share which show that on June 30, 2011 cable held 54 percent of the wireline market, DSL held 37 percent, and FTTH held six percent.

Summary statistics of broadband quality by the number of ISPs are provided in *Table 5*. For all four measures of quality, there appears to be a positive relationship between quality and the number of wireline ISPs. A positive relationship is also apparent between qualities, measured by burst and sustain downstream, and the number of wireless ISPs. The econometric analysis in Section 5 will test whether these relationships hold when the controls for observed and unobserved cost and demand conditions are added to the empirical model.

Variable	**Burst downstream**	**Sustained downstream**	**Burst upstream**	**Sustained upstream**
w_1 (wireline competitor 1)	9.384	7.395	1.299	1.249
	(9.980)	(8.328)	(3.806)	(3.755)
w_2 (wireline competitor 2)	14.76	11.00	2.466	2.295
	(10.30)	(8.202)	(5.071)	(5.041)
w_3 (wireline competitor 3)	15.59	11.95	3.418	3.227
	(10.17)	(8.746)	(6.608)	(6.640)
w_4 (wireline competitor 4)	15.89	12.41	4.076	3.866
	(10.51)	(9.007)	(7.398)	(7.550)
w_5 (wireline competitor 5)	16.15	12.10	3.787	3.488
	(10.04)	(8.071)	(6.586)	(6.626)
w_6 (wireline competitor 6)	15.52	12.28	4.166	3.995
	(9.245)	(7.535)	(7.284)	(7.309)
w_7 (wireline competitor 7)	18.65	14.12	5.520	5.077
	(11.01)	(9.046)	(8.192)	(8.377)
Wireless				
m_1 (wireless competitor 1)	9.502	6.891	1.073	0.976
	(7.610)	(4.780)	(0.987)	(0.563)
m_2 (wireless competitor 2)	9.306	7.369	0.985	1.003
	(8.019)	(5.778)	(1.054)	(0.927)
m_3 (wireless competitor 3)	11.45	8.191	1.252	1.186
	(8.598)	(5.765)	(1.851)	(1.785)
m_4 (wireless competitor 4)	15.46	11.96	3.366	3.145
	(10.37)	(8.594)	(6.107)	(6.187)
m_5 (wireless competitor 5)	16.35	12.69	4.329	4.127
	(10.64)	(9.485)	(8.090)	(8.166)
m_6 (wireless competitor 6)	15.06	11.20	2.727	2.535
	(9.896)	(8.154)	(5.521)	(5.511)
m_7 (wireless competitor 7)	15.85	11.81	2.865	2.638
	(10.32)	(7.968)	(5.189)	(5.169)
w_1 (wireline competitor 1)	9.384	7.395	1.299	1.249
	(9.980)	(8.328)	(3.806)	(3.755)

NOTES. Mean speed in Mbps. Standard deviation in parenthesis.

Table 5: Broadband quality by number of ISPs

Table 6 compares market structure and demographics from our sample of 5,281 census block groups with the remaining population of 209,040 census block groups in the US with at least one wireline ISP and one wireless ISP. There are statistically significant differences in the means of most of the variables between the sample and population groups. However, the mean differences for the number of wireline ISPs, number of wireless ISPs, mean maximum bandwidth in the market, age, education and education diversity are relatively small. In contrast, census block groups in our sample have, on average, 389 more persons living with 27.6 less square kilometers and earning about $4,875 more in household income per capita. As such, the results presented in Section 5 should be interpreted with the qualification that they pertain to markets that are located in wealthier, higher-density, urban regions of the US.

	Population		Sample		Diff.
Variable	Obs.	Mean	Obs.	Mean	
W	209,106	2.904	5,281	3.121	-0.217***
		(1.143)		(1.171)	(0.016)
M	209,106	4.710	5,281	5.041	-0.331***
		(1.358)		(1.243)	(0.017)
BANDDOWN	209,106	25.88	5,281	26.11	-0.230
		(35.23)		(30.04)	(0.420)
BANDUP	209,106	7.927	5,281	7.352	0.575*
		(32.22)		(27.32)	(0.382)
MOBILE BANDDOWN	209,106	3.383	5,281	3.470	-0.087***
		(3.287)		(2.037)	(0.029)
MOBILE BANDUP	209,106	1.438	5,281	1.435	-0.003***
		(3.199)		(1.967)	(0.029)
S	209,106	1,416	5,281	1,805	-389.0***
		(782.5)		(1,234)	(17.07)
INCOME	209,106	63,634	5,281	68,509	-4,875***
		(27,031)		(27,994)	(389.7)
AGE	209,106	38.05	5,281	37.90	0.150*
		(5.847)		(5.439)	(0.076)
EDUC	209,106	12.95	5,281	13.27	-0.320***
		(1.516)		(1.486)	(0.021)
SD EDUC	209,106	4.403	5,281	4.610	-0.207***
		(1.683)		(1.774)	(0.025)
AREA	209,106	37.68	5,281	10.04	27.64***
		(332.1)		(39.81)	(0.910)
REG	209,106	0.616	5,281	0.599	0.017***
		(0.384)		(0.384)	(0.005)
HDEN	209,106	942.8	5,281	942.6	0.200
		(2,564)		(2,354)	(32.87)
CPDEN	209,106	2,209	5,281	2,075	143.3*
		(6,748)		(6,379)	(89.01)
ROADS	209,106	27.28	5,281	14.85	12.430***
		(109.7)		(27.64)	(0.450)
INTERSECTIONS	209,106	72.01	5,281	63.20	8.810***
		(120.8)		(64.06)	(0.920)
BEDROCK	208,974	0.046	5,281	0.036	0.010***
		(0.154)		(0.120)	(0.002)
WETLANDS	209,106	0.027	5,281	0.024	0.003***
		(0.090)		(0.080)	(0.001)
HOUSES	209,106	604.6	5,281	764.3	-159.7***
		(324.7)		(484.3)	(6.702)

NOTES. Diff. is the population mean less the sample mean. Obs. is number of observations. Standard deviation of mean in parenthesis. Standard error of diff. in parenthesis. For comparison, we include MOBILE BANDDOWN, the mean maximum advertised downstream speed (Mbps) of all ISPs in the census block group that provide residential wireless broadband Internet connections at June 2011, and MOBILE BANDUP, the mean maximum advertised upstream speed (Mbps) of all ISPs in the census block group that provide residential wireless broadband Internet connections at June 2011.

Table 6: Population, market structure, and demographics

5. Results

The empirical model and data described in Sections 3 and 4 are used to investigate the effects of the number of ISPs and their product-type on broadband Internet quality. Section 5 begins by estimating profit functions for wireline and wireless ISPs, respectively. Then, it estimates alternative specifications of the quality equation, with and without corrections for the endogeneity of market structure, for each of our four measures of speed.

5.1 Step-one: broadband market structure

Following Bresnahan and Reiss (1991), we assume a specification for firm profits that is additively separable in variable profits per capita, fixed costs and an unobserved error, and with variable profits per capita not changing with market size:

$$\Pi^{*}_{Hj} = S_j . V_{Hj}(X_j, W_j, M_j; \alpha_H, \gamma_H, \eta_H) - F_{Hj}(Y_j, W_j; \mu_H, \theta_H) + e_{Hj}$$

$$\Pi^{*}_{Lj} = S_j . V_{Lj}(X_j, W_j, M_j; \alpha_L, \gamma_L, \eta_L) - F_{Lj}(Y_j, M_j; \mu_L, \theta_L) + e_{Lj} \qquad \text{(II-7)}$$

where $V_H = \alpha_{H1} + X\eta_H - \sum_{w=2}^{7} \alpha_{Hw} w_w - \sum_{m=2}^{7} \gamma_{Hm} m_m$ and $V_L = \alpha_{L1} + X\eta_L - \sum_{m=2}^{7} \alpha_{Lm} m_m - \sum_{w=2}^{7} \gamma_{Lw} w_w$ are per-capita variable profits, $F_H = \mu_{H1} + Y\theta_H + \sum_{w=2}^{7} \mu_{Hw} w_w$ and $F_L = \mu_{L1} + Y\theta_L + \sum_{m=2}^{7} \mu_{Lm} m_m$ are fixed costs, X_j = [*INCOME*$_j$, *AGE*$_j$, *EDUC*$_j$, *SD EDUC*$_j$, *AREA*, *REG*$_j$], Y_j = [*ROADS*$_j$, *INTERSECTIONS*$_j$, *BEDROCK*$_j$, *WETLANDS*$_j$, *HOUSES*$_j$] is a vector of observable proxies for fixed costs, w_w is a dummy variable that measures the change in profits following the entry of the *wth* wireline provider and m_m is a dummy variable that measures the change in profits following the entry of the *mth* wireless provider.

Table 7 presents maximum likelihood estimates of the profit functions.[15] Because we do not observe the scale of profits, it is difficult to interpret the estimated parameters other than reporting the relative signs and significance between wireline and wireless ISPs. Comparison of the constant terms, $\alpha_{H1} = 4.940$ and $\alpha_{L1} = 6.559$, indicates that, all other things held constant, wireline markets are less profitable than wireless markets. The estimated parameters on w_w and m_m show the effects on profits from competition. Focusing on wireline profits in column two, we observe relatively large and negative impacts on profits following the entry of the second ($\alpha_{H2} = 3.787$) and third ($\alpha_{H3} = 4.538$) wireline ISPs. The marginal effects from additional entry quickly diminish, and the negative impact on profits from the seventh wireline ISP is $\alpha_{H7} = 4.850$. In contrast, the negative impacts on wireline profits following wireless entry are relatively small and less precisely estimated, ranging from $\gamma_{H2} = 0.346$ to $\gamma_{H7} = 0.834$. A reasonably similar pattern is observed for wireless profits in column five. There are large negative impacts on profits following the entry of the third ($\alpha_{L3} = 4.087$) and fourth ($\alpha_{L4} = 4.765$) wireless ISPs, but the marginal effects become smaller thereafter. The negative impact on profits from the seventh wireless ISP is $\alpha_{L7} = 6.024$. The impacts on wireless profits from wireline entry are small and imprecisely estimated, ranging from $\gamma_{L2} = 0.367$ to $\gamma_{L7} = 0.811$. Overall, the results

---

[15] For brevity, we do not report the effects on fixed costs from entry. To permit easier estimation by maximum likelihood we follow Greenstein and Mazzeo (2006) and scale all the continuous variables in the latent profit functions by dividing by their sample mean and applying the log transformation. Because of missing *BEDROCK* data for Alaska and Hawaii, the final sample was reduced to 5,281.

suggest that wireline and wireless ISPs do not compete aggressively against one another. Once each product-type is provided by three to four incumbents, additional entry has relatively little effect on competitive conduct.

	Wireline Profits			**Wireless Profits**		
	Coefficient	**Estimate**	**s.e.**	**Coefficient**	**Estimate**	**s.e.**
Variable profits						
CONSTANT	α_{H1}	4.940^{***}	1.522	α_{L1}	6.559^{***}	2.287
w_2 (Wireline competitor 2)	α_{H2}	3.787^{***}	1.215	γ_{L2}	0.367^{*}	0.221
w_3 (Wireline competitor 3)	α_{H3}	4.538^{***}	1.226	γ_{L3}	0.336	0.228
w_4 (Wireline competitor 4)	α_{H4}	4.869^{***}	1.228	γ_{L4}	0.417	0.257
w_5 (Wireline competitor 5)	α_{H5}	5.036^{***}	1.232	γ_{L5}	0.526^{**}	0.239
w_6 (Wireline competitor 6)	α_{H6}	5.049^{***}	1.249	γ_{L6}	0.451^{*}	0.246
w_7 (Wireline competitor 7)	α_{H7}	4.850^{***}	1.259	γ_{L7}	0.811^{**}	0.326
m_2 (Wireless competitor 2)	γ_{H2}	0.346	0.310	α_{L2}	1.861	2.101
m_3 (Wireless competitor 3)	γ_{H3}	0.533^{*}	0.302	α_{L3}	4.087^{*}	2.232
m_4 (Wireless competitor 4)	γ_{H4}	0.744^{**}	0.294	α_{L4}	4.765^{**}	2.232
m_5 (Wireless competitor 5)	γ_{H5}	0.628^{**}	0.295	α_{L5}	5.320^{**}	2.235
m_6 (Wireless competitor 6)	γ_{H6}	0.545^{*}	0.295	α_{L6}	5.784^{***}	2.233
m_7 (Wireless competitor 7)	γ_{H7}	0.834^{***}	0.297	α_{L7}	6.024^{***}	2.238
INCOME	η_{H1}	-0.033	0.096	η_{L1}	0.108	0.078
AGE	η_{H2}	0.048	0.174	η_{L2}	-0.363	0.346
EDUC	η_{H3}	0.537	0.366	η_{L3}	-0.645^{**}	0.304
SD EDUC	η_{H4}	0.133^{*}	0.072	η_{L4}	-0.134^{*}	0.076
AREA	η_{H5}	0.012	0.019	η_{L5}	-0.051^{***}	0.020
REG	η_{H6}	0.076	0.086	η_{L6}	0.521^{***}	0.083
Fixed costs						
ROADS	θ_{H1}	-0.178^{***}	0.023	θ_{L1}	-0.153^{***}	0.025
INTERSECTIONS	θ_{H2}	-0.027	0.018	θ_{L2}	0.053^{***}	0.017
BEDROCK	θ_{H3}	-0.014	0.007	θ_{L3}	-0.033^{***}	0.008
WETLANDS	θ_{H4}	-0.016^{***}	0.004	θ_{L4}	-0.046^{***}	0.004
HOUSES	θ_{H5}	0.226^{***}	0.067	θ_{L5}	-0.346^{***}	0.070
Instrument relevance			$\chi^2(5)$ 204.4^{***}			$\chi^2(5)$ 353.4^{***}
log likelihood			-7,324			-7,625

NOTES. Ordered probit models for the wireline profit function and the wireless profit function were estimated by maximum likelihood. s.e. denotes robust standard errors. ***significant at the 0.01 level; **significant at the 0.05 level; *significant at the 0.1 level. Continuous variables are scaled by dividing by their sample mean and applying the log transformation. The estimated coefficients for the effects on fixed costs from entry are not reported. χ^2 is a likelihood ratio test of the null of zero correlation between the fixed costs instruments and firm profits (or, market entry).

Table 7: Step-one estimates of profits

The estimated parameters on fixed cost controls are interesting because they highlight similarities and differences between wireline and wireless entry strategies and because they serve as excluded instruments in the estimation of the quality equation. All other things held constant, both wireline and wireless ISPs are less likely to enter markets with more road miles as this reflects a higher cost for outside plant. Similarly, both wireline and wireless ISPs are less likely to enter markets with more variation in the potential for rocky soils and markets with more wetlands. Both of these physical conditions make deployment and maintenance more difficult and raise costs. Furthermore, construction and maintenance in areas with relatively more wetlands often requires additional local government approvals and more specialized engineering techniques. Wireless providers are more likely to enter markets with more road intersections and with fewer household. Intersections are popular spots because if both roads are major there is apt to be a truck stop, regular gas stations, and a few motels; places wireless providers want to cover. Outside the urban areas, base stations are placed for coverage purposes, not so much for capacity, and wireless providers try to maximize geographic coverage by putting antennas in areas where people live or travel.

Because the fixed cost controls, *ROADS*, *INTERSECTIONS*, *BEDROCK*, *WETLANDS*, and *HOUSES*, enter the profit equation (5) but not the quality equation (4), the estimated effects on quality from wireline and wireless competition can be identified when the fixed cost controls are relevant. Likelihood ratio tests from the ordered probit models indicate relevant instruments. The null that the estimated coefficients on the excluded instruments in the wireline profit equation are jointly equal to zero is rejected at the one percent level ($204.4 > \chi^2(5) = 15.1$), as is the null from the wireless profit equation ($353.4 > \chi^2(5) = 15.1$).

5.2 Step-two: market structure and broadband Internet quality

Table 8 presents the two-step corrected estimates of the Internet quality equation (4) with controls for the endogeneity of market structure. Quality is measured by burst and sustained speeds for the entire day in the downstream and upstream directions. Because the dependent variable is positive, the log transformation is applied to q in all regressions. For the purpose of comparison, *Table 9* provides similar estimates to *Table 8* but excludes the correction terms as control variables. Since the correction terms are calculated from step-one estimates, the asymptotic variance of the step-two estimator is not valid. We report bootstrapped standard errors for the quality estimates with 500 replications.[16]

---

[16] Since multiple household observations on actual speed are recorded from about five percent of census block groups, the effective sample size for estimating equation (4) exceeds 5,281. For robustness, we also weighted markets with multiple household observations with one divided by the total number of sampled households in each market. The results, not reported, are similar to those reported in Table 8.

		Burst downstream	Sustained downstream	Burst upstream	Sustained upstream
w_2 *(Wireline competitor 2)*	β_{W2}	0.088	0.081	0.095**	0.137***
		(0.056)	(0.051)	(0.046)	(0.050)
w_3 *(Wireline competitor 3)*	β_{W3}	0.111	0.122**	0.147**	0.201***
		(0.068)	(0.061)	(0.061)	(0.061)
w_4 *(Wireline competitor 4)*	β_{W4}	0.126	0.149**	0.193***	0.263***
		(0.081)	(0.071)	(0.075)	(0.074)
w_5 *(Wireline competitor 5)*	β_{W5}	0.152	0.177**	0.230***	0.304***
		(0.091)	(0.079)	(0.089)	(0.087)
w_6 *(Wireline competitor 6)*	β_{W6}	0.177	0.243***	0.222**	0.351***
		(0.109)	(0.092)	(0.104)	(0.100)
w_7 *(Wireline competitor 7)*	β_{W7}	0.158	0.185*	0.205*	0.327***
		(0.117)	(0.098)	(0.112)	(0.108)
m_2 *(Wireless competitor 2)*	β_{M2}	-0.127	-0.048	-0.011	0.039
		(0.121)	(0.106)	(0.074)	(0.075)
m_3 *(Wireless competitor 3)*	β_{M3}	-0.101	-0.056	-0.002	0.019
		(0.117)	(0.102)	(0.068)	(0.073)
m_4 *(Wireless competitor 4)*	β_{M4}	-0.112	-0.062	-0.006	0.017
		(0.115)	(0.100)	(0.070)	(0.070)
m_5 *(Wireless competitor 5)*	β_{M5}	-0.140	-0.073	-0.005	0.017
		(0.117)	(0.100)	(0.077)	(0.074)
m_6 *(Wireless competitor 6)*	β_{M6}	-0.141	-0.066	-0.025	0.054
		(0.119)	(0.100)	(0.081)	(0.080)
m_7 *(Wireless competitor 7)*	β_{M7}	-0.208*	-0.128	0.016	0.052
		(0.126)	(0.105)	(0.093)	(0.088)
S	β_1	7.1e-06	8.3e-06**	-9.6e-07	-4.4e-07
		(4.9e-06)	(4.0e-06)	(5.6e-06)	(5.0e-06)
λ_H	ρ_H	-0.024	-0.034*	-0.040*	-0.057***
		(0.022)	(0.017)	(0.022)	(0.020)
λ_L	ρ_L	0.039**	0.029**	0.013	0.009
		(0.017)	(0.013)	(0.016)	(0.016)
CONSTANT	β_0	1.191***	1.015***	-0.606***	-0.548***
		(0.162)	(0.139)	(0.133)	(0.127)
Observations		5,566	5,566	5,561	5,561
R-squared		0.795	0.829	0.814	0.809
Instrument validity					
$Wald_{23}$		0.61	3.29*	3.47*	6.58**
$Wald_{34}$			1.78	3.24*	6.89***
$Wald_{45}$			1.25	1.25	1.87
$Wald_{27}$		0.63	2.15		5.50**
$Wald_{37}$		0.48	1.25		3.71*
$Wald_{47}$			0.58	0.04	1.36

NOTES. Dependent variable is *log q*. Bootstrapped standard errors in parenthesis. ***Significant at the 0.01 level; **significant at the 0.05 level; *significant at the 0.1 level. Coefficient estimates of other control variables, *INCOME, AGE, EDUC, SD_EDUC, AREA, REG, BANDDOWN*, *BANDUP*, *ISPDOWN*, *ISPUP*, *CABLE*, *FTTH*, $BRAND_b$ and $REGION_r$, are not reported. $Wald_{kl}$, k (= 2, 3, …, 7) $\neq$ l (= 2, 3, …, 7), tests the null that $\beta_{Wk} = \beta_{Wl}$.

Table 8: Step-two corrected estimates of quality: all-day

Columns one and two display estimates of burst and sustained speeds for the downstream direction and columns three and four display similar estimates for the upstream. The estimated coefficients on λ_H and λ_L represent correlations between the unobserved factors that affect quality and the latent profit functions of wireline and wireless ISPs. The coefficient on λ_H is negative implying that unobserved factors affect quality and the probability of wireline entry in an opposite manner.[17] The coefficient on λ_H is positive implying that unobserved factors affect quality and wireless entry in the same way. Comparing the corrected estimates of quality to *Table 9* we observe that correcting for the endogeneity of market structure is important. First, when the correction terms are excluded, the estimated positive relationships between the number of wireline ISPs and wireline quality are substantially lower and one may incorrectly conclude that market structure has relatively little effect on broadband quality. Second, the biases are more pronounced for sustained speeds. For example, the uncorrected estimates from column four of *Table 9* show that upstream sustained speeds are about nine percent higher, respectively, in markets with two and three wireline ISPs than in monopoly markets. In contrast, the corrected estimates from column two of *Table 8* shows that sustained speeds are 14 to 20 percent higher, respectively, for the same changes in market structure.[18]

---

[17] The descriptive specification of equation (4) does not permit a clear structural interpretation of the estimated coefficients on λ_H and λ_L. However, the estimate of $\sigma_{ue_H} < 0$ is consistent with our public access story in Section 3. That is, a market with lower unobserved demand due to high quality public access at libraries, museums, post offices and/or work places is likely to attract fewer wireline entrants but also supports higher quality.

[18] As noted by Petrin and Train (2010), the appropriate control function to include in the quality equation can be a specification issue. For robustness, we estimated alternative specifications of the modified quality equation with second- and third-order expressions for the correction terms. The results, not reported, are similar to those reported in Table 8 although the effects from wireline competition are a little smaller for upstream burst speeds.

		Burst downstream	**Sustained downstream**	**Burst upstream**	**Sustained upstream**
w_2 *(Wireline competitor 2)*	β_{W2}	0.062	0.049	0.061	0.090^{**}
		(0.052)	(0.047)	(0.040)	(0.042)
w_3 *(Wireline competitor 3)*	β_{W3}	0.057	0.054	0.072^{*}	0.095^{**}
		(0.052)	(0.047)	(0.041)	(0.043)
w_4 *(Wireline competitor 4)*	β_{W4}	0.054	0.055	0.088^{**}	0.113^{***}
		(0.053)	(0.048)	(0.043)	(0.044)
w_5 *(Wireline competitor 5)*	β_{W5}	0.064	0.061	0.100^{**}	0.119^{**}
		(0.056)	(0.050)	(0.048)	(0.048)
w_6 *(Wireline competitor 6)*	β_{W6}	0.081	0.113^{**}	0.073	0.138^{**}
		(0.063)	(0.055)	(0.059)	(0.058)
w_7 *(Wireline competitor 7)*	β_{W7}	0.052	0.040	0.036	0.085
		(0.067)	(0.059)	(0.061)	(0.061)
m_2 *(Wireless competitor 2)*	β_{M2}	-0.081	-0.011	0.031	0.056
		(0.121)	(0.106)	(0.068)	(0.076)
m_3 *(Wireless competitor 3)*	β_{M3}	-0.059	-0.022	0.019	0.039
		(0.116)	(0.102)	(0.063)	(0.072)
m_4 *(Wireless competitor 4)*	β_{M4}	-0.036	0.002	0.033	0.056
		(0.114)	(0.100)	(0.061)	(0.068)
m_5 *(Wireless competitor 5)*	β_{M5}	-0.032	0.015	0.043	0.061
		(0.114)	(0.099)	(0.061)	(0.068)
m_6 *(Wireless competitor 6)*	β_{M6}	-0.010	0.039	0.082	0.104
		(0.114)	(0.100)	(0.062)	(0.068)
m_7 *(Wireless competitor 7)*	β_{M7}	-0.042	0.003	0.085	0.110
		(0.115)	(0.100)	(0.064)	(0.070)
S	β_1	$8.1\text{e-}06^{*}$	$9.2\text{e-}06^{**}$	-3.7e-07	2.4e-07
		(4.8e-06)	(4.0e-06)	(5.2e-06)	(4.8e-06)
CONSTANT	β_0	1.132^{***}	0.991^{***}	-0.585^{***}	0.494^{***}
		(0.160)	(0.138)	(0.127)	(0.126)
Observations		5,566	5,566	5,561	5,561
R-squared		0.795	0.829	0.814	0.809

NOTES. Dependent variable is *log q*. Robust standard errors in parenthesis. ***Significant at the 0.01 level; **significant at the 0.05 level; *significant at the 0.1 level. Coefficient estimates of other control variables, *INCOME, AGE, EDUC, SD_EDUC, AREA, REG, BANDDOWN*, *BANDUP*, *ISPDOWN*, *ISPUP*, *CABLE*, *FTTH*, $BRAND_b$ and $REGION_r$, are not reported.

Table 9: Uncorrected estimates of quality equation: all-day

The corrected estimates of quality in *Table 8* provide several informative results. There is no relationship between wireline speeds and the number of wireless ISPs in the market. There is no relationship between wireline downstream burst speeds and the number of wireless ISPs in the market (although this relationship becomes significant when examining peak-periods of the day). Wireline speeds are often higher in markets with two or more wireline ISPs than a market with a single wireline ISP, and these effects are much larger in the upstream direction. Interestingly, once the market has about three or four incumbents, the marginal effect on quality from additional competitors is relatively small or zero. The estimated coefficients on the wireline dummy variables, w_w, in column three shows that upstream burst speeds in markets with four wireline ISPs are about 20 percent higher than a monopoly market. The marginal effects thereafter are effectively zero. A Wald test is unable to reject the null that the difference in quality effects between the seventh and fourth competitors is different from zero ($\chi^2(1) = 0.05$; Prob > $\chi^2 = 0.82$). A similar pattern is apparent for sustained speeds but the effects of competition on quality are maximized at three and four incumbents, respectively. Downstream sustained speeds in markets with three or more wireline ISPs are about twelve percent higher than monopoly markets. Upstream sustained speeds in markets with four or more wireline ISPs are about 27 percent higher than monopoly markets.

Since the Internet is often congested during evening hours, time of day may be an important aspect of the quality-competition relationship. *Table 10* and *Table 11* present corrected step-two estimates of the Internet quality equation, where speeds are measured during peak and off-peak periods of the day. Overall, the results are qualitatively similar to those displayed in *Table 8*, although the presence of a significant effect of wireline competition for peak periods, but not off-peak, could be because competition occurs primarily based on peak and not off-peak speed. Another noticeable difference is that the positive effects on downstream burst speeds from wireline competition are smaller and less precisely estimated in the off-peak period relative to the peak period. A potential explanation might be that the ISPs who employ transient speed-enhancing technologies can much easier maintain good burst speed performance during off-peak period than during peak hours.

		Burst downstream	**Sustained downstream**	**Burst upstream**	**Sustained upstream**
w_2 *(Wireline competitor 2)*	β_{W2}	0.122^{**}	0.103^{**}	0.097^{**}	0.150^{***}
		(0.055)	(0.049)	(0.045)	(0.047)
w_3 *(Wireline competitor 3)*	β_{W3}	0.168^{**}	0.159^{***}	0.144^{**}	0.213^{***}
		(0.069)	(0.061)	(0.057)	(0.060)
w_4 *(Wireline competitor 4)*	β_{W4}	0.192^{**}	0.192^{***}	0.190^{***}	0.275^{***}
		(0.082)	(0.072)	(0.070)	(0.072)
w_5 *(Wireline competitor 5)*	β_{W5}	0.186^{**}	0.196^{**}	0.220^{***}	0.320^{***}
		(0.095)	(0.081)	(0.086)	(0.088)
w_6 *(Wireline competitor 6)*	β_{W6}	0.239^{**}	0.273^{***}	0.211^{**}	0.360^{***}
		(0.113)	(0.093)	(0.104)	(0.099)
w_7 *(Wireline competitor 7)*	β_{W7}	0.232^{*}	0.225^{**}	0.202^{*}	0.351^{***}
		(0.122)	(0.102)	(0.111)	(0.109)
m_2 *(Wireless competitor 2)*	β_{M2}	-0.125	0.058	-0.0002	0.005
		(0.149)	(0.121)	(0.073)	(0.081)
m_3 *(Wireless competitor 3)*	β_{M3}	-0.099	-0.053	-0.019	0.024
		(0.138)	(0.119)	(0.070)	(0.081)
m_4 *(Wireless competitor 4)*	β_{M4}	-0.073	-0.034	-0.019	-0.0005
		(0.140)	(0.117)	(0.069)	(0.078)
m_5 *(Wireless competitor 5)*	β_{M5}	-0.070	-0.022	-0.020	-0.003
		(0.144)	(0.120)	(0.075)	(0.081)
m_6 *(Wireless competitor 6)*	β_{M6}	-0.057	-0.006	0.009	0.028
		(0.149)	(0.123)	(0.084)	(0.086)
m_7 *(Wireless competitor 7)*	β_{M7}	-0.117	-0.060	-0.002	0.027
		(0.157)	(0.130)	(0.095)	(0.095)
S	β_1	6.1e-06	6.6e-06	-1.0e-06	-2.3e-07
		(5.6e-06)	(4.6e-06)	(5.4e-06)	(4.9e-06)
λ_H	ρ_H	-0.036	-0.040^{**}	-0.038^{*}	-0.058^{***}
		(0.023)	(0.018)	(0.022)	(0.021)
λ_L	ρ_L	0.024	0.020	0.015	0.013
		(0.018)	(0.016)	(0.015)	(0.015)
CONSTANT	β_0	0.911^{***}	0.798^{***}	-0.590^{***}	0.549^{***}
		(0.190)	(0.154)	(0.137)	(0.132)
Observations		5,551	5,549	5,536	5,533
R-squared		0.766	0.803	0.813	0.805
Instrument validity					
$Wald_{23}$		2.21	5.39^{**}	2.95^{*}	5.29^{**}
$Wald_{34}$			2.46	3.41^{*}	7.05^{***}
$Wald_{45}$			0.03	0.87	2.04
$Wald_{27}$		1.38	2.84^{*}		5.86^{**}
$Wald_{37}$		0.77	1.34		4.42^{**}
$Wald_{47}$			0.47	0.04	1.93

NOTES. Dependent variable is *log q*. Bootstrapped standard errors in parenthesis. ***significant at the 0.01 level; **significant at the 0.05 level; *significant at the 0.1 level. Coefficient estimates of other control variables, *INCOME, AGE, EDUC, SD_EDUC, AREA, REG, BANDDOWN, BANDUP, ISPDOWN, ISPUP, CABLE, FTTH,* $BRAND_b$ and $REGION_r$, are not reported. $Wald_{kl}$, k (= 2, 3, …, 7) $\neq l$ (= 2, 3, …, 7), tests the null that $\beta_{Wk} = \beta_{Wl}$.

Table 10: Step-two corrected estimates of quality equation: peak

		Burst downstream	Sustained downstream	Burst upstream	Sustained upstream
w_2 *(Wireline competitor 2)*	β_{W2}	0.079	0.079	0.087^{*}	0.125^{**}
		(0.059)	(0.051)	(0.048)	(0.050)
w_3 *(Wireline competitor 3)*	β_{W3}	0.096	0.118^{**}	0.138^{**}	0.197^{***}
		(0.072)	(0.060)	(0.061)	(0.061)
w_4 *(Wireline competitor 4)*	β_{W4}	0.111	0.144^{**}	0.187^{**}	0.268^{***}
		(0.085)	(0.069)	(0.075)	(0.075)
w_5 *(Wireline competitor 5)*	β_{W5}	0.150	0.179^{**}	0.222^{**}	0.309^{***}
		(0.094)	(0.075)	(0.088)	(0.088)
w_6 *(Wireline competitor 6)*	β_{W6}	0.182	0.250^{***}	0.216^{**}	0.370^{***}
		(0.113)	(0.089)	(0.109)	(0.107)
w_7 *(Wireline competitor 7)*	β_{W7}	0.149	0.186^{*}	0.189	0.331^{***}
		(0.122)	(0.096)	(0.118)	(0.117)
m_2 *(Wireless competitor 2)*	β_{M2}	-0.122	-0.041	0.031	0.043
		(0.114)	(0.103)	(0.076)	(0.079)
m_3 *(Wireless competitor 3)*	β_{M3}	-0.100	-0.060	0.015	-0.0002
		(0.107)	(0.100)	(0.071)	(0.071)
m_4 *(Wireless competitor 4)*	β_{M4}	-0.126	-0.070	0.012	0.018
		(0.108)	(0.096)	(0.070)	(0.074)
m_5 *(Wireless competitor 5)*	β_{M5}	-0.169	-0.090	0.018	0.019
		(0.112)	(0.098)	(0.076)	(0.079)
m_6 *(Wireless competitor 6)*	β_{M6}	-0.175	-0.089	0.050	0.055
		(0.116)	(0.100)	(0.082)	(0.085)
m_7 *(Wireless competitor 7)*	β_{M7}	-0.248^{**}	-0.157	0.047	0.056
		(0.124)	(0.105)	(0.092)	(0.095)
S	β_1	7.9e-06	$9.6\text{e-}06^{**}$	-1.6e-06	-1.1e-06
		(5.0e-06)	(4.3e-06)	(5.4e-06)	(5.0e-06)
λ_H	ρ_H	-0.024	-0.035^{**}	-0.038^{*}	-0.063^{**u}
		(0.023)	(0.017)	(0.023)	(0.022)
λ_L	ρ_L	0.046^{***}	0.035^{**}	0.011	0.008
		(0.017)	(0.014)	(0.016)	(0.016)
CONSTANT	β_0	1.291^{***}	1.08^{***}	-0.625^{***}	-0.547^{***}
		(0.148)	(0.136)	(0.137)	(0.137)
Observations		5,527	5,525	5,512	5,506
R-squared		0.795	0.827	0.814	0.808
Instrument validity					
$Wald_{23}$		0.30	3.03^{*}	3.42^{*}	6.82^{***}
$Wald_{34}$			1.72	3.59^{*}	8.37^{***}
$Wald_{45}$			2.69	1.26	1.62
$Wald_{27}$		0.57	2.41		5.10^{**}
$Wald_{37}$		0.56	1.59		3.36^{*}
$Wald_{47}$			0.86	0.00	1.06

NOTES. Dependent variable is *log q*. Bootstrapped standard errors in parenthesis. ***significant at the 0.01 level; **significant at the 0.05 level; *significant at the 0.1 level. Coefficient estimates of other control variables, *INCOME, AGE, EDUC, SD_EDUC, AREA, REG, BANDDOWN*, *BANDUP*, *ISPDOWN*, *ISPUP*, *CABLE*, *FTTH*, $BRAND_b$ and $REGION_r$, are not reported. $Wald_{kl}$, k (= 2, 3, …, 7) $\neq l$ (= 2, 3, …, 7), tests the null that $\beta_{Wk} = \beta_{Wl}$.

Table 11: Step-two corrected estimates of quality equation: off-peak

In summary, the results of this research have several economic and policy implications. Examining the estimates of market entry and quality together, there appears to be little competition between wireline and wireless broadband at 2011. This finding suggests that consumers likely view these product-types as more complementary than substitutable with wireline providing bandwidth and wireless providing mobility. These estimates also indicate that much of the interesting competitive conduct in broadband markets occurs in markets with two to four ISPs. As such, the FCC's Form 477 database, which has provided semi-annual counts of wireline and wireless firms by county, zip code and/or census tract since June, 2000 is not that useful since it does not distinguish between one, two or three ISPs due to data confidentiality. The size of the quality improvements, along with the assumptions of the empirical model, also indicate that ISPs have some flexibility in the short to medium run to perform smaller, low-cost modifications to the network to affect the speeds delivered to customers. Finally, the emphasis on upstream quality improvements could reflect relatively more upstream capacity and/or responsiveness to changing consumer preferences for uploading files to the cloud, operating their own web servers, social networking, and remote connectivity to their place of employment.

6. Conclusions

The FCC-sponsored Measuring Broadband America program installed automated performance-measuring devices in over 7,000 households across the U.S. Located between the ISP and the households' network, these devices provide regular quality measurements during different times of the day and year in terms of downstream and upstream data rates, packet loss and latency. We used these data, along with information on the number of ISPs, to examine empirically the relationship between broadband Internet market structure and product quality.

Using four measures of speed, burst and sustained speed in the downstream and upstream direction, we found that wireline speeds are often higher in markets with two or more wireline ISPs than with a single wireline ISP. Excluding the correction terms from the analysis understates this effect. In addition, we also found that increases in wireline speeds are larger in the upstream direction, and there is no relationship between wireline speeds and the number of wireless ISPs.

CHAPTER III.

QUALITY COMPETITION IN THE BROADBAND SERVICE PROVISION INDUSTRY

1. Introduction

This chapter presents an empirical analysis of quality competition among ISPs. The analysis focuses on competition in nearly a thousand local broadband markets in California. The markets are small geographic areas around central offices[19] of Incumbent Local Exchange Companies (ILECs). In these markets, we examine how ILEC ADSL[20] players respond to competition from Competitive Local Exchange Companies (CLECs) and cable players. The analysis focuses on an important quality attribute, the Internet data rate, and estimates the strategic choices of maximum advertised download data rates for ILEC broadband ISPs. Empirical studies to assess the relationship between competition in service provision and the quality of service offered to consumers are few. Furthermore, much of the scant empirical work that examines the broadband Internet Service

---

[19] A central office (or "wire center") is the location where the telephone company's network switching equipment is installed. There is one central office for each local telephone exchange in the traditional Public Switched Telephone Network (PSTN).

[20] Asynchronous Digital Subscriber Line (ADSL) is a technology that enables broadband data transmission over the copper telephone lines in the local loop of the telephone exchange. ADSL service, as defined for purposes of the National Broadband Map and therefore this study, may involve use of copper lines all the way from the central office to the subscriber's premises, or may make use of fiber from the central office to a remote terminal. The common element is that ADSL involves using the copper telephone wires for the last part of the transmission path to the subscriber's premises. ADSL offers faster download than upload speed, as opposed to symmetric DSL (SDSL). SDSL is primarily a business product.

Providers' (ISPs') choice of quality is not based on rigorous microfoundations for the firm's strategic decisions for entry and quality choice.

The work presented in Chapter III takes advantage of recent advances in the industrial organization literature on feasible estimation of discrete games to model and estimate the determinants of an ISP's decision to enter a local market, and what speed of service to offer upon entry. The objective of the research efforts is to characterize the fundamentals of broadband services provision in the US with a structural model and to address various hypotheses concerning how firms respond to the entry and quality decisions of their rivals.

The econometric model used in this research draws on the work of Bajari, et al. (2010), who propose a two-stage method to estimate models of strategic interactions for discrete strategy spaces. In the first stage, we estimate reduced form choice probabilities for the entry and quality decision of each potential entrant in a market. The estimated choice probabilities are then used in the second stage to estimate the structural parameters of the firms' profit functions. The method thus accounts for both the strategic aspect of each firm's decision, as well as the endogeneity problems inherent in the estimation.[21]

The economic literature on competition and quality shows that a higher degree of market competition may lead to higher or lower quality of service. While more competition increases the firm's incentives to supply high quality (holding output prices fixed), more competition also reduces the price–cost margin, which reduces the incentives to invest in quality.

---

[21] The endogeneity problem arises from the familiar simultaneity problem in incomplete information games: in equilibrium, the competitors' actions depend on their expectations about the firm's action, and vice versa.

Considered another way, if greater competition leads to stronger share-stealing effects, there will be higher equilibrium quality. However, given that the elasticity of demand with respect to quality need not increase with competition, the premise may not hold. Thus, the net effect of competition on quality is a priori uncertain and empirical measurement is necessary. Outside of a few markets such as healthcare, airlines, and retail gasoline, this has rarely been done in the literature.

Empirical studies to assess the relationship between competition in service provision and the quality of service offered to consumers are few. Furthermore, much of the scant empirical work that examines the broadband Internet Service Providers' (ISPs') choice of quality is not based on rigorous microfoundations for the firm's strategic decisions for entry and quality choice.

This research work contributes to the existing literature by showing empirically that ILECs appear to respond to intermodal but not intramodal competition. ILECs improve the quality of their ADSL offerings when a cable operator enters the market, or when the incumbent cable operator deploys DOCSIS 3.0. The research also found evidence that ILECs do not raise their ADSL service quality when the competing CLEC is offering ADSL only, regardless of speed, but that ILECs do boost their speed when CLECs deploy fiber in the market. It represents the first step toward a major advance in the empirical analysis of broadband provision, where little structural econometric work has been done.[22]

---

[22] The main structural contribution is by Nevo et. al. (2013), who examine the welfare effects from usage-based pricing and demand for residential broadband.

The structure of Chapter III is as follows. After this introductory section, Section 2 reviews the literature on competition and quality and provides an overview of the literature on entry games with a special focus on static games. Section 3 describes the empirical model, and Section 4 details the data. Results are presented in Section 5. Section 6 concludes with a discussion of the results and outlines the plans for future work.

2. Background and Review

The relationship between competition and quality has attracted much recent attention from United States policy makers. In particular, the education, electricity, finance, health, media and telecom sectors have experienced extensive legislative reform by state and federal governments intended to promote consumer choice, greater product variety, and increased quality through greater competition. In his examination of healthcare, Katz (2013) explains the key justification for this approach as being the

> "… intuition that, due to the potential to steal market share from rivals, a competitive care provider has stronger incentives to raise its quality to attract patients."

Furthermore, this intuition is conditioned on the

> "…belief that greater competition leads to stronger share-stealing effects and, thus, higher equilibrium quality."

This section reviews the theoretical and empirical evidence concerning this intuition.

2.1 Theoretical literature on competition in quality

The straightforward intuition that share stealing will lead to a positive association between competition and quality is supported by economic theory when prices are fixed, as is the case for some regulated markets (e.g., Douglas and Miller, 1974; Schmalensee, 1977). As long as the fixed price is set above marginal cost at some base level of quality, firms will increase quality in an attempt to gain market share.

When prices are not fixed, more competition will also lower the price–cost margin, and this may reduce the incentives to invest in quality. Gaynor (2006) uses the Dorfman-Steiner (1954) condition, adapted to monopolistic competition, to show that the amount spent on quality by the firm depends on the ratio of the quality elasticity of demand to the price elasticity of demand. When an increase in market power reduces both elasticities, quality may increase or decrease depending on the relative strengths of the two effects. Gaynor notes that similar intuition is provided by several other studies but within a different modeling framework. For example, Kranton (2003) studied the effect of competition on quality when consumers are imperfectly informed about quality. Her model shows that if firms compete in price for market share, both price and quality can be lower, which is analogous to the price elasticity of demand exceeding the quality elasticity of demand.

Matsa (2011) describes the tradeoffs facing the firm in the short and long run. He notes that lower profit margins under more competition reduce the immediate cost of losing a “sale” so firms may shade quality. In the long run, however, competition may raise the likelihood that unhappy consumers switch to a competitor, so firms improve quality. In their growth model with incremental innovations, Aghion and Howitt (2009) show that competition fosters innovation in sectors where firms operate at the same technological level. Here, competition reduces pre-innovation rents and thereby increases the incremental profits from innovating and becoming a leader. In other sectors, competition reduces the post innovation rents of laggard firms and thus their incentive to catch up with the leader. Chen and Schwartz (2013) also use a model of innovation to outline conditions where the incentive to add a new, higher-quality product can be greatest under monopoly. The monopolist loses more profit on the old product but may earn more profit on the new one because it prices the old product in a way that internalizes the effect on the new one.[23] While these studies represent only a few examples, they are illustrative of the overall finding from the theoretical literature. Competition can lead to lower or higher quality, depending on the underlying properties of demand, costs and information.[24]

2.2 Empirical literature on competition in quality

Given the theoretical ambiguity in outcomes, it is not surprising that recent empirical studies have produced mixed results on the relationship between competition and quality for different industries with different market conditions.

---

[23] The key factor is the extent to which the monopolist can divert sales to the new product as opposed to leaking sales to outside goods if it raises the price of its old product (Chen and Schwartz, 2013).

[24] This ambiguity has a long history in industrial organization (IO) theory. Chamberlin (1933) and Abbott (1955) show that firms with market power may reduce product quality to save costs and maximize their profits. Swan (1970, 1971) demonstrated no relationship between monopoly power and product quality and defined conditions under which a competitive and monopoly market introduce a product with the same level of quality but the monopoly will charge a higher price. Schmalensee (1979) shows that this result holds up under some relaxation of the original assumptions but questions whether quality choice under oligopoly will be well approximated by either the competitive or monopoly models.

The basic empirical approach has been to write down a firm's equilibrium quality function as the implicit solution to their profit maximization problem. A reduced-form quality equation is then specified that relates some measure of quality to cost and demand shifters and to a measure of the number of firms in the market. For example, as mentioned in Chapter II already, Mazzeo (2003) shows that average flight delays are longer in more concentrated airline markets. Goolsbee and Petrin (2004) estimate that cable television (TV) channel capacity, number of over-the-air channels and number of premium movie channels increased in response to satellite entry, while Savage and Wirth (2005) document a similar effect with respect to potential entry from cable overbuilders.[25] Matsa (2011) finds that supermarkets facing more intense competition have more products available on their shelves, while Olivares and Cachon (2009) show that the inventories of General Motors dealerships increases with the number of competitors.

In contrast, Domberger and Sherr (1989) find no correlation between the threat of new entry and customers' satisfaction with their attorney used for home purchases. Prince and Simon (2013) show that flight delays for incumbent airlines worsen in response to entry threats by Southwest Airlines. Chen and Gayle (2013) examine mergers and product quality (i.e., the ratio of non-stop flight distance to total flight distance used to get passengers from origin to destination) for the airline industry. They find that quality increased in markets where the merging airlines did not compete *ex-ante*, and decreased in markets where they did. This is

---

[25] An overbuilder in the cable industry is a second entrant into an existing cable franchise area to compete with the incumbent. Overbuilding using hybrid fiber-coax networks (i.e., a traditional cable system architecture) is relatively rare in the US in general and in California in particular.

consistent with their theory that mergers improve coordination, but diminish competitive pressure for firms to provide high-quality products.

Similar studies have also been conducted with advertising as the proxy for quality. Dick (2001) examined the United States retail banking industry using higher advertising intensity (i.e., marketing expenses divided by total asset value) as a measure of higher customer service quality. He found that dominant banks provide a higher level of service quality than fringe banks. Crawford (2007) analyzed the relationship between TV station ownership and the quality of their programming. He found no relationship between cross ownership with a local newspaper or radio station and the number of minutes of advertising included in TV programming, where more minutes are indicating lower quality TV service. Hiller et. al. (2014) analyzed consumer media bundles and showed a positive correlation between the number of independent TV stations and the amount of time and space devoted to advertising.

In telecommunications, as also mentioned in Chapter II, Wallsten and Mallahan (2013) found that the number of wireline Internet service providers (ISPs) in United States census tracts is positively correlated with the highest advertised downstream speeds. Nardotto et. al. (2012) showed a positive relationship between lower barriers to entry, measured by the presence of local loop unbundling, and average broadband download speeds in the United Kingdom.[26] In addition, Chapter II showed that wireline speeds are higher in census block groups with two

---

[26] Unbundling requires the incumbent telephone company to lease the connection from their central office to the household (“local loop”) to new entrants so they can compete in the final product market for broadband Internet.

or more wireline ISPs than with a single wireline ISP, and there is no relationship between wireline speeds and the number of wireless ISPs.

While reduced-form quality equations provide useful insights into the general relationship between competition and quality, they say nothing about the strategic interactions between firms with respect to their quality choices. Kugler and Weiss (2013) use a reaction-function approach to estimate the strategic quality choices for Austrian gas stations. Their empirical reaction function relates the opening hours of a station to those of its competitors. Their results suggest significant but imperfect coordination, in opening hours among stations of the same network, which implies that opening hours are strategic complements. They find a similar but weaker effect between independent stations or between stations from competing networks. Brueckner and Luo (2013) use a similar model to investigate strategic interaction among United States airlines in flight frequency. Using instrumental variables estimation, a positive reaction function is found in some specifications, suggesting complementarity in the choice of frequencies.

In summary, there is much work in industrial organization (IO) on the effect of competition on prices, but not nearly enough has been done on quality. Both parties to a merger lawsuit would benefit from some empirical evidence in this direction. Lack of evidence on this is mostly due to the lack of data on quality. Schmalensee (1979) noted 35 years ago that

> "…it is far from obvious that any single mathematical representation of quality can serve for a broad spectrum of products."

Even more so today, most industries sell highly differentiated products, making standardized quality measures difficult to collect, and the few previous studies looked at flight delays, product availability in supermarkets, and the number of TV channels. These are worthwhile quality dimensions in their respective industries, but pale compared to the importance of Internet speed in the broadband industry. Chapter III investigates an essential and standardized quality attribute – Internet speed in a highly relevant industry in the digital age – and estimates the strategic choices of download and upload speeds for Californian broadband providers. Estimation of the static model of strategic interactions with discrete-choice methods determines the probability of a particular level of quality for a representative broadband provider as a function of the expected quality choice of rivals and various market characteristics.[27]

2.3 Broadband Market Entry

When entering new markets, or re-evaluating their business plans regarding technology or quality in an existing market, ISPs face many decisions: which technologies to offer, what packages to create, how much to invest in service quality, what prices to charge, and how to promote service offerings. The ISP's customers can decide on the type of contract, what service level they purchase, and what additional products they take with a service bundle. ISPs must also consider the strategic reactions of the rival firms. Will they enter the market? How will a competing firm position its market play? The interrelated nature of these decisions

---

[27] Xiao and Orazam (2011) estimate a simple discrete-choice model of broadband entry and find that sunk costs are an important determinant of wireline entry in US zip codes. However, they are unable to distinguish between one, two or three providers due to data confidentiality, and they do not estimate the direct effects of entry on market outcomes such as quality.

suggests modeling them with empirical discrete games that can assume sequential or simultaneous move by the players.

This essay models the broadband Internet markets as a repeated simultaneous game. Simultaneous games are imperfect-information games; players do not have the knowledge about the actions of the others. In the model, ISPs choose a quality simultaneously, and they do not know the current-period actions of the other firms.[28] When we observe broadband markets and make an attempt to understand how the players behave, we also lack information on price, cost, or demand data. We can observe, however, the entry and exit of players, the speeds that they provide, along with market demographics, and make inferences even in the case of incomplete information.

Inference about structural parameters of the profit function from observations on entry was made possible by Bresnahan and Reiss (1990) and the subsequent stream of literature triggered by their seminal paper. Bresnahan and Reiss inferred the effects of entry on competition from the relationship between the number of market entrants and the market size. By observing strategic entry decisions of small retail firms in isolated rural markets, they argued that firms must pay a fixed and sunk cost to enter the market. They also argued that the total industry profit depends on the number of firms on the market but not on the identity of the entrants. Bresnahan and Reiss proposed an estimator that maximizes the likelihood for the number of firms and introduced the idea of entry

---

[28] The reality is, or course, more complex; ISPs do not actually choose a service quality once per six months all at the same time. However, our modeling approach is commonly adopted in the literature when there is no clear first mover and is best seen as an approximate structure designed to reflect uncertainty regarding competitors' plans.

thresholds, i.e., the market size required to support a given number of firms. The two main disadvantages of their model are that firms' costs are homogenous and that the firms do not offer differentiated products. In their later work (1994), they estimated firms' sunk costs from differences in the thresholds for entry and exit. Berry (1992) relaxes this limitation by allowing heterogeneity between firms entering the markets. He develops a model of market entry considering a large number of heterogeneous potential entrants and applies the model to analyze competition in airline markets. Berry recommends using simulation methods to address the computational problem of calculating the linear combination of integrals that define the probability of events. Mazzeo (2002) extends the Bresnahan-Reiss model by allowing firms to offer heterogeneous (high-quality and low-quality) products. Using data from motel markets along U.S. interstate highways, and endogenizing the quality choice of firms, he founds that hoteliers have strong incentives to differentiate. Ciliberto and Tamer (2009) broaden the literature by allowing for heterogeneity without making equilibrium selection assumptions. Applying a pseudo maximum likelihood estimation method to the US airline industry, and expanding on Tamer's earlier work (2003), they find evidence of heterogeneity across airlines in their profit functions.

Additional recent contributions include Seim (2006), Aguirregabiria and Mira (2007), and Bajari et al. (2010). Seim's static equilibrium model makes early use of spatial econometrics in market structure and product type choice studies. Her simulation results demonstrate the firms' incentives for spatial differentiation in the video rental industry and the importance of incorporating product-type choices into the market entry process. Aguirregabiria and Mira (2007) propose a two-step method to estimate static games of incomplete information and illustrate it using an example of a static game of market entry. Their method greatly reduces the computational complexity of earlier approaches, and the present work derives from theirs. In the spirit of Aguirregabiria and Mira (2007), Bajari, et al. (2010) implement a two-stage method to estimate models of strategic interactions for discrete strategy spaces. In the first stage, they estimate reduced form choice probabilities for the entry and quality decision of each potential entrant in the market. Then, in the second stage, they use these computed choice probabilities to estimate the structural parameters of the firm's profit function. As an application for the two-stage model, they study the determination of stock recommendation issued by equity analysis for high-tech stocks between the years of 1998-2003.

In telecommunications, employing the entry model in Mazzeo's (2002) study, Greenstein and Mazzeo find (2006) evidence that competitors are heterogeneous and that firms account for both potential market demand and the business strategies of their competitors when making their entry decisions. Following Bresnahan and Reiss (1991), Xiao and Orazem (2011) estimate a discrete-choice model of broadband entry, as discussed above.

In addition to these two works, in which estimation is based directly on theoretical entry models, most studies of market entry in broadband are nonstructural (reduced-form). Almost all of the existing works have been extracted from cross-sectional, static studies of existing players; the impacts of potential new entrants are not studied. In a typical broadband market entry study, a cross-section of either the number of ISPs or an indicator for the presence of at least one competitor in the local area is regressed on demographic and other market characteristics (Prieger, 2003; Grubesic and Murray, 2004; Flamm, 2005; Prieger and Church, 2012; Prieger, 2013). These studies show that the decisions of telecom service providers to deploy network resources and offer service in a local market depend on both economic and regulatory considerations. Demand factors such as market size, average income, and other demographic characteristics all been shown to affect broadband penetration (Prieger, 2003; Grubesic and Murray, 2004; Flamm, 2005; Flamm and Chaudhuri, 2007; Prieger and Hu, 2008; Prieger, 2013). Some of these papers also show that population density or terrain also influence broadband penetration in the expected ways, and can be used as proxies for cost. For example, rural areas are more likely to be served only with lower-speed broadband or by few providers, or less likely to have broadband available at all than urban areas, due to low population density and rougher topography, (Stenberg et al., 2009; Li et al., 2011; Prieger, 2013). Intermodel and intramodal competition among broadband ISPs, both actual and potential, also affects the incentives to enter the local markets (Denni and Gruber, 2007; Prieger and Hu, 2008; Wallsten and Mallahan, 2013). Like the work of Greenstein and Mazzeo (2006) and Xiao and Orazem (2011), this essay performs structural estimation to identify parameters of the potential entrant's profit function. Unlike the earlier studies, this research is particularly interested in those parameters relating to quality competition, and will thus be a significant addition to the scant structural empirical work on broadband competition.

3. Econometric Model

3.1 Game-theoretic underpinnings

Our structural econometric model is based on a static game theoretic approach to the profit maximization decision of a broadband provider. We adopt the approach of Bajari et al (2010) for estimation of static games of incomplete information with multiple equilibria, and we refer the interested reader to their article for presentation of the model at a high level of mathematical formality. The static game approach is a generalization of a discrete choice model that allows the quality choice of a firm to depend on the actions of the other firms. Firm *i* in market *m* at time *t* chooses an alternative $a_{imt} \in A = \{0,\dots J\}$ representing a quality level, where alternative 0 represents offering no broadband at all. Let the firm's profit u_{imt} be

$$u_{imt}(a_{imt}, a_{-imt}, s_{imt}; \theta) = \pi_{imt}(a_{imt}, a_{-imt}, s_{imt}; \theta) + \varepsilon_{imt}(a_{imt}) + \eta_{imt} \quad \text{(III-1)}$$

where a_{-imt} represents the actions (chosen alternatives) of the other potential entrants in the market and period, s_i is a vector of firm *i*'s state variables affecting profit, and θ is a finite vector of parameters. The state vector is assumed to be common knowledge to all firms, but $\varepsilon_{imt}(a_{imt})$ is private information for firm *i*. For identification, we assume that after accounting for the actions of the other firms through argument a_{-imt}, the state variables of the other firms (s_{-imt}) do not affect directly firm *i*'s profits. This exclusion restriction will be used to identify the parameters of the deterministic part of the profit function in the two-step estimation. The final term, η, can be either private or common information, and includes all factors specific to the market, period, firm, or any combination of these that affect the profit of all alternatives equally. For example, η_{imt} can be a market-firm fixed effect η_{im} such as a firm's long-standing reputation in the area or a period fixed effect η_t stemming from the business cycle. The state variables include some factors, x_{imtj}, that vary across alternatives $j \in A$ and others, z_{imt}, that do not. The actions of the other firms enter observed profit through a set of competition variables $w_{imt}(a_{-imt})$. Observable profits are assumed to be linear in the state and competition variables:

$$\pi_{imt}(a_{imt}, a_{-imt}, s_{imt}; \theta) = y_k^{'} z_{imt} + \beta^{'} x_{imtj} + \delta^{'} w_{imt}(a_{-imt}) \qquad \text{(III-2)}$$

The firm does not observe the actions of other firms before choosing its action. Suppressing time and market subscripts, the firm's *expected* profit is

$$U_i(a_i, s_i; \theta) = E\pi_i(a_i, a_{-i}, s_i; \theta) + \varepsilon_i(a_i) = y_k^{'} z_{imt} + \beta^{'} x_{imtj} + \delta^{'} E w_{imt}(a_{-imt}) \qquad \text{(III-3)}$$

where the expectation is taken over the space of other firms' private information. See Bajari et al (2010) for a precise statement of this expectation. Informally, if it could observe the private information, firm *i* would know what each other firm would do. For example, by assuming the other firms want to maximize profit, firm *i* can calculate the other firms' decision rules for quality choice that map their private information into their action space *A*. Taking expectation over the private information of the other firms yields an expected set of resulting competition variables, Ew_{imt}.

3.2 Estimation

We assume that $\epsilon_i(a_i)$ is drawn from the extreme value distribution as in the logit model. Then the firm chooses alternative $a_i = j$ such that $U_i(j,s_i;\theta) = U_i(k,s_i;\theta)$ for all $k \neq j$. Given the logit structure of the error terms, the probability that the firm chooses quality level j is thus

$$\Pr(a_{imt} = j|s_{imt}) = \frac{\exp\left(\gamma_k' z_{imt} + \beta' x_{imtj} + \delta_j{}'(Ew_{imt})\right)}{1 + \sum_k \exp\left(\gamma_k' z_{imt} + \beta' x_{imtj} + \delta_j{}'(Ew_{imt})\right)}$$

This formulation incorporates the usual identification assumption that the profit of base alternative 0, not entering, is normalized to zero. The expression above allows estimation of $\theta = (\beta,\gamma,\delta)$ by maximum likelihood estimation (MLE) using a conditional logit model for choice of quality.[29] Note that the choice-invariant fixed effects η drop out of the conditional likelihood. For the same reason, the coefficients on z_{imt} must be alternative specific for the impact of the z to be estimable (as is familiar from the multinomial logit model).

The econometrician observes actual w_{imt} in the data, but not Ew_{imt}. We cannot substitute w_{imt} for Ew_{imt} in estimation, because the former is endogenous due to the simultaneity of the game. In the two-step method of Bajari et al (2010), reduced form choice probabilities for competitors are estimated in the first step.

---

[29] Estimation was performed using the `asclogit` command in Stata 13.1.

By observing quality choices in a large number of markets, the econometrician forms a consistent estimate of the equilibrium choice probabilities. In our application, we use conditional logit in the first step, where the quality choices of each firm are regressed on z_{imt} and x_{imtj}, but not the competition variables. We then use the estimated choice probabilities to form the expected values of competition variables Ew_{imt} for the second-step estimation.

Since the state variables x and z are used to identify both the effects of (β,γ) and δ on the observed choices and profit is linear, an exclusion restriction avoids collinearity problems and helps identification. As mentioned above, the state variables specific to competitors are not included in the second-step estimation for the ILEC's decision. However, those excluded state variables are used to estimate the choice probabilities of competitors $k \neq i$ in the first step, which then appear in the estimate of the expected action and consequence $Ew_{imt}(a_{-imt})$. In our application, the excluded instruments are infrastructure variables of the other firms, which we describe in the next subsection. Although these affect the actions of the other firms, conditional on those actions the infrastructure costs of the other firms should not affect directly the quality choice of firm i.

3.3 Specifics

This essay focuses on ADSL provision by ILECs, the dominant telecommunications firms in the local markets. In California, this includes the "U-verse" DSL service offered by AT&T, but not the "FIOS" fiber-to-the-home service from Verizon, which instead counts as fiber-based broadband. The competitors include ADSL and fiber broadband from CLECs and cable modem broadband offerings.[30] For now, we set aside competition from less closely related offerings such as fixed or mobile wireless broadband, but recognize that these competitors may be more relevant in the future.

For estimation, the many speed categories in the NBM are collapsed into four alternatives for ADSL: greater or equal to 768kbps (kilobit per second) but less than 3 Mbps (Megabit per second), 3 Mbps to 6 Mbps, 6 Mbps to 10 Mbps, and 10 to 25 Mbps. No ILECs report offering ADSL with maximum speed below 768 kbps or greater than or equal to 25 Mbps during our time period. Choice $j = 0$ of not offering ADSL in the market gives the base alternative. These speed categories are presented in *Table 12* for reference. Variables in $\boldsymbol{z}$ include demographic variables reflecting market characteristics and infrastructure variables. Variable *NearestAnySpeed* is the distance in miles (or log miles, in one of the estimations) to the Census block nearest to the center of market m where firm i offered broadband using the same technology (ADSL, for the ILECs here) last period, per the NBM.[31] Thus, when the ILEC already offered ADSL in the market at t-1, *NearestAnySpeed* will be small. When the ILEC did not offer ADSL in the market the previous period but did in a nearby area, *NearestAnySpeed* will be smaller than if the firm offered ADSL in some distant location. Due to the presence of sunk costs in broadband

---

[30] The distinction between ILECs and CLECs is clear within a market, because the NECA tariff identifies the locations of ILECs, and all other ADSL or fiber providers in that market must be CLECs. However, AT&T (and other large firms) may be treated as ILECs in some markets and CLECs in other markets if they do out-of-region entry.

[31] All distance variables were calculated based on the latitude and longitude of the central offices and Census block centroids. The great circle distance metric was computed and the nearest broadband locations for each firm and market were found using a FORTRAN program.

infrastructure deployment, we thus expect that higher values of *NearestAnySpeed* will lower the probability of higher ADSL quality.

There are two *x* variables in the model: *NearestSameSpeed* and *SameSpeednotFound*. The former is constructed similarly to *NearestAnySpeed* but only ADSL in the same quality category counts in the calculations. Since this variable was missing when the firm did not offer a particular quality level anywhere in California the previous period, it is set to zero for such cases, and an indicator variable *SameSpeednotFound* is set to one. *SameSpeednotFound* thus captures the impact of the variable *NearestSameSpeed* when the latter would logically be infinite. By logic similar to the above, we expect *NearestSameSpeed* and *SameSpeednotFound* for quality *j* both to impact negatively the probability of the firm offering quality *j*.[32]

The competition variables we choose are indicators for the presence of at least one competitor in a quality category. These are the *w* variables, which are functions of quality choice decision a_{-i} as introduced above. The indicators are cumulative, defined as $w_{imt}^{bj}=1$ if broadband of type *b* and speed *j* or higher is offered in the market this period, with $w_{imt}^{bj}=0$ otherwise. In the second step estimations, Ew_{imt} are the expected values of these variables, where the w_{imt}^{bj} are arranged into a column vector including all *b* and *j*, and thus can take values between 0 and 1. For CLECs offering ADSL, there is an additional alternative $j = 5$ of greater or equal to 25 Mbps but less than 50 Mbps. For cable modem, the categories are $j = 1$ (less

---

[32] Each *x* variable also has cross-impacts. For example, *NearestSameSpeed* for the highest quality level has a marginal effect on the probability of the firm offering ADSL in the lower quality categories. We calculate but do not report these cross-effects in the tables for the sake of brevity.

than 10 Mbps), 2 (10 Mbps to 25 Mbps), 3 (25 Mbps to 50 Mbps) 4 (50 Mbps to 100 Mbps), and 5 (100+ Mbps). For fiber, the quality categories are 0 (no entry or any fiber below 1 Gbps, grouped because there is little fiber below that speed) and 1 (gigabit fiber). The ADSL and cable modem speed categories are also presented in *Table 12* for reference.

Alternative	**Speed Category**		
	ILEC ADSL	**CLEC ADSL**	**Cable Modem**
0	No entry	No entry	No entry
1	768K ≤ MADTR <3M	768K ≤ MADTR <3M	MADTR <10 M
2	3M ≤ MADTR <6M	3M ≤ MADTR <6M	10M ≤ MADTR < 25M
3	6M ≤ MADTR <10M	6M ≤ MADTR <10M	25M ≤ MADTR < 50M
4	10M ≤ MADTR <25M	10M ≤ MADTR <25M	50M ≤ MADTR < 100 M
5	NA	25M ≤ MADTR <50M	100M ≤ MADTR

Table notes: *MADTR* is the maximum advertised downstream transmission rate.

Table 12: Quality alternatives – downstream speed categories

Since the demographic variables are specific to the market and apply to all firms and periods, and because it is unrealistic to assume that two observations from different periods for the same firm and market are independent, we use standard errors that are robust to clustering within markets. Finally, even though we have panel data, in this essay we do not exploit the panel structure of the data in estimation. All estimations use pooled data from the latest four periods. Data from the earliest (fifth) period is used only to calculate the distance regressors for the first period included in the regression. We pool the data for several reasons. First, note that any market or market-firm fixed effect (η_m or η_{im}) that affects identically the profits of all quality levels is already accounted for in the conditional logit formulation. More practically, adding alternative-specific market fixed effects would add about four thousand coefficients to the model, making estimation difficult and possibly leading to incidental parameter bias. Finally, most of the variation in the data occurs in the cross section, not the time series within each market, and so fixed effect modeling would reduce greatly the effective size of the sample.

4. Data

The research presented in Chapter III focuses on competition in broadband service provision in local markets in California. California is large enough to contain many local markets, yet not so large as to make working with the voluminous NBM data unwieldy.

4.1 Market definition

Any definition of the broadband Internet market is only an approximation of how ISPs may view their market play. Market definition is made difficult because the natural areas of deployment for different types of broadband providers, e.g., wire center serving areas and cable franchise areas, do not exactly match. Previous studies have used counties, census tracts, ZIP codes, and local telephone exchange boundaries to define the geographical market for broadband Internet (Gillett and Lehr, 1999; Prieger, 2003; Wallsten and Mallahan, 2013; Xiao and Orazem, 2011; Nardotto et. al., 2012; Prieger and Conolly, 2013; Prieger, 2013). Our market definition instead is similar in spirit to that of Prieger and Hu (2008b), who carefully examine the distance from the local phone company's central office to define the local markets for ADSL. Roughly speaking, the 965 broadband Internet markets analyzed in this study are small geographic areas within the distance of 2.3 miles of the ILECs' central offices in California.[33] The main reason behind our market definition is that in areas close to an existing ILEC wire center, the incumbent ADSL player has the greatest ability to match the speeds of fiber and cable modem competitors, due to the degradation of DSL speed with line distance. The rest of this subsection gives further details on the market definition process and can be skipped by readers not interested in the technical details.

The definition of the markets for the entry game is a three-step process. Market definition begins with drawing a circle of radius 12,000 feet (12 kilofeet (kf), about 2.3 miles) around each California ILEC wire center found in the NECA tariff #4. The threshold of 12 kf (along with a secondary threshold of 18 kf, discussed below) was chosen in accord with California Public Utility Commission methodology for validating information on provision of DSL (CPUC, 2013). A radius of 12 kf from the equipment in the wire center also corresponds to the straight-line threshold for provision of DSL at 6.3 Mpbs.[34] Since last-mile network may be constrained to run along right-angled streets, a 12 kf radius by the Euclidean metric has a worst-case situation where the lines from the wire center have a taxicab distance of 18 kf

---

[33] For robustness, the research also analyzed markets with a circle of radius of 18 thousand feet from an existing central office.

[34] For the relationship between distance and ADSL speed, see http://whatis.techtarget.com/reference/Fast-Guide-to-DSL-Digital-Subscriber-Line.

long,[35] in which case DSL of at least 1.5 Mbps is possible. These speed limitations are relaxed in many markets by the installation of remote terminals to neighborhoods farther from the wire center, as in AT&T's fiber-to-the-node (FTTN) architecture for its U-verse service. However, we have no data on which markets include remote terminals.

In a second step, we limit the areas defined by the circles to actual wire center serving areas of the ILECs. Using GIS data from GDT on the service territory of the ILECs associated with each wire center, parts of the 12 kf radius circles not also in the actual service territory were excluded from each market. This step ensures that in dense urban areas, where wire centers are closer to each other, the market area associated with one wire center does not overlap with the territory served by an adjacent wire center.[36]

A third step is necessary to match the second-step market areas to the broadband provision data in the NBM, which is keyed to Census geography. The third-step market area, therefore, consists of the union of all Census block groups (CBGs) that lie wholly within the area from the second step. In a few rural locations, no CBG are contained within the areas from step two, and in such cases we instead use the set of CBGs that overlap the step-two area. All GIS processing for market definition was performed using ArcMap. The market definitions result in

---

[35] The taxicab distance, defined as the distance between two points measured along axes at right angles, is also called Minkowski's L_1 distance. The worst-case scenario is the maximum taxicab distance for a fixed Euclidean distance, and occurs when the communications lines run along the legs of a right triangle with hypotenuse equal to 12 kf in length. The line length to reach the 12 kf radius is about 16.97 kf in this case.

[36] In a few markets (29) the purported coordinates of the wire center from the GDT and NECA data sources did not agree to within 2 V&H units (about 0.6 miles) using the L2 norm. (The telephone industry uses a unique "V&H" coordinate system for central office locations.) For such wire centers we did not use the GDT wire center-serving area to limit the market. Instead we used the entire area defined by the 12 kf radius around the wire center (as located using the NECA data) less any area already part of another market.

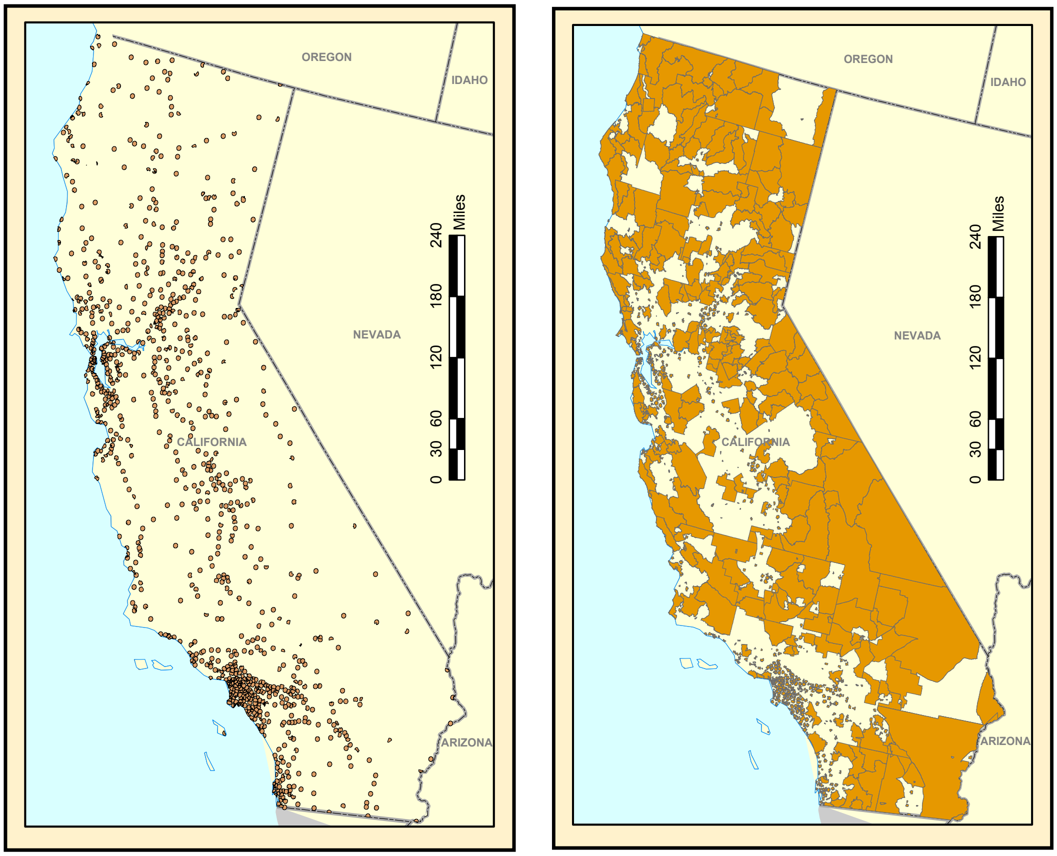
965 markets.[37] For use in testing the robustness of the econometric conclusions, an equivalently constructed set of markets based on an 18 kf (about 3.4 miles) radius was also created.

In summary, our definition results in a set of local broadband markets that are distinct, small enough to represent an ILEC's infrastructure decisions in a single wire center area for DSL, yet large enough so that local decisions about infrastructure and quality of service do not affect multiple markets.[38] *Figure 3* shows the step-two and step-three market areas throughout the state. *Figure 4* and *Figure 5* show a detailed view of some of these markets. *Figure 4* shows some of the Los Angeles urban area, in which the markets are often constrained more by wire center boundaries than by the 12kf radius. This is most apparent in the West L.A. markets in the upper left and the downtown L.A. markets in the upper right areas of the figure. The heavy dots on the map mark the ILEC central office locations, the blocky areas surrounding the points are the market areas from step three (each a collection of CBGs), and the larger, circular or smooth-bordered areas are the market areas from step two (the intersection of the 12kf and wire center area boundaries). *Figure 5* shows some extremely rural markets.

---

[37] Four of the potential market areas were dropped after the second step because they were very small, and an additional market defined for a wire center on the Oregon border was dropped because it appeared to serve customers in Oregon instead of California.

[38] By which we mean infrastructure deployment in the central office and within the same wire serving area. Of course, backhaul infrastructure such as high capacity transmission lines between central offices or connections to the Internet backbone may affect multiple markets.



Step-Two Market Areas

Step-Three Market Areas

Figure 3: The Market Areas in California

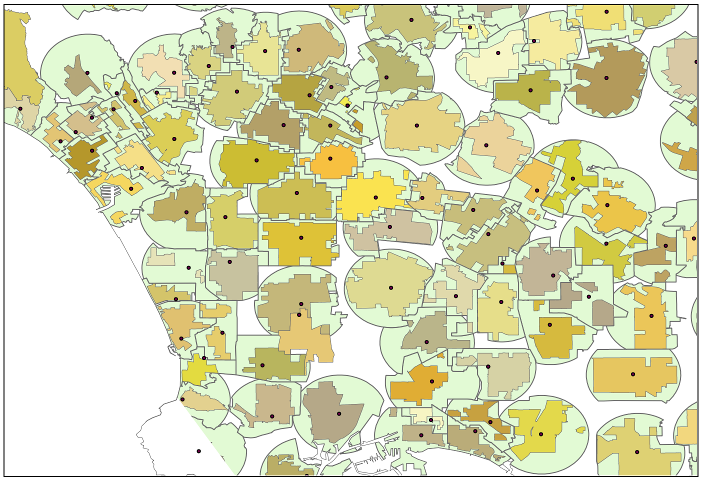


Figure 4: Example of Urban Market Areas: Los Angeles Area

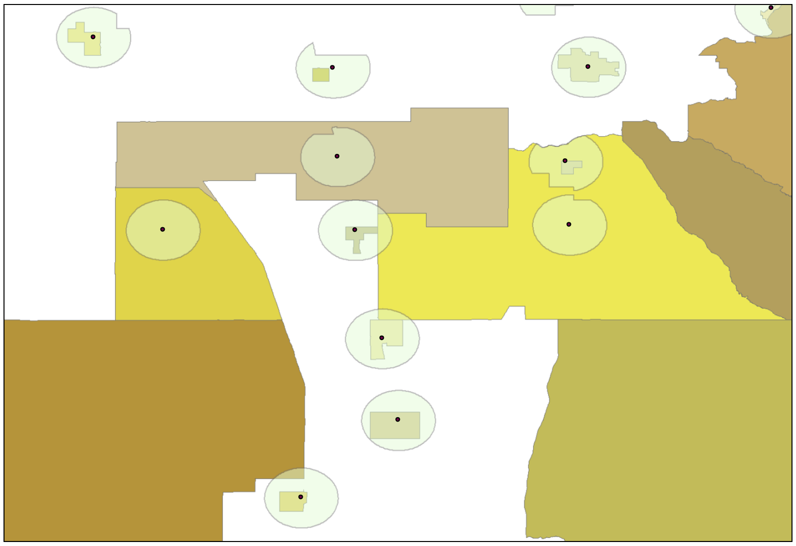


Figure 5: Example of Rural Market Areas: Rural Fresno County

A few markets at the top of *Figure 5* are like those in *Figure 4*, where at least one CBG falls entirely within the step-two market area. The large markets in the middle of *Figure 5* show examples of the few markets composed of CBGs that overlap with the step-two market area (because no CBG is wholly contained within it). These CBGs with low population density can be quite large, and this accounts for much of the rural areas in the right panel of *Figure 3* being colored in. This is not likely to lead to overstating broadband presence, however. Whether an extremely rural central office is placed into a large or small market area, the maximum speed of any broadband provision of any type is highly likely to be present near the central office, which is typically in the center of town.

4.2 Broadband data

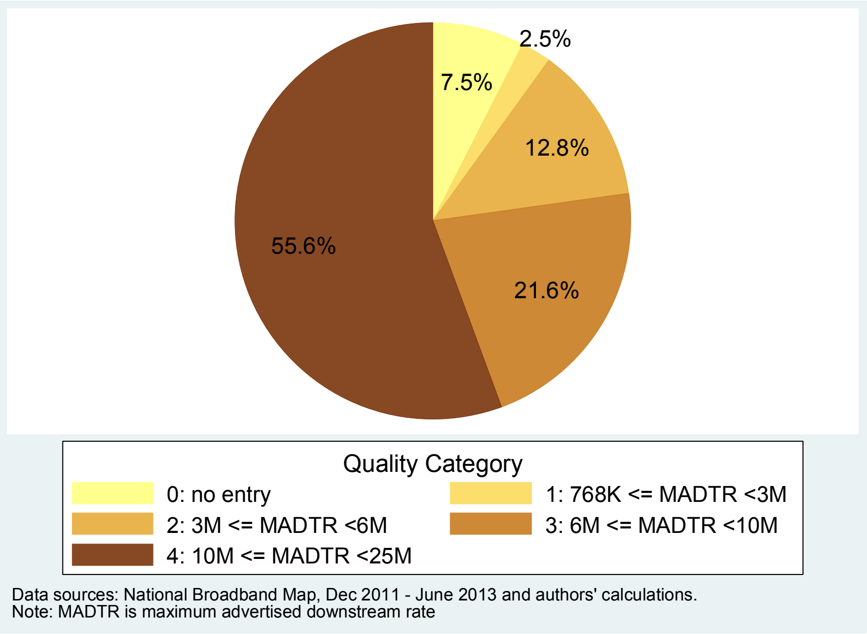
We draw data on the location and quality of broadband service in California from five semiannual waves of the US National Broadband Map, June 2011 to June 2013.[39] These were the latest data available at the start of this project. We chose not to include the first two rounds of the NBM data, from 2010, because those rounds used an earlier Census geography. We matched ISPs offering service anywhere in the market areas to the corresponding markets and recorded each firm's maximum advertised downstream rate, separately by technology and holding company.[40] While in theory the NBM also records typical transmission rates, those fields are missing for many firms, and we use the advertised rates instead. Technologies covered in the NBM include the ADSL, fiber, and cable modem services we investigate in this essay, as well as wireless and less commonly used wireline technologies.

In general, information on potential entry that did not occur cannot be found in the NBM. However, since the markets are defined around ILEC locations, the ILEC ADSL potential entrant is obvious. The ADSL quality choices of ILECs in California markets are shown in *Figure 6*.

---

[39] Created from a collaboration between the National Telecommunications and Information Administration, the FCC, and all states, territories and Districts of the US, the NBM is an online tool that provides semi-annual information on the broadband service providers, their product type, technology, and their maximum advertised upload and download speeds in each US census block.

[40] Service providers are aggregated to the level of their holding company, even if they operate in the same market with multiple operating companies. We used a master list of holding companies constructed by one of the authors for previous broadband research that includes all firms appearing in the FCC Form 477 broadband filings in recent years. Our list of holding companies account for variation in company names, mergers, acquisitions, spin-offs, and cable system area swaps.



Data sources: National Broadband Map, Dec 2011 - June 2013 and authors' calculations.
Note: MADTR is maximum advertised downstream rate

Figure 6: ILEC ADSL Quality Choice in California

The data include 14 ILECs in 965 markets over four periods, choosing among five quality alternatives, for a total of 3,860 cases and 19,300 observations available for the estimations. The first-stage estimations include observations for 25 cable companies (4,365 cases and 26,190 observations) and 16 CLECs (37,743 cases and 226,458 observations for ADSL; 52,234 cases and 313,404 observations for fiber). The set of potential entrants for CLECs for a market and technology type includes any CLEC offering service anywhere in California (except when the CLEC is already an ILEC in the market). The set of cable modem entrants includes any firm with a franchise area that at least partially overlaps a market.

For cable firms, the locations for entry into broadband service provision are limited by the extent of their franchise areas. In California, new cable franchises are awarded by the state, and the CPUC makes available GIS shapefiles of statefranchised areas.[41] We used these data to construct a variable measuring what fraction of the market area the franchise area covers. In one of our robustness tests, we weight the cable modem competition variables by this variable to account for market coverage that is less than complete.

4.3 Demographic data

Most of the demographic data for the markets come from Geolytics, based on the 2010 Census and 2008-2012 American Consumer Survey data from U.S. Census Bureau for CBGs. However, to improve the precision of the population and household density variables, we instead counted population and households in Census blocks falling into our step-two market areas, and divided each by the step-two market area in square miles.[42] Similarly, the regressor for market area is for the step-two definition. We used the County Business Patterns 2011 data of the U.S. Census Bureau to get information on finance, insurance, and real estate (FIRE) employment in our markets.[43] *Table 13* contains summary statistics for the variables used in the study (see next page).

---

[41] Not all franchise areas were awarded by the state in the past, however. Some legacy locally-awarded franchise areas with long terms are missing, and the fraction coverage variables described below in the text are missing for those. As local franchises expire, they are converted to state franchises.

[42] This avoids the difficulty that some of the step-three market areas are overly large in rural areas, even though the locus of economic activity is near the central office. If we included population and area of the entire step-three market areas, the resulting densities would be misleadingly low.

[43] The county level employment data were linked to our markets by calculating which county or counties each market is in. Since the FIRE employment variable describes the composition of employment in the county instead of counting employees, it is reasonable to apply these county-level data to our markets. When a market falls into more than one county, the data from the multiple counties is averaged, weighted by the market area falling into each.

	Mean	s.d.	Min	Max
Y (chosen alternatives)	3.154	1.197	0	4
Competition variables†				
Cable modem service				
CM <10 M	0.853	0.355	0	1
10M < CM < 25M	0.826	0.379	0	1
25M < CM < 50M	0.761	0.427	0	1
50M < CM <100 M	0.738	0.440	0	1
100M < CM	0.471	0.499	0	1
CLEC ADSL				
768K < CLEC ADSL <3M	0.634	0.482	0	1
3M < CLEC ADSL <6M	0.509	0.500	0	1
6M < CLEC ADSL <10M	0.458	0.498	0	1
10M < CLEC ADSL <25M	0.384	0.486	0	1
25M < CLEC ADSL <50M	0.109	0.312	0	1
CLEC fiber > 1G	0.223	0.417	0	1
(i) Main Demographics				
Area (log mi)	2.473	0.460	-1.150	2.986
Pop. density (log)	6.479	3.082	-2.387	11.327
Pop. growth	0.110	0.801	-0.916	22.752
Age	37.854	6.136	20.093	62.661
Education (grade)	13.694	1.690	7.895	17.445
Rental housing (%)	0.413	0.177	0.010	1.000
Work at home	0.204	0.070	0.000	0.638
Water area (%)	0.013	0.035	0.000	0.493
Income (log)	11.090	0.399	9.322	12.544
Vacancy rate (log)	-2.339	0.786	-4.615	-0.109
FIRE employment (log %)	-2.831	0.272	-3.597	-2.054
(ii) Additional Demographics				
HH density (log)	5.424	3.044	-3.294	10.205
Nonwhite %	0.270	0.174	0.000	0.884
Female %	0.493	0.051	0.038	0.631
Education, s.d.	514.379	342.021	63.649	2262.061
Long commute %	0.199	0.106	0.000	0.868

†The competition variables are defined to represent all possible speeds that competitors could offer in the market. When the maximum speed in the NBM of competitors is in a particular category, then the indicator variables equal 1 for that and lower speed categories.

Table 13: Summary Statistics

5. Results

The conditional logit estimations return a large number of estimated coefficients, since each regressor not varying over alternatives has a different coefficient for each of the four alternatives apart from the baseline choice. In our main estimation, we have 84 coefficients.

While exponentiated coefficients from a conditional logit estimation have meaning as odds ratios relative to the base alternative, it is often more natural for econometricians and policy analysts to think in terms of marginal effects. The marginal effect of a regressor is the impact of a one unit increase in the regressor on the probability that the firm chooses a particular quality alternative. When the regressor is log(x), 0.01 times the marginal effect measures the impact of a 1% increase in x. Given our interest in the top end of the quality ladder, we show marginal impacts on the top three quality categories.[44]

In the tables, we present the marginal effects at the median (*MEMdn*) and average marginal effects (*AME*) for the regressors of interest instead of the coefficients.[45] With MEMdn, the marginal effect is calculated once, setting all covariates at their median values. With AME, the marginal effect is calculated for each observation in the sample using actual values of regressors, with the results then averaged over the sample.

---

[44] We focus on the top end, in addition, because few ADSL offerings by ILECs are in the 0 or 1 categories anyway.

[45] Another reason not to present the coefficients is that the marginal effects are functions of all the coefficients, and thus it is possible for the ME of a regressor to be statistically significant even when its coefficient is not. Thus checking for significance stars on coefficients can give a misleading sense of which regressors are truly important.

5.1 First-step results

In the first step of estimation, the quality choices of competitors are regressed on the market demographics and the infrastructure variables proxying the costs of the firm. Since the first step is akin to a reduced form forecasting exercise, we err on the side of including a large set of predictors without regard for causal meaning of the coefficients, parsimony, or the significance of the estimates. The demographics include those also included in the second step: market area in miles, population density and growth rate, median household income (averaged across Census block groups in the market) average age and age squared, average highest educational grade level achieved, the fraction of housing units that are rented or vacant, the fraction of the labor force working at home, the proportion of area employment that is in the financial, insurance, or real-estate (FIRE) sector, and the fraction of the market area that is under water. Where these variables are right skewed they are in logs, as noted in the tables. These variables were chosen based on a review of previous literature on the determinants broadband entry decisions. An additional set of demographics are included only in the first step estimations: the density of households in the market, the fractions of nonwhite people, the percentage female, the standard deviation of education attainment, and the proportion of workers with long commutes. These variables are not included in the second step because of concerns about near multicollinearity with other demographic variables or insignificance and for the sake of parsimony in presenting results.

While we do not present the results from the first step in tables here, we note two things. The infrastructure variables are clearly highly relevant. The coefficients for *NearestSameSpeed* and *SameSpeednotFound* are statistically significant at the 1% level for all competing broadband types. The coefficients for *NearestAnySpeed* and *AnySpeednotFound* (where the latter variable is defined similarly to *SameSpeednotFound* but across all speed categories > 0) are also generally (but not uniformly) highly significant. The high significance and impact of these infrastructure variables implies that they are likely to be effective instruments to identify separately the impact of the competition variables in the second-step estimation. We also note that some of the demographic variables have insignificant coefficients, even those that we would expect to have strong impacts on quality choice. While that does not mean that they have no significant marginal effect on choice probabilities (see footnote 45), it does mean that the infrastructure variables alone are capturing much of the variation in quality choice. The same first step estimates of Ew_{imt} are used for all the second-step specifications.

5.2 Second-step results

Estimation 1. We begin with a simple specification for ILEC ADSL quality choice in which only demographic variables are included. *Table 14* contains the marginal effects, MEMdn and AME.

	Prob(3M < ADSL speed <6M)				Prob(6M < ADSL speed <10M)				Prob(10M < ADSL speed <25M)			
	ME at median		Average ME		ME at median		Average ME		ME at median		Average ME	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE
Demographics												
Area (log mi)	-1.022	1.953	-1.294	2.037	-9.903***	3.238	-8.533***	3.044	9.883**	3.897	7.194*	3.674
Pop. density (log)	-0.447	0.435	0.145	0.443	0.937	0.815	1.551**	0.643	0.284	0.821	1.355**	0.635
Pop. growth	0.600	1.100	0.470	1.190	-7.035	4.691	-6.111	4.022	6.104	4.086	4.408	3.001
Age	-1.645	1.046	-1.475	1.101	-1.498	2.023	-0.921	1.753	3.246	2.045	3.015*	1.711
Age squared	0.021*	0.012	0.019	0.013	0.008	0.025	0.003	0.021	-0.029	0.025	-0.028	0.021
Education (grade)	-0.068	0.751	0.110	0.768	1.475	1.358	1.464	1.160	-1.142	1.349	-0.549	1.136
Rental housing (%)	-14.497**	7.313	-13.237*	7.016	-37.210***	12.456	-28.700***	10.136	53.210***	12.739	45.438***	10.415
Work at home	-8.695	11.971	-10.910	12.095	-31.780	22.402	-28.841	18.529	37.473	23.134	26.151	19.458
Water area (%)	32.732*	17.782	39.309**	18.081	-87.169*	48.996	-67.503	41.034	61.927	44.643	55.993	35.660
Income (log)	-6.483*	3.761	-4.939	3.656	-8.468	6.624	-4.726	5.486	16.684**	6.867	16.539***	5.747
Vacancy rate (log)	2.466	1.841	2.258	1.790	10.395***	3.244	8.296***	2.622	-12.939***	3.247	-10.712***	2.682
FIRE employment (log %)	-5.347*	2.828	-4.563	2.774	-22.731***	6.094	-17.710***	4.899	28.384***	6.190	24.037***	4.945

*10% sig level ** 5% sig level ***1% sig level.

Table notes: Estimation method is Conditional Logit. Since there are no competition variables, the two step estimation described in the text (refer to section IV.B in the text) is not required in this specification. SE's are robust to clustering on markets. "ME at median" is the marginal effect of a one unit increase in the regressor in the row label on the choice probability given in the column superheading, calculated at the median values of all regressors. "Average ME" is the marginal effect of a one unit increase in the regressor in the row label on the choice probability given in the column superheading, calculated at actual regressor values and averaged over the sample. Choice alternatives not shown in the table but included in the estimation are 1) no ADSL broadband, and 2) Prob(768K < ADSL speed <3M).

Table 14: Estimation 1 – demographics variables only

The marginal effects are expressed in percentage points. Here we focus mainly on the marginal effects for the highest speed ADSL, contained in the rightmost set of columns. This speed, from 10 to 25 Mbps, ("high-speed ADSL" in the following discussion) is offered in California by ILECs held by 10 holding companies, the largest of which is AT&T offering its U-verse service.[46] Another four firms offer ADSL only with lower speeds.[47] The results show that several of the demographic variables significantly[48] increase the probability of offering high-speed ADSL: income, population density, age (at the 10% level only), rental housing %, and FIRE employment. One variable, the vacancy rate, significantly lowers the firm's

---

[46] The holding companies of these service providers are AT&T Inc., Calaveras Telephone Company, Frontier Communications Corporation, LICT Corporation, Ponderosa Communications, Inc., Sebastian Enterprises , Sierra Tel Communications Group, SureWest/Consolidated, Telephone and Data Systems, Inc., and Volcano Communications Company.

[47] The holding companies of these service providers are: Bryan Family Inc., Siskiyou Telephone Co., VARCOMM, Inc., and Verizon Communications. Verizon offers lower speed ADSL in some markets, but for higher qualities offers subscribers fiber (FIOS) instead.

[48] Here we mean "significant in either MEMdn or AME," and so below as well.

probability of offering high-speed ADSL. For an example of interpreting the numbers, consider the income variable. The MEMdn for log income, 16.68, implies that an increase of market-area household income of 10% increases the probability of the ILEC offering high-speed ADSL by 1.67 percentage points. The MEM for log income, 16.54, is similar in this case, although we observe that often the AME's are somewhat smaller than the MEMdn's.

Estimation 2. Estimation 2 repeats the previous specification but with the competition variables included (see *Table 15*).

	Alternative 2 Prob(3M < ADSL speed <6M)				Alternative 3 Prob(6M < ADSL speed <10M)				Alternative 4 Prob(10M < ADSL speed <25M)			
	MEMdn		AME		MEMdn		AME		MEMdn		AME	
	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE	Est.	SE
Competition variables†												
CM <10 M	-9.846	7.332	-5.055		39.410***	14.640	32.850		-26.950**	12.110	-19.072	
10M < CM < 25M	43.690***	12.730	33.305		-31.520**	15.870	-22.652		-11.910	11.700	-12.425	
25M < CM < 50M	40.230***	11.610	43.626		-28.470***	10.240	-27.666		-11.770*	6.197	-15.258	
50M < CM <100 M	-76.51***	5.089	-71.584		31.460***	5.296	29.042		45.820***	5.658	46.794	
100M < CM	-11.15***	3.926	-8.644		-9.640*	4.925	-8.585		20.520***	4.956	15.782	
768K < CLEC ADSL <3M	1.388	5.856	3.524		-2.155	6.883	0.319		2.129	7.114	5.269	
3M < CLEC ADSL <6M	-5.243	6.984	-5.304		2.416	14.730	1.387		2.220	14.070	0.391	
6M < CLEC ADSL <10M	-8.570	5.639	-8.353		40.230***	15.550	26.689		-33.750**	14.270	-30.193	
10M < CLEC ADSL <25M	0.869	2.938	0.225		-23.050***	8.682	-17.865		22.030***	7.460	17.120	
25M < CLEC ADSL <50M	-1.176	3.037	-0.088		-39.920***	5.804	-28.432		42.920***	5.925	34.986	
CLEC fiber > 1G	-13.11***	3.540	-9.706		11.280	8.536	12.109		2.555	8.439	8.028	
Demographics												
Area (log mi)	-2.676	2.545	-1.680	1.821	-7.311**	3.703	-4.717*	2.697	10.048**	4.471	7.659**	3.446
Pop. density (log)	0.116	0.664	0.598	0.467	1.061	0.950	1.344**	0.638	-0.971	0.908	0.205	0.627
Pop. growth	2.463	2.027	1.299	1.515	-13.576**	5.813	-9.899**	4.128	10.956**	4.957	7.180**	3.140
Age	-0.010	1.581	-0.020	1.114	-2.889	2.288	-1.975	1.709	2.875	2.272	2.051	1.673
Age squared	0.005	0.018	0.004	0.013	0.025	0.027	0.017	0.020	-0.030	0.027	-0.020	0.020
Education (grade)	-0.286	1.048	0.067	0.753	0.617	1.689	0.740	1.225	-0.218	1.632	0.321	1.185
Rental housing (%)	3.607	9.950	2.297	7.010	-53.267***	15.314	-37.238***	10.400	49.788***	14.795	35.859***	10.390
Work at home	-7.971	15.331	-8.766	11.236	-37.606	24.265	-29.685*	16.733	44.447*	23.856	26.861	17.726
Water area (%)	38.599*	23.005	35.697**	16.719	-83.318	52.120	-48.197	35.932	48.603	45.621	50.388	32.816
Income (log)	3.346	4.424	2.934	3.183	-5.649	8.181	-3.133	5.828	2.465	7.845	2.732	5.728
Vacancy rate (log)	-0.668	2.302	-0.650	1.612	13.733***	3.784	9.300***	2.498	-13.188***	3.554	-9.925***	2.491
FIRE employment (log %)	-0.588	4.511	-0.284	3.168	-38.049***	8.272	-26.087***	5.267	38.768***	7.835	28.438***	5.356

*10% sig level ** 5% sig level ***1% sig level.

†For the competition variables, the marginal effect is for moving from the speed category below to the category in the row label. E.g., for the row labeled "10M < CM < 25M" the ME is for moving from facing CM competitors with a max speed of less than 10M ("CM <10 M") to facing CM competitors with a max speed of between 10M and 25M.

Table notes: Estimates are from the second step, and all the regressors labeled "competition variables" are the expected values calculated from the first step (refer to section IV.B in the text). See also notes to previous table.

Table 15: Estimation 2 – competition and demographics variables

In this estimation, the infrastructure variables are still not included in the second step. Thus, the impacts of the competition variables may be biased due to endogeneity. For example, cost factors for firm *i* that are omitted in this regression, such as the presence of previously installed or nearby infrastructure, may be correlated with the quality choices of rivals through unobserved local factors. The estimates show apparently large impacts of competition on the ADSL quality choice. We consider the impact of each type of competitor in turn. The marginal effect given for a speed category *j* is calculated to pertain to changing the competitors' maximum speed category from $j - 1$ to *j*.

The cable modem quality choice seems to affect the ILEC ADSL quality decisions a lot. When the cable modem service is relatively slow, the negative marginal effects for high-speed ADSL indicate that the ILECs are less likely to offer fast service themselves.[49] Once the cable companies move up into the DOCSIS 3.0 speed tiers, 50 Mbps and above, however, the ILEC is more likely to offer high-speed ADSL. The apparent effect of CLEC ADSL competition is similar: when the competitors' offerings are worse quality, the ILECs quality is less likely to be of the highest. The impact of CLEC gigabit fiber is small and insignificant. Given the omission of the infrastructure variables, we do not yet assign causal interpretation to these results. Comparing the results of the first two estimations, we see that the addition of the competition variables did not greatly change the marginal effects of

---

[49] In this work, not all standard errors (s.e.'s) are available. Difficulties with numerical derivatives, and lack of time to program analytic derivatives, leads to the omitted SE's in this and following tables. Furthermore, as is common in the IO literature, the second step s.e.'s do not account for estimation error in the first step. Thus our reported s.e.'s are smaller than those from the valid asymptotic variance-covariance matrix.

the demographics. For this reason and to save space in the tables, we will not show the impacts of the demographic variables in *Table 16* and *Table 17*.

Estimation 3. The addition of the infrastructure variables in the second-step estimation brings us to our preferred estimation (see *Table 16*).

	Alternative 2			Alternative 3			Alternative 4		
	Prob(3M < ADSL speed <6M)			Prob(6M < ADSL speed <10M)			Prob(10M < ADSL speed <25M)		
	MEMdn		AME	MEMdn		AME	MEMdn		AME
	Est.	SE	Est.	Est.	SE	Est.	Est.	SE	Est.
Competition variables									
CM <10 M	-8.215***	2.050	-14.204	-2.950	5.733	9.159	12.190**	6.210	13.605
10M < CM < 25M	11.000*	6.371	17.335	5.262	8.081	-7.888	-16.960	11.820	-15.807
25M < CM < 50M	88.550***	6.431	77.247	-11.400*	6.359	-18.942	-76.460***	10.870	-46.781
50M < CM <100 M	-96.230***	1.008	-84.814	7.890***	1.793	22.649	88.310***	2.359	56.720
100M < CM	-1.865**	0.906	-3.333	-2.413	1.645	0.029	4.285**	2.160	3.874
768K < CLEC ADSL <3M	-2.262	1.529	-0.276	-2.829	2.837	0.784	5.422*	3.277	5.784
3M < CLEC ADSL <6M	-0.466	1.777	-1.912	-0.822	3.334	-1.535	1.198	3.803	0.234
6M < CLEC ADSL <10M	-1.228	1.681	-8.962	10.180*	5.387	9.726	-9.436	6.110	-6.464
10M < CLEC ADSL <25M	-0.629	0.789	1.546	-8.855*	4.568	-3.453	10.030*	5.207	8.767
25M < CLEC ADSL <50M	3.318**	1.571	12.446	-5.660***	1.485	-22.302	1.310	2.409	-3.930
CLEC fiber > 1G	-3.101***	0.822	-9.931	-3.069*	1.774	3.048	6.212***	2.033	7.570
Demographics									
Included but not shown									
Cost/Infrastructure									
SameSpeednotFound (own effect)	-3.142***	0.657	NA	-7.237***	1.293	NA	-88.510***	1.586	NA
NearestSameSpeed (log mi, own effect)	-0.100***	0.018	NA	-0.221***	0.043	NA	-0.307***	0.048	NA
NearestAnySpeed (log mi)	0.313**	0.152	0.324	0.527	0.321	0.036	-0.853*	0.436	-0.785

*10% sig level ** 5% sig level ***1% sig level.

Table notes: Estimates are from the second step. See also notes to previous table.

Table 16: Estimation 3 - competition, demographics, and nearest infrastructure variables

SameSpeednotFound and *NearestSameSpeed* have the expected negative impacts on same-choice alternatives. *NearestAnySpeed* has a further negative impact on the high-speed ADSL choice. After controlling for the same-speed infrastructure, the marginal effect for *NearestAnySpeed* can be interpreted as the impact of the distance to lower-speed infrastructure.

As expected from the discussion of the potential endogeneity problems in Estimation 2, the impacts of the competition variables, while qualitatively similar to before, have very different magnitudes in Estimation 3. The differing results show how the infrastructure variables help control for omitted variable bias. Since this is our main specification, we go through the results in greater detail here. When cable competitors switch from having no service to offering the slowest service (< 10 Mbps), the probability of high-speed ADSL rises by 12 percentage points. Looking at the columns in *Table 16* for alternatives $j = 2$ and $j = 3$, we see that about two-thirds of these 12 percentage points come from upgrading from ADSL of speed between 3 and 6 Mbps, while about a quarter come from upgrading from ADSL of speed between 6 Mbps and 10 Mbps). Thus, whether an ILEC faces any cable competition at all appears to spur investment in ADSL speed. This impact (and those that follow) is not merely from the coincidence of DSL and cable modem service in more attractive markets, because the demographic regressors in the model control for the key market factors of income, population density, and so on. Furthermore, these apparent impacts are not merely reflections of favorable cost conditions for broadband provision, since last period's infrastructure variables account for that. Thus, they likely reflect the strategic considerations of ILECs in California and we interpret them as such.

However, as the cable competitors rise up the quality ladder, the impacts are not monotonic. When cable modem quality rises from between 10 and 25 Mbps to between 25 and 50 Mbps, the ILEC is 76 percentage points less likely to offer high-speed ADSL. Looking at the other columns of the table, we find that probability lost from alternatives 3 and 4 went to alternative 2. There are relatively few observations (88 out of 3,860 cases) with ILEC DSL entrants facing cable competition in speed category 3, and the large negative impact may merely be a small sample phenomenon, statistical significance notwithstanding. However, it may also be that the ILEC is responding with slower broadband to what the cable company did *not* do: upgrade to DOCSIS 3.0.[50] When the maximum speed of the cable modem service rises to the two highest speed categories, the marginal effects on high-speed ADSL are positive. The largest marginal effects are for when the cable competitors upgrade from 25-50 Mbps to 50-100 Mbps (DOCSIS 3.0). The MEMdn for high-speed ADSL is 88.3 percentage points, and the AME is 57.7. The probability gained comes mainly from alternative 2. In summary, ILECs generally respond to cable competition by upgrading their ADSL quality when they face any competition at all and when the quality of the competition becomes high.

In contrast with cable modem competition, there is no strongly significant evidence that ILECs pay much attention to the quality of their CLEC ADSL competition. The results are also in contrast to the previous estimation, in which CLEC ADSL had some highly significant marginal effects on high-speed ILEC ADSL. This difference shows the importance of the infrastructure variables in controlling for omitted variable bias. The largest impact on high-speed ILEC ADSL, 10 percentage points for the MEMdn and 8.8 for the AME, comes from when the CLECs move into the same speed category (between 10 and 25 Mbps). However, the MEMdn is not significant at the 5% level. The generally weak response to CLEC ADSL quality may indicate that ILECs do not perceive CLECs in California to be

---

[50] The DOCSIS 2.0 standard can provide maximum usable throughput up to 38 Mbps downstream and up to 27 Mbps upstream. DOCSIS 3.0 can dramatically increase downstream and upstream capacity by a factor of the number of channels used in the network. This means that DOCSIS 2.0 could be used to offer service (typically) in the 10 to 25 Mbps range, whereas moving to the 50 to 100 Mbps range requires an upgrade to a DOCSIS 3.0 platform.

much of a competitive threat to their largely residential-oriented ADSL service. The largest CLEC ADSL provider, by far, is MegaPath (held by Platinum Equity, Inc.), which targets the business market. Finally, the presence of gigabit CLEC fiber spurs a 6.2 percentage point increase (per MEMdn; 7.6 for AME) in high-speed ADSL. While Google fiber does not appear in our data, these results are in line with anecdotal accounts of incumbent broadband providers increasing their quality of service in response to Google fiber elsewhere in the country.[51]

Estimations 4 to 6. Here we briefly consider three additional estimations performed as robustness checks. A subset of the results is in *Table 17,* where only the marginal effects for high-speed ADSL are shown.

	Estimation 4 **Log distance variables**			**Estimation 5** **CM coverage-adjusted**			**Estimation 6** **Expanded set of demographics**		
	MEMdn		AME	MEMdn		AME	MEMdn		AME
	Est.	SE	Est.	Est.	SE	Est.	Est.	SE	Est.
Competition variables									
CM <10 M	4.910	3.128	11.231	11.000**	5.141	13.161	11.740*	7.002	12.096
10M < CM < 25M	-4.850	5.424	-9.504	-19.310	12.740	-17.818	-20.680	13.420	-16.731
25M < CM < 50M	-91.610***	5.722	-46.370	-75.860***	12.170	-46.378	-72.210***	12.130	-45.683
50M < CM <100 M	91.710***	2.702	47.058	88.880***	2.330	56.697	86.870***	2.676	56.384
100M < CM	4.659***	1.788	7.546	3.925*	2.183	3.722	4.810**	2.446	4.166
768K < CLEC ADSL <3M	1.081	2.527	3.047	6.265*	3.276	6.507	1.618	3.382	2.540
3M < CLEC ADSL <6M	2.997	2.431	-0.033	1.524	3.820	0.337	1.257	4.400	-0.084
6M < CLEC ADSL <10M	-5.894	3.731	-3.093	-9.273	6.161	-6.138	-7.285	6.234	-3.641
10M < CLEC ADSL <25M	2.784	3.559	0.996	9.293*	5.228	7.787	6.009	5.192	4.600
25M < CLEC ADSL <50M	-0.760	3.418	-4.793	1.613	2.502	-2.887	2.285	3.955	-3.664
CLEC fiber > 1G	0.475	2.050	3.022	6.791***	2.028	8.025	2.673	2.916	2.594
Demographics									
Included but not shown									
Cost/Infrastructure									
SameSpeednotFound†	-90.240***	1.541		-88.310***	1.605		-87.410***	1.811	
NearestSameSpeed (log mi)†	-4.660***	0.860		-0.313***	0.047		-0.330***	0.054	
NearestAnySpeed (log mi)	-2.366***	0.553	-4.071	-0.886**	0.448	-0.798	-0.924*	0.487	-0.777

*10% sig level ** 5% sig level ***1% sig level. †Own effect (effect on alternative 4, in this case).

Table Notes: Each set of three columns pertains to a different estimation specification. 1st: distances in the infrastructure variables are in logs instead of levels. 2nd: cable modem competition variables are weighted by fraction of market covered by the cable service area. 3rd: all available demographics are included.

Table 17: Additional ILEC ADSL Estimations (Robustness Checks)

---

[51] For example, in the Austin area where Google Fiber is available, AT&T's U-verse DSL service has offered downstream speeds up to 300 Mbps. See http://www.fiercetelecom.com/story/att-begins-upgrading-austin-customers-1-gbps-service/2014-08-11 (accessed on 8/13/2014).

In Estimation 4, we use the log of the distances for the infrastructure variables instead of the level. The results are qualitatively similar to Estimation 3, except that the marginal effects of the slowest cable modem service and CLEC gigabit fiber lose significance. In Estimation 5, we replace the cable modem competition variables with their market-coverage-weighted counterparts. The results are very close to those of Estimation 3, even for the cable competition variables. Finally, we include all available demographics in Estimation 6. Again, the results are similar, except that the impact of CLEC gigabit fiber loses significance.

6. Discussion

The analysis presented in this chapter shows that ILECs respond to the quality choices of rival broadband providers. Their responses are heterogeneous to the type of provider and to the level of quality. Specifically, ILECs appear to care more about rivals using cable modem or fiber technologies than rivals using a similar ADSL technology. The likely expectation for the ILEC is that if consumers are going to switch services they are more likely to switch to a rival with a technology that can provide a very fast Internet service. Moreover, the level of speed matters in strategic responses. Particularly, when cable modem rivals move from no service to a “low” speed service tier or from a “medium” service tier to a “high” speed service tier, the ILEC also increases speed. This suggests strategic complementarity in the provision of quality and is consistent with the findings from Kugler and Weiss (2013) and Brueckner and Luo (2013) in the reaction-function literature as well as with Goolsbee and Petrin (2004), Savage and Wirth (2005) and Matsa (2011) in the reduced-form competition and quality literature.

Interestingly, however, when cable modem rivals move from "low" to "medium" speed service tiers, the ILEC is less likely to provide high quality ADSL, which suggests strategic substitutability in the provision of quality. This response may be due to changes in the price and quality elasticity of demands as suggested by theory. However, it is possible that this reflects the ILEC's reaction to an underlying capacity constraint facing its rival that does not have DOCSIS 3.0 installed. The result may even merely be an artifact of the data that will not persist once we expand our analysis. More empirical and theoretical analysis is required to fully understand this intriguing result. Overall, our empirical finding of a non-monotonic relationship between the quality choices of rival broadband providers also resembles the findings of Chen and Gayle (2013) from the airline industry.

CHAPTER IV.

THE IMPACT OF HIGH-SPEED BROADBAND AVAILABILITY ON REAL ESTATE VALUES

1. Introduction

An extensive literature on broadband adoption is available. However, empirical studies related to the economic impact of fiber technology are fewer in number. Chapter IV explores whether people are willing to pay more for real estate located in areas where fiber broadband access is available than for a property that does not offer this amenity.[52]

Numerous factors influence the value of residential real estate, including the energy efficiency of buildings, the proximity of good schools, or the amount of crime in neighborhoods. For some people, an important consideration when buying a home might be the availability of fiber-based broadband services to the property. To test this assumption, the research presented in this chapter aims to evaluate whether access to fiber broadband is associated with any measurable effect on property values. Using a hedonic price framework and data from the National Broadband Map (NBM) the research goal is to investigate how constant-quality real estate prices vary, where constant-quality real estate is defined as a property where structural, land, and community attributes are all held constant. The focus of this

---

[52] I would like to thank Professor Dale N. Hatfield for his contribution in coming up with the initial research idea.

investigation is the hypothesized impact of variations in fiber service availability on residential single-family house prices. This research adds to the existing literature by conducting an empirical analysis of the assumed neighborhood effect of fiber availability, with the ultimate objective of measuring the value of broadband Internet throughout real estate markets across the United States.

Existing literature has examined the economic impact of broadband penetration, but not that of fiber-based Internet access. Using 2011 June data from the NBM, the recent Broadband Brief by the Department of Commerce's National Telecommunications & Information Administration and the Economics and Statistics Administration confirmed that "broadband is less available in rural areas than in urban areas" (NTIA-ESA, 2013, p. 11). The NTIA-ESA analysis also showed that proximity to central cities within a Metropolitan Statistical Area (MSA) is likely to be "more strongly associated with the availability of the highest speed levels of broadband service than population density" (p. 10). The broadband brief, however, leaves the question open whether the location within an MSA is simply associated with increased broadband availability or whether it is a contributing factor to increased broadband availability.

Real estate economists often quantify the impact of variables that are specific to neighborhoods by applying the hedonic method outlined in the seminal paper of Rosen (1974).[53] These hedonic valuation models assume that the main considerations of property values, such as structural characteristics, neighborhood characteristics, and relative location of the property, are known.

---

[53] Hedonic valuation models are regressions of real estate value against property characteristics that determine this value.

The first models focused on the structural characteristics of the property, including the size of the building, the number of bedrooms and bathrooms, and other lot characteristics. Later, other area amenities, such as air pollution (Rosen, 1979), local climate (Haurin, 1980), and crowding (Roback, 1980) were added. Roback (1982) also considered labor markets in her approach. The empirical studies of Beeson and Eberts (1989), Peek and Wilcox (1991), Blomquist and Berger (1992), and Potepan (1994) found that crime, recreational opportunities, and population demographics should also be considered for real estate valuation models.

Despite recent advances in real estate economics, spatial econometrics, and the increasing number of studies that support the existence of neighborhood effects, the impact of fiber-based broadband on property prices is still not a well-researched area. Academic research regarding this topic has been limited by a lack of good quality data on fiber broadband access availability. Although a recent research by RVA LLC found that “a fiber connection adds between $5,300 and $6,450 to the value of a home” (RVA, 2013, p. 31), their study was based on surveying homebuyers and developers, and it was not an empirical analysis of transactional data.

The NBM has data to allow for investigation of this research question (NBM, 2011). The NBM shows where broadband is available, the technology used to provide the service, the maximum speeds, and the name of the service providers. Created from collaboration between the National Telecommunications and Information Administration, the Federal Communications Commission (FCC), and all states of the US and territories and the District of Columbia, the NBM is an online tool that provides semi-annual information on the availability, technology, speed, and location of broadband Internet access at the census block level. Matching fiber broadband availability information from the NBM with factual information on real estate sales transactions and property characteristics will not only make it possible to investigate the economic impact of superior broadband, but it also provides another approach to measure the value of fiber broadband in monetary terms.

This third chapter and the model used in the essay were inspired by the empirical research of Haurin and Brasington (1996). Using two variants of a random coefficients model and testing transactional data from six MSAs in Ohio, Haurin and Brasington studied the impact of school quality on real estate prices. They found that public school quality positively influences real constant-quality house prices. For simplicity, we decided to follow their hedonic model for the study in Chapter IV. Using information from the NBM and county assessors' data for residential single-family houses, we tested 2011 transactions from nine counties of three MSAs in the State of New York. The early results suggest that the availability of fiber broadband might be as important in explaining spatial variations in real constant-quality house prices as the presence of cooling capability/air conditioning, fireplaces, or a pool.

The next section gives a brief overview about real estate valuation techniques. The empirical model is described in Section 3 and Section 4 details the data. The preliminary results are presented in Section 5. Section 6 concludes and describes future work.

2. Real Estate Valuation Techniques

Malpezzi (2002) divides real estate valuation model into three main groups: hedonic valuation techniques, repeat sales methodologies[54], and hybrid models[55]. Hedonic valuation models are essentially regressions of real estate value against property characteristics that determine this value. Hedonic valuation models assume that the main considerations of property values, such as structural characteristics, neighborhood characteristics, and relative location of the property are known. Hedonic price models are derived from Lancaster's (1966) consumer theory, Rosen's (1974) trading model, and Maclennan's (1977) theoretical works. Lancaster suggests that consumer utility is generated not by goods but instead by the characteristics of the goods. Rosen modeled how suppliers and consumers interact assuming a framework of bids and offers for product characteristics. Maclennan's model recognized that observed real estate transaction prices cannot be equilibrium prices and laid down the theoretical foundation for the hedonic models.

---

[54] Repeat sales methods are based on data that directly measure property price appreciation over different periods. Prices from these known time periods are combined to create matched pairs, providing observations of actual transactions on the same property. Repeat sales models have the advantage of controlling for unobserved characteristics of a given property (no omitted variable bias). Bailey, Muth, and Nourse (1963) were the first to propose repeat sales regressions, simply using ordinary least squares (OLS). Case and Shiller (1989) pointed out the disadvantages of using OLS and suggested using another regression technique, generalized least squares (GLS), a statistical technique for estimating the unknown parameters in a regression model. There are two disadvantages of the repeat sales methods. First, frequently traded properties are not necessarily a random sample of all real estate available. Second, the methods often do not consider improvements to properties; the property sold at t_1 may not be identical to the property sold at t_0.

[55] Hybrid valuation models combine hedonic and repeat sales models. They estimate the two models as imposing a constraint that price changes over time are equal in both models. According to Malpezzi (2002), the basis for the hybrid valuation theory is contained in the influential works of Case and Quigley (1991), Quigley (1995), Hill, Knight and Sirmans (1997), and Knight, Dombrow and Sirmans (1995). The primary disadvantage of the hybrid method is that it requires careful matching of time-series and cross-section observations.

Due to the importance of location and neighborhood characteristics in explaining house price variations, more recent developments in house price models are leveraging advances in spatial econometrics. Dubin (1988), Laakso (1997), Karakozova (2005), Kiel and Zabel (2007) all found empirically that characteristics of the vicinity significantly affect real estate prices. According to LeSage and Pace (2009), there is sound justification to use spatial econometric models in all of the valuation methods described above, as the omitted location and neighborhood variables are considered to be autocorrelated. A future version of this research will also expand on the hedonic model, which is described in the next section, to include the latest advances of spatial econometrics to address potential issues due to selection bias and endogeneity bias.

3. Empirical Model

In the spirit of Haurin and Brasington (1996), this essay tests a simple hedonic price equation.

$$\ln V_{ij} = X_{ij}\beta + J_j\delta_j^{'} + \varepsilon_{ij} \qquad \textit{(IV-1)}$$

In this equation, the coefficient $\delta_j^{'}$ represents percentage deviation of an average house price in district j from the price of a constant-quality property.

The capitalization test for the community and MSA variables is as follows:

$$\delta_j^{'} = Z_j\gamma^{'} + \mu_j^{'} \qquad \textit{(IV-2)}$$

In equation *(IV-2)*, $\delta_j^{'}$ is related to the community and MSA level variables Zj.

Equations *(IV-1)* and *(IV-2)* test for an impact through changes in the lot price. Depending on the land size, the impact varies amongst properties within a given census block group. Combining equation *(IV-1)* with *(IV-2)*, the hedonic price equation can be written as:

$$\ln V_{ij} = X_{ij}\beta + Z_j\gamma' + \mu_j' + \varepsilon_{ij} \qquad \textit{(IV-3)}$$

Equation *(IV-3)* relates the natural log of the real transaction prices for houses *(lnV)* to a set of structural and land characteristics **X**. Using GLS is appropriate because we test for CBG-specific mean zero random errors in house prices.

The parameter of interest in equation (*IV-3*) is $\left(\frac{\partial \ln V_{ij}}{\partial FIBER\_D}\right) = \gamma_f'$. The coefficient γ_f' indicates the percentage deviation of an average house price in CBG j, where fiber broadband is available, from the price of a constant-quality property. Failure to reject the null hypothesis $\gamma_f' = 0$ provides evidence that the presence of fiber in the census block group may have an impact on real estate value.

4. Data and Variables

The primary source of the property information is a dataset containing real estate transactional data and property characteristics for single-family detached houses in three MSAs in the State of New York.[56] The data set was obtained from DataQuick (2013), a property information service provider. The master dataset included property characteristics and assessor data for a total of 24,784 sale

---

[56] The three MSAs are: Buffalo-Cheektowaga-Tonawanda, Poughkeepsie-Newburgh-Middletown, and Rochester. Table 20 shows key characteristics of the nine counties in these three MSAs.

transactions for single-family detached houses in 2011. Fiber broadband availability data were obtained using the June 2011 version of the NBM. Detailed definitions of all variables are listed in *Table 18.*

Variable	**Description and data source**
logHOUSEPRICE	Log of transaction amount for residential single-family house. Source: DataQuick (2013)
AGE10	Age of house in ten years. Source: DataQuick (2013)
LOT SIZE10k	Lot size in ten thousand square feet. Source: DataQuick (2013)
HOUSESIZE1k	House size in thousand square feet. Source: DataQuick (2013)
GARAGESIZE1k	Garage size in thousand square feet. Source: DataQuick (2013)
NBRBATH	Number of bathrooms. Source: DataQuick (2013)
PATIOPORCH_D	Patio & porch dummy. Source: DataQuick (2013)
FIREPLACE_D	Fireplace dummy. Source: DataQuick (2013)
COOLING_D	Cooling solution dummy. Source: DataQuick (2013)
POOL_D	Pool dummy. Source: DataQuick (2013)
NBRBATH	Number of bathrooms. Source: DataQuick (2013)
Q1SALE	Dummy variable to indicate Quarter 1 sales. Source: DataQuick (2013)
Q2SALE	Dummy variable to indicate Quarter 2 sales. Source: DataQuick (2013)
Q3SALE	Dummy variable to indicate Quarter 3 sales. Source: DataQuick (2013)
FIBER_D	Availability of fiber-based Internet access technology in the CBG in 2011. Source: NBM (2011)
TAX_CNTY	Nominal property tax rate. Source: The Tax Foundation (2013)
INCOME1k_CBG	Per capita income in the CBG (in thousands). Source: ACS (2011)
NWHITE_CBG	The percentage of nonwhite households in the CBG. Source: ACS (2011)
DISTANCE_CBD	Calculated distance from the property to the MSA's center (in miles). Source of MSA center geocodes: Holian and Kahn (2012)
TURNOVER_CNTY	Percentage of households who lived in the same house or in the same county 12 month ago. ACS (2011)
ACCESS_CNTY	Weighted average of the average commuting time to work. Source: ACS (2011)
ARTREC_CNTY	Percentage of employees in the art & recreation sector. This is a measure of art & recreation opportunities. Source: CBP (2011)
POPGR_CNTY	2010 county population divided by 2000 county population. Source: GeoLytics (2012)
CRIME_MSA	Serious crimes including murder, robbery, etc. This is an MSA-level variable. Source: FBI (2011)

Table 18: Variable Descriptions

Other explanatory variables in (3) are drawn from various sources, including data from US Census (2011), ACS (2011), and Geolytics (2012). As described in equation (IV-3), our test relates the natural log of the real estate transaction prices to a set of structural and neighborhood characteristics and several jurisdictional amenities.

Measures of the *house and lot characteristics* included the age of the house (AGE10, measured in ten years), lot size (LOTSIZE10k, measured in 10,000 square feet), house size (HOUSESIZE1k, measured in 1,000 square feet), garage size (GARAGESIZE1k, measured in 1,000 square feet), and number of bathrooms (NBRBATH). We used dummy variables to indicate the presence of a patio and/or a porch (PATIOPORCH_D), a pool (POOL_D), air conditioning or some cooling solution (COOLCODE_D), and a fireplace (FIREPLACE_D).

Measures of *neighborhood characteristics* included average income per capita (INCOME1k_CBG), expected county population growth (POPGR_CNTY), tax rate (TAX_CNTY, measured in percentage), and the number of serious crimes per capita in the MSA (CRIME_MSA). INCOME1k_CNTY is defined as the 2011 per-capita income in the county, measured in thousands of dollars. We measured expected county population growth by the ratio of 2010 to 2000. TAX_CNTY is a public sector variable, and it is the nominal tax rate used in the county. The MSA-level measure of crime is the number of serious crimes per capita.

For *neighborhood amenities*, we adopted recreational and arts opportunities (ARTREC_CNTY), accessibility (ACCESS_CNTY), and the distance of the real estate to the central business district (DISTANCE_CBD, measured in miles).

ARTREC_CNTY is a variable we used to proxy the recreational and arts opportunities. We defined ARTREC_CNTY as the percentage of employees in the county who work in the arts, entertainment, and recreation sector. We measured accessibility by the average time in minutes to get to the workplace by those who commute to work. The distance to the central business district was defined as the geographic distance between the geocoded location of the property and the latitude and longitude coordinates for the central business district of the principal city in each MSA.[57]

We are interested to study the impact of *fiber availability* on real constant-quality house prices. Therefore, we used information from the NBM (2011) to identify census block groups where fiber technology was present. FIBER_D is the dummy variable indicating the presence of fiber in a census block group.[58] Since our analysis is at the census block group level, we considered fiber available in a census block group if the technology was reported in at least one of the census blocks.[59]

Other *jurisdictional variables* in the estimation are the percentage of non-white households (NONWHITE) and the percentage of people who lived in the same county twelve months ago (TURNOVER_CNTY). The former is used to capture variations in house price resulting from discrimination, and the latter is a measure to proxy community stability.

---

[57] The location of each MSA's CBD was obtained from the research of Holian and Kahn (2012)

[58] Census block groups are small statistical subdivisions of census tracts. Census tracts typically coincide with the limits of cities, towns or other administrative areas. They contain 1,500 to 8,000 people; on average, they made up approximately four census block groups. There are 217,740 block groups nationwide, as of the 2010 census.

[59] Previous market structure studies have used census tract, county, local telephone exchange and zip-code boundaries to define the geographical market for broadband Internet (Gillett and Lehr, 1999; Prieger, 2003; Wallsten and Mallahan, 2010; Xiao and Orazem, 2011; Nardotto et. al., 2012). Because ISP decisions to roll out and promote new services are usually made for smaller geographical footprints, we define the market for fiber broadband to be a census block group, which generally contains between 600 and 3,000 people.

Table 19 presents summary statistics for all the variables used in our empirical analysis. Because some of the properties in our dataset have missing or incomplete data, our final dataset is comprised of 20,521 real estate transactions in the three MSAs of the study. *Table 20* shows descriptive statistics for the selected nine counties.

Variable	Obs.	Mean	s.d.	Min	Max
HOUSEPRICE	20521	155036.7	128501.7	4200	2250000
lnHOUSEPRICE	20521	11.65183	0.810398	8.336308	14.59855
AGE10	20521	5.062516	3.231351	0	292
LOTSIZE10k	20521	1.632267	1.630365	0.04356	8.712
HOUSESIZE1k	20521	1.63024	0.638391	0.384	9.146
GARAGESIZE1k	20521	0.343415	0.248996	0	12.324
PATIOPORCH_D	20521	0.744359	0.436231	0	1
FIREPLACE_D	20521	0.44145	0.496572	0	1
COOLING_D	20521	0.392232	0.48826	0	1
POOL_D	20521	0.078944	0.269657	0	1
NBRBATH	20521	1.971395	0.862973	0	8
Q1SALE	20521	0.186687	0.389669	0	1
Q2SALE	20521	0.26295	0.440246	0	1
Q3SALE	20521	0.297987	0.457385	0	1
Q4SALE	20521	0.252376	0.434386	0	1
TAX_CNTY	20521	2.676215	0.360779	1.77	3.02
FIBER_D	20521	0.680181	0.466418	0	1
INCOME1k_CBG	20521	63.96973	26.28247	8.466	175.481
NWHITE_CBG	20521	0.148656	0.200824	0	1
DISTANCE_CBD	20521	11.90923	9.044666	0.238185	46.71218
TURNOV_CNTY	20521	0.958216	0.008421	0.921552	0.965054
ACCESS_CNTY	20521	23.70592	4.07874	20.92004	33.88074
ARTREC_CNTY	20521	1.599454	0.353301	1.209812	2.04769
POPGR_CNTY	20521	0.66925	4.000225	-3.28593	9.211786
CRIME_MSA	20521	3.44403	0.816763	2.41	4.392

NOTES. Obs. is number of observations. s.d. is standard deviation.

Table 19. Summary Statistics

MSA/COUNTY	Total Population	Area Size	Pop. Density	Housing Units	Properties in Sample
Buffalo-Cheektowaga-Tonawanda, NY					
Niagara County	216,469	1139.7	414.4	99,120	1,629
Erie County	919,040	1226.9	881.4	419,974	6,914
Poughkeepsie-Newburgh-Middletown, NY					
Orange County	372,813	838.6	459.3	137,025	1,695
Dutchess County	297,488	825.3	373.9	118,638	1,407
Rochester, NY					
Livingston County	65,393	640.3	103.5	27,123	309
Monroe County	744,344	1366.7	1132.6	320,593	6,937
Ontario County	107,931	662.5	167.6	48,193	822
Orleans County	42,883	817.4	109.6	18,431	222
Wayne County	93,772	1383.1	155.3	41,057	586
9 County average	317,792	988.9	421.9	136,684	2,280
US average (3143 counties)	98,232	1208.0	259.3	41,904	-

Source: US Census (2010)

Table 20. County Population, area size, and housing units.

The average house price is $155,036. The average lot size is 16,322 square feet, and average building size is 1,630 square feet. The average garage size is 343 square feet. On average, 74% of the properties have a patio or porch, 44% have fireplace, 39% have a cooling solution installed, and 8% have a pool. A typical house has two bathrooms and is 50.6 years old.

5. Results

Table 21 reports the results based on the 20,521 observations located across 2,180 census block groups in the nine counties of the three selected MSAs in New York State.

		OLS		**OLS (*outliers disregarded*)**		**Robust regression**	
	Coefficient	Estimate	s.e.	Estimate	s.e.	Estimate	s.e.
House and lot characteristics:							
CONSTANT	β_0	4.971***	(0.739)	5.766***	(0.620)	4.463***	(0.394)
AGE10	β_1	-0.053***	(0.004)	-0.059***	(0.003)	-0.070***	(0.002)
AGE_SQ	β_2	0.0005	(0.0003)	0.0008***	(0.0002)	0.002***	(0.0001)
LOTSIZE10k	β_3	0.006**	(0.003)	0.059***	(0.006)	0.065***	(0.004)
LOT10k_SQ	β_4	-0.007***	(0.001)	-0.007***	(0.001)	-0.008***	(0.001)
HOUSESIZE1k	β_5	0.360***	(0.027)	0.387***	(0.020)	0.372***	(0.013)
HOUSE1k_SQ	β_6	-0.024***	(0.006)	-0.023***	(0.004)	-0.018***	(0.002)
GARAGESIZE1k	β_7	0.207***	(0.022)	0.164***	(0.018)	0.170***	(0.015)
GARAGE1k_SQ	β_8	-0.020***	(0.004)	-0.016***	(0.003)	-0.044***	(0.010)
PATIOPORCH_D	β_9	-0.031***	(0.008)	0.016***	(0.006)	0.011**	(0.005)
FIREPLACE_D	β_{10}	0.165***	(0.008)	0.154***	(0.006)	0.130***	(0.005)
COOLING_D	β_{11}	0.100***	(0.007)	0.088***	(0.006)	0.060***	(0.005)
POOL_D	β_{12}	0.043***	(0.010)	0.031***	(0.008)	0.025***	(0.008)
NBRBATH	β_{13}	0.117***	(0.007)	0.100***	(0.005)	0.087***	(0.004)
Q1SALE	β_{14}	-0.028**	(0.011)	-0.029***	(0.009)	0.002	(0.006)
Q2SALE	β_{15}	0.035***	(0.010)	0.023***	(0.008)	0.013**	(0.006)
Q3SALE	β_{16}	0.057***	(0.010)	0.044***	(0.007)	0.027***	(0.006)
Neighborhood characteristics:							
FIBER_D	γ_1	**0.026*****	**(0.009)**	**0.029*****	**(0.007)**	**0.040*****	**(0.006)**
TAX_CNTY	γ_2	-0.443***	(0.033)	-0.417**	(0.026)	0.003***	(0.000)
INCOME1k_CBG	γ_3	0.004***	(0.0002)	0.004***	(0.000)	-0.709***	(0.012)
NWHITE_CBG	γ_4	-0.948***	(0.029)	-1.015***	(0.040)	-0.002***	(0.000)
DISTANCE_CBD	γ_5	-0.002*	(0.0008)	-0.003***	(0.001)	-0.370***	(0.019)
TURNOV_CNTY	γ_6	9.084***	(0.776)	8.104***	(0.663)	8.889***	(0.424)
ACCESS_CNTY	γ_7	-0.019***	(0.005)	-0.016***	(0.004)	-0.005**	(0.003)
ARTREC_CNTY	γ_8	-0.292***	(0.034)	-0.298***	(0.026)	-0.231***	(0.018)
POPGR_CNTY	γ_9	0.019***	(0.003)	0.019***	(0.003)	0.011***	(0.002)
CRIME_MSA	γ_{10}	-0.224***	(0.018)	-0.197***	(0.014)	-0.158***	(0.010)
R-squared		0.621		0.722		0.782	

NOTES. Dependent variable is 2011 Log Real Transaction House Price. Sample size is 20,521 transactions in nine counties. S.e. denotes robust standard errors in parenthesis. ***Significant at the 0.01 level; **significant at the 0.05 level; *significant at the 0.1 level.

Table 21: OLS and Robust Regression

5.1 Ordinary Least Square Regression

The house and lot characteristics have the expected signs, and most of them are significant. Increasing age (AGE10) reduces house value. The positive coefficient for AGES10_SQ, although imprecisely estimated, suggests that housing depreciates at a decreasing rate.

The OLS regression shows that increased square footage of the house (HOUSESIZE10k) and square footage of the lot (LOTSIZE10k) both increase the price of the property at a decreasing rate. The presence of a fireplace (FIREPLACE_D), cooling solution (COOLING_D), pool (POOL_D), patio and/or porch (PATIOPORCH_D), and bathrooms (NBRBATH) all increase the value of the house.

The coefficients of the *jurisdictional variables* generally also have the expected sign. Increasing non-white population (NWHITE_CBG), county tax rates (TAX_CNTY), and crime (CRIME_MSA) were associated with decreasing property prices. Increasing per-capita income (INCOME_CBG) and population stability (TURNOV_CNTY) were both associated with greater house value. Positive population growth (POPGR_CNTY) was associated with increased real estate value. Geographic distance from the central business district (Distance_CBD) and better accessibility (ACCESS_CNTY), as expected, were negatively correlated with price. The positive coefficient of *FIBER_D* ($\gamma_1 = 0.026$) implies that the presence of fiber increases the property value in the neighborhood. The percentage effect of the coefficient is 2.6%.[60] The coefficient is significant at the one percent level. The R-square value is 0.621.[61,62]

---

[60] For more details on the interpretation of dummy variables in semi logarithmic equations see Halvoeren and Palmquist (1980).

[61] For robustness, we also tested equation (IV-3) for location differences and grouped the nine counties into two groups. Group East included Dutchess county and Orange County. Group West included all the other counties. When controlling for the location differences in the regression by adding a group identifier dummy, the coefficient for FIBER_D decreased from 0.026 to 0.016 and was less precisely estimated (the result was significant at the ten percent level). The coefficient for the group dummy was 1.05 (significant at the one percent level). The R-squared value and other coefficients in the regression were similar to those reported in Table 21.

[62] We also estimated equation (IV-3) with city-specific dummy variables. The OLS regression results, not reported, are similar to those reported in *Table 21*

5.2 Outliers

The mean of the residuals is 2.5e-11, and the standard error is 0.499. *Figure 7* below shows the distribution of the residuals after running the OLS regression.

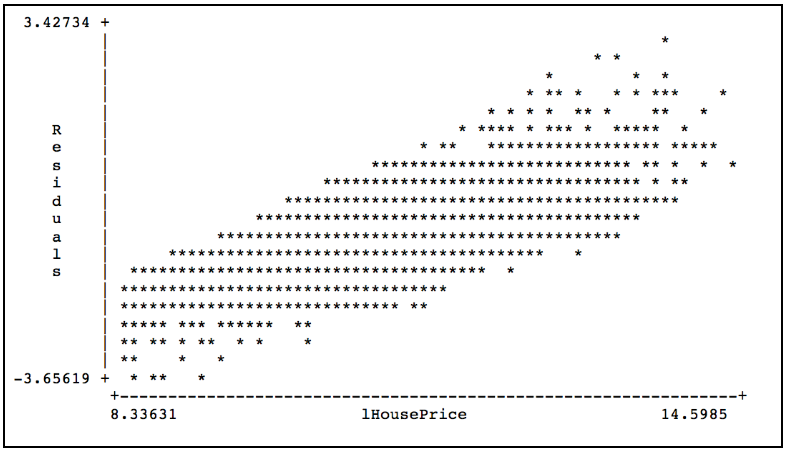


Figure 7: Distribution of the residuals (OLS regression)

Figure 7 also suggests that our observations are contaminated with outliers and/or with influential observations.[63] There are 489 observations, or about 2%, which are three standard errors away from zero. Indeed, when analyzing the residuals, we can find that 353 of the 489 outliers have a transaction price of less than fifty thousand dollars. Column 5 of *Table 20* shows the result of the OLS regression after excluding the outliers them from the analysis. The results of the regression are very similar to those of the OLS model. The coefficient of *FIBER_D* is significant at the one percent level R-square increases from 0.621 to 0.722 if outliers are dropped.

Robust regression is another alternative to use when data is contaminated with outliers or influential observations. It is a compromise between excluding the

---

[63] For the purpose of this analysis, we consider the observations with large residuals (+/- 3 standard deviations from zero) as outliers.

observations entirely from the analysis and including and treating all of them equally in OLS regression. Column 7 of *Table 21* shows the result of our robust regression. The `rreg` command of Stata first runs the OLS regression, then gets the Cook's D value for each observation, and finally it drop any observation with Cook's distance greater than one.[64] Generally, the results are very similar to those of the OLS model. Using robust regression, the R-squared value has increased from 0.621 to 0.782. Also, most of the model coefficients are significant at the 1% level. The house and lot characteristics have the expected signs. Increasing age (AGE10) was associated with reduced house value, and the positive coefficients for AGE10_SQ suggest that housing depreciates at an increasing rate. We were able to confirm that the increased square footage of the house (HOUSESIZE10k) and the square footage of the lot (LOTSIZE10k) were both associated with an increase in the price of the property at a decreasing rate. The presence of a fireplace (FIREPLACE_D), cooling solution (COOLING_D), pool (POOL_D), and bathrooms (NBRBATH) were all associated with increased value of the house. The coefficients of *jurisdictional variables* generally also have the anticipated sign. An increasing non-white population (NWHITE_CBG), county tax rates (TAX_CNTY), and crime (CRIME_MSA) were associated with decreasing property prices. Increasing per-capita income (INCOME_CBG) and population stability (TURNOV_CNTY) were both associated with greater house value. Better accessibility (ACCESS_CNTY) and positive population growth (POPGR_CNTY) were both found to be associated with

---

[64] Stata's ***rreg*** performs an iteractive regression. It first calculates case weights from absolute residuals, and then regresses again using those weights. Iterations stop when the maximum change in weights drops below tolerance. Weights derive from Huber weights and biweights (Li, 1985).

increased real estate value. Distance from the central business district, as expected, was found to be negatively correlated with the price. The positive coefficients of *FIBER_D* ($\gamma'_1 = 0.004$) implies that the presence of fiber in the neighborhood increases property value. The percentage effect of the coefficient is 4.1%. The coefficient of *FIBER_D* is significant at the one percent level.

6. Discussion

The essay in this Chapter IV presented an empirical study of the impact of access to fiber-delivered Internet on real estate values. The research goal was to determine if people are willing to pay more money for real estate located in areas where fiber broadband access is available than for a property that does not offer this amenity. Using information from the NBM and county assessors' data for residential single-family houses from three MSAs in New York State, we applied a hedonic pricing model used in real estate economics. The OLS and the robust regression models signified that the presence of fiber-based broadband was associated with a positive effect on property values in the neighborhood.

However, caution is warranted in drawing conclusions at this point for two reasons. First, correlation does not necessarily equal causation. Fiber availability may drive real estate prices upwards. An unobserved variable may jointly determine both real estate prices and fiber presence. Alternatively, both might be correct. Residential properties in markets with high-speed broadband access would be expected to have greater value. However, good quality broadband infrastructure is also expected to be rolled out first in high-income areas with high-valued real estate. Estimating the value of high-speed Internet availability through property markets creates challenges in addressing this potential endogeneity. Second, this proof-of-concept study was using data from three MSAs only. The real estate dataset is unrepresentative; the selection of these three MSAs was arbitrary. In addition, our exposure time was rather short; the analysis focused on 2011 data alone. Regardless of these limitations, the results in Chapter IV are strong enough to justify future research efforts.

CHAPTER V.

SUMMARY AND FUTURE WORK

1. Summary

This dissertation comprised three multidisciplinary and empirical essays focusing on a different aspect of broadband Internet access markets in the United States. Chapter II investigated the effects of the number of firms and their product-type on broadband Internet quality. The study used an econometric model that relates the actual speeds delivered in census block groups to the number of wireline and wireless Internet service providers (ISPs), cost and demand conditions, and correction terms for the endogeneity of market structure. The research found that wireline speeds are often higher in markets with two or more wireline ISPs than with a single wireline ISP. It also found that increases in wireline speeds are larger in the upstream direction, and there is no relationship between wireline speeds and the number of wireless ISPs. The research work adds to the existing literature by investigating empirically the effects of the number of firms and their product-type on broadband Internet quality.

Chapter III followed a static game theoretic approach to the profit maximization decision of ADSL providers and used a simple two-stage method of estimation of the structural parameters of the ISPs' profit functions. The results included two main findings. First, ILECs improve the quality of their ADSL offerings when a cable player enters the market, and also when cable operators start to offer DOCSIS 3.0 speeds. Second, ILEC ADSL providers do not raise their service quality in response to ADSL competition from CLECs, but they do boost their speed when CLECs deploy fiber in the market. This research represents the first step in a major advance in the empirical analysis of broadband provision, where little structural econometric work has been done.

Chapter IV presented a study to show the impact of fiber-based broadband service availability on real estate values. Using information from the National Broadband Map and county assessors' data for residential single-family houses from three Metropolitan Statistical Areas in the State of New York, this research used a hedonic pricing model to test the hypothesis. Based on the investigation of 20,521 real estate transactions, this study found that the presence of fiber-based broadband is associated with a positive effect on property values, and the value increase is in the range of 2.6 percent to 4.1 percent.

2. Future Research

The first essay investigated the effects of the number of firms and their product-type on broadband Internet quality using performance data in 5,281 CBGs from June, 2011. Future research will include the estimation of wireline market structure on a wider sample of census block groups and time-series analysis. The next step in this research project will also consider combining cross-sectional and longitudinal data with pricing information, and assess the relationship between broadband Internet market structure and prices. An interesting area for enhancement is to use more definitions of quality, including other technical dimensions, functional quality, and the ISP's image characteristics.

The second essay focused on the market entry strategies of ISPs but it only focused on the main competitors of the ADSL players. The next step in this research project will be to re-estimate the model using the alternative 18 kf market definition, as a further robustness check, and to add additional competitors to the estimation of ILECs' quality choice for ADSL. Future work will investigate how CLECs ADSL and cable modem players respond to intermodal and intramodal competition, and will also include competition from fixed and mobile wireless broadband.

The third essay found that fiber availability is positively correlated with real estate prices, but the research had two limitations. First, the real estate dataset was not representative; the analysis used data from three MSAs and the selection of the MSAs was arbitrary. Second, the limited data set and research budget did not allow the study of causation. Simple regression estimates of the correlation are likely confounded by unobserved factors that jointly affect the value of real estate and the decision to deploy fiber. For example, high-quality broadband infrastructure is expected to be rolled out first in high-income and high-educated locations, but these locations also contain high-value real estate. Failure to account for these omitted factors would result in an over-statement of the effect of fiber on real estate values. Future research will work to alleviate this bias with an econometric method that models the firm's decision to deploy fiber; and control for demand-side factors, such as income and education, during the estimation of real estate prices.

3. Concluding Remarks

The combination of large data sets for multidisciplinary academic research unlocks significant value. It makes newly discovered empirical facts available, and the results can improve transparency and policymaking. With the trends that lead to decentralized innovation and decentralized market forces, the use of big data is becoming a key basis of greater understanding of Internet Service Providers' behavior. In addition, it will also lead to improved policy making. Former FCC Chairman Julius Genachowski once remarked,

"policymaking is only as good as the facts and data on which decisions are based." Indeed, fact-based communication policy-making is important for two reasons. First, quality data drives quality decisions. Second, factual and public data improves transparency, and transparency can protect against decision-making processes being captured by partial interests.

BIBLIOGRAPHY

Abbott, L. (1955). *Quality and Competition; an Essay in Economic Theory*. Columbia University Press.

Aghion, P., & Howitt, P. (2009). *The Economics of Growth*. MIT Press.

Aguirregabiria, V., & Mira, P. (2007). Sequential estimation of dynamic discrete games. *Econometrica, 75*(1), 1–53.

Aguirregabiria, V. (2014). Empirical Industrial Organization: Models, Methods, and Applications. Unpublished Manuscript, Cambridge University Press.

Bailey, M. J., Muth, R. F., & Nourse, H. O. (1963). A Regression Method for Real Estate Price Index Construction. *Journal of the American Statistical Association*, 58, 933-942.

Bajari, P., Han, H., Krainer, J., & Nekipelov, D. (2010). Estimating Static Models of Strategic Interactions. *Journal of Business & Economic Statistics*, *28*(4), 469-482.

Beeson, P. E., & Eberts, R. W. (1989). Identifying Productivity and Amenity Effects in Interurban Wage Differentials, *Review of Economics and Statistics*, 71, 443-452.

Berry, S. (1992), Estimation of a Model of Entry in the Airline Industry. *Econometrica*, *60*(4), 889–917.

Blomquist, G. C., & Berger, M. C. (1992). Mobility and Destinations in Migration Decisions: The 20 Roles of Earnings, Quality of Life, and House Prices. *Journal of Housing Economics*, 2, 37-59.

Bresnahan, T. F., & Reiss, P. C. (1990). Entry in Monopoly Markets. *Review of Economic Studies, 57*, 531-553.

Bresnahan, T. F., & Reiss, P. C. (1991). Entry and Competition in Concentrated Markets. *The Journal of Political Economy*, 99(5), 977-1009.

Brueckner, J., & Luo, D. (2013). Measuring Strategic Firm Interaction in Product-Quality Choices: The Case of Airline Flight Frequency. Work-in-Progress, Department of Economics, University of California, Irvine

California Public Utility Commission (CPUC). (2013). *Data Processing Methods*. October 1st. Retrieved from http://www.cpuc.ca.gov/NR/rdonlyres/73AE2E37-3E8C-45E8-8B13-97C3B8434D86/0/CA_Methodology_2013_10_01.pdf

Case, B., & Quigley, J. M. (1991). The Dynamics of Real Estate Prices. *Review of Economics and Statistics*, 22(1), 50-58.

Case, K. E., & Shiller, R. J. (1989). The Efficiency of the Market for Single Family Homes. *American Economic Review*, 79(1), 125-37.

Chamberlin, E. (1933). *The Theory of Monopolistic Competition*. Harvard University Press.

Chen, Y., & Gayle, P. (2013). Mergers and Product Quality: Evidence from the Airline Industry. MPRA Paper No. 51238, Retrieved from ideas.repec.org/p/pra/mprapa/51238.html

Chen, Y., & Schwartz, M. (2013). Product Innovation Incentives: Monopoly vs. Competition. *Journal of Economics and Management Strategy*, *22*(3), 513-528.

Ciliberto F., & Tamer, E. (2009). Market Structure and Multiple Equilibria. *Econometrica*, *77*(6), 1791-1828.

Craig, J., & Savage, S. (2013). Market Restructuring and the Efficiency of Electricity Generation: Plant Level Evidence From the United States 1995 to 2006. *The Energy Journal*, 34(1), 1-31.

Crawford, G. (2007). Television Station Ownership Structure and the Quantity and Quality of TV Programming. FCC Media Ownership Study #3, Washington, D.C., FCC. Retrieved from https://apps.fcc.gov/edocs_public/attachmatch/DA-07-3470A4.pdf

DataQuick. (2013). Assessor Data for Specific Jurisdictions. DataQuick, Inc., San Diego, CA.

Denni, M., & Gruber, H. (2007). The diffusion of broadband telecommunications in the U.S.—the role of different forms of competition. *Communications & Strategies*, *68*, 139-157.

Dick, A. (2007). Market Size, Service Quality, and Competition in Banking. *Journal of Money, Credit and Banking*, *39*(1), 49-81.

Domberger, S., & Sherr, A. (1989). The Impact of Competition on Pricing and Quality of Legal Services. *International Review of Law and Economics*, *9*, 41-56.

Dorfman, R., & Steiner, P. (1954). Optimal Advertising and Optimal Quality. *American Economic Review*, *44*, 826-836.

Douglas, G., & Miller, J. (1974). Quality Competition, Industry Equilibrium, and Efficiency in the Price-Constrained Airline Market. *American Economic Review*, *64*, 657-669.

FBI. (2011). Crime in the United States 2011. Washington,DC. Retrieved May 31, 2013, from www.fbi.gov/about-us/cjis/ucr/crime-in-the-u.s/2011/crime-in-the-u.s.-2011/

FCC. (2011a). *Internet Access Services: Status as of June 30, 2010.* FCC, Washington, D.C.

FCC. (2011b). *Measuring Broadband America – August 2011.* FCC, Washington, D.C.

FCC. (2012). *Measuring Broadband America – July 2012.* FCC, Washington, D.C.

FCC. (2013). *Measuring Broadband America – February 2013.* FCC, Washington, D.C.

Flamm, K. (2005). The Role of Economics, Demographics, and State Policy in Broadband Availability. Paper presented at the PURC/London Business School Conference on "The Future of Broadband: Wired and Wireless, 2005", Gainesville, Florida.

Flamm, K., & Chaudhuri, A. (2007). An analysis of the determinants of broadband access. *Telecommunications Policy*, *31*, 312–326.

Gaynor, M. (2006). What do We Know About Competition and Quality in Health Care Markets? NBER Working Paper 12301, Cambridge, MA.

GeoLytics, Inc. (2012). *Census 2010 and ACS 2008-2012* [Custom Data Files]. Available from www.geolytics.com.

Gillett, S., & Lehr, W. (1999). Availability of Broadband Internet Access: Empirical Evidence. *27th Annual Telecommunications Policy Research Conference*, Washington, D.C.

Goolsbee, A., & Petrin, A. (2004). The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable TV. *Econometrica*, *72*, 351-81.

Greenstein, S., & Mazzeo, M. (2006). The Role of Differentiation Strategy in Local Telecommunication Entry and Market Evolution: 1999-2002. *The Journal of Industrial Economics*, *54*(3), 323-350.

Greenstein, S., and McDevitt, R. (2011). “The Broadband Bonus: Estimating Broadband Internet's Economic Value.” *Telecommunications Policy*, 35(7), 617-632.

Greenstein, S., and Prince, J. (2013). Does Service Bundling Reduce Churn? Available at http://ssrn.com/abstract=1966221 or http://dx.doi.org/10.2139/ssrn.1966221.

Grubesic, T. H., & Murray, A. T. (2004). Waiting for broadband: Local competition and the spatial distribution of advanced telecommunication services in the United States. *Growth and Change*, *35*(2), 139-165.

Halvorsen, R., & Palmquist R. (1980). The Interpretation of Dummy Variables in Semilogarithmic Equations, *The American Economic Review*, *73*(3), 474-475.

Haurin, D. R. (1980). The Regional Distribution of Population, Migration, and Climate, *Quarterly Journal of Economics*, XCV, 293-308.

Haurin, D. R., & Brasington, D. (1996). The impact of school quality on real house prices: Interjurisdictional effects. *Journal of Housing Economics*, 5(4), 351-368.

Heckman, J. (1979). Sample Selection Bias as a Specification Error. *Econometrica*, 47, 153-161.

Hill, R. C., Sirmans, C.F., and Knight, J. R. (1999). A Random Walk Down Main Street? *Regional Science and Urban Economics*, 29(1), 89-103.

Hiller, S., Savage, S., & Waldman, D. (2014). Market Structure and Media Diversity. forthcoming in *Economic Inquiry*.

Holian, M. J., & Kahn, M. E. (2012). The Impact of Center City Economic and Cultural Vibrancy on Greenhouse Gas Emissions from Transportation. San Jose, CA: MTI Publications.

Katz, M. (2013). Provider Competition and Healthcare Quality: More Bang for the Buck? *International Journal of Industrial Organization*, *31*, 612-625L

Knight, J. R., Dombrow, J., & Sirmans, C.F. (1995). A Varying Parameters Approach to Constructing House Price Indexes. *Real Estate Economics*, 23(2), 187-206.

Kranton, R. (2003). Competition and the Incentive to Produce High Quality. *Economica*, *70*, 385-404.

Kugler, A., & Weiss, C. (2013). Time as a Strategic Variable: An Empirical Analysis of Opening Hours in the Austrian Gasoline Market. Work-in-Progress, Vienna University of Business and Economics, Vienna.

Lardinois, F. 2012. Google Fiber: 20% Of Kansas City, MO Neighborhoods Have Already Met Their Sign-Up Goals. Techcrunch, [online] 29 July, Accessed on 4-11-14 at http://techcrunch.com/2012/07/29/kc-google-fiber-threshold-20-percent/

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy*, 74, 132-157.

Lee, S. [Sangwon] , Marcu, M., & Lee, S. [Seonmi]. (2011). An empirical analysis of fixed and mobile broadband diffusion. *Information Economics and Policy*, *23*, 227-233.

Li, G. (1985). Robust regression. In Exploring Data Tables, Trends, and Shapes, ed. D. C. Hoaglin, C. F. Mosteller, & J. W. Tukey, 281–340. New York: Wiley

Leichtman Research Group. (2011). *Nearly 1.3 Million Add Broadband in the First Quarter of 2011.* Accessed on 5-1-13 at http://www.leichtmanresearch.com/press/051711release.html

Leichtman Research Group. (2014). *Cable, DBS & Telcos: Competing for Customers 2014*, Leichtman Research Group, Durham, NH.

Macher, J., Mayo, J., Ukhaneva, O., & Woroch, G. (2012). "Demand in a Portfolio-Choice Environment: The Evolution of Telecommunications." *Georgetown McDonough School of Business Research Paper No. 2012-19.*

Maclennan, D. (1977). Some Thoughts on the Nature and Purpose of Hedonic Price Functions. *Urban Studies*, 14, 59-71.

Malpezzi, S. (2002). Hedonic Pricing Models and House Price Indexes: A Select Review. In Kenneth Gibb and Anthony O'Sullivan (eds.), Housing Economics and Public Policy: Essays in Honour of Duncan Maclennan. Oxford: Blackwell Publishing, 67-89.

Manuszak, M., & Moul, C. (2008). Prices and Endogenous Market Structure in Office Supply Superstores. *The Journal of Industrial Economics*, 56, 94–112.

Manning, C. A. (1989). Explaining Intercity Home Price Differences, *Journal of Real Estate Finance and Economics*, 2, 131-149.

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

Mazzeo, M. (2002). Product Choice and Oligopoly Market Structure. *Rand Journal of Economics 33*(2), 221-242.

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

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

Mills, E. S. (1967). An Aggregative Model of Resource Allocation in a Metropolitan Area, *Amer. Econ. Rev. Papers and Proceedings*, 57, 197-211.

Miller, D. A. & White, R. A.(1998). A Conterminous United States Multi-Layer Soil Characteristics Data Set for Regional Climate and Hydrology Modeling. Retrieved May 15, 2014, from http://www.soilinfo.psu.edu/

Nardotto, M., Valletti, T., & Verboven, F. (2012). Unbundling the Incumbent: Evidence From UK Broadband. CEPR Discussion Papers 9194.

NBM. (2011). NTIA & FCC National Broadband Map. June 2011, Retrieved December 30, 2012, from http://www2.ntia.doc.gov/Jun-2011-datasets.

Nevo, A., Turner, J., & Williams, J. (2013). Usage-Based Pricing and Demand for Residential Broadband. Presented at the NBER Economics of IT and Digitization Workshop, July. Retrieved from: http://dx.doi.org/10.2139/ssrn.2330426

NTIA-ESA (2013). Broadband Availability Beyond the Rural/Urban Divide, Retrieved June 9, 2013, from www.ntia.doc.gov/report/2013/broadband-availability-beyond-ruralurban-divide.

Olivares, M., & Cachon, G. (2009). Competing Retailers and Inventory: An Empirical Investigation of General Motors' Dealerships in Isolated U.S. Markets. *Management Science*, *55*(9), 1586–1604.

Ozanne, L., & Thibodeau, T. G. (1983). Explaining Metropolitan Housing Price Differences, *J. Of Urban Economics*, 13, 51-66.

Peek, J., & Wilcox, J. A. (1991). The Measurement and Determinants of Single Family House Prices, *AREUEA J.*,19, 353-382.

Petrin, A., & Train, K. (2010). A Control Function Approach to Endogeneity in Consumer Choice Models, *Journal of Marketing Research,* 47(1), 3-13.

Pew Research Center (2010). Spring Change Assessment Survey 2010. Accessed on 2-14-13 at http://www.pewinternet.org/~/media//Files/Questionnaire/2010/Spring%202010%20Topline%20-%20Broadband.pdf.

Potepan, M. J. (1994). Intermetropolitan Migration and House Prices: Simultaneously Determined? *Journal of Housing Economics*, 3, 77-91.

Prieger, J. E. (2003). The supply side of the Digital Divide: Is there equal availability in the broadband Internet access market? *Economic Inquiry*, *41*(2), 346-363.

Prieger, J. E. (2007). The Impact of Cost Changes on Industry Entry and Exit, *Journal of Economics*, *91*(3), 211–243.

Prieger, J. E. (2012). The Broadband Digital Divide and the Economic Benefits of Mobile Broadband for Rural Areas. *Telecommunications Policy*, *37*(6-7), 483-502.

Prieger, J. E. & Church, T. (2012). Deployment of Mobile Broadband Service in the United States. In *Mobile Services Industries, Technologies, and Applications in the Global Economy*, I. Lee (Ed.), Hershey, PA: IGI Global.

Prieger, J. E. & Hu, W.-M. (2008a). Competition in broadband provision and the digital divide. In Y.K. Dwivedi, et al. (Eds.), *Handbook of Research on Global Diffusion of Broadband Data Transmission*, (pp. 241-259). Hershey, PA: IGI Global.

Prieger, J. E., & Hu, W.-M. (2008b). The broadband digital divide and the nexus of race, competition, and quality. *Information Economics and Policy*. *20*(2), 150-167.

Prieger, J. E. & Lee, S. (2008). Regulation and the Deployment of Broadband. In Y.K. Dwivedi, et al. (Eds.), *Handbook of Research on Global Diffusion of Broadband Data Transmission*, (pp. 278-303). Hershey, PA: IGI Global.

Prince, J. E., & Simon, D. (2013). Do Incumbents Improve Service Quality in Response to Entry? Evidence from Airlines' On-Time Performance. *Indiana University School of Public and Environmental Affairs Research Paper No. 2010-09-01*. Retrieved from http://ssrn.com/abstract=1670203 or http://dx.doi.org/10.2139/ssrn.1670203

Rosston, G., (2009). "The Evolution of Third-Part High-Speed Internet Access." *Information Economics and Policy,* 21(1), 21-33.

SamKnows. (2012). SamKnows Test Methodology, Methodology and Technical Information Relating to the SamKnows Testing Platform. Accessed on 10-10-2012 at www.samknows.com.

Savage, S., & Wirth, M. (2005). Price, Programming, and Potential Competition in U.S. Cable Television Markets. *Journal of Regulatory Economics*, *27*, 25-46.

Schmalensee, R. (1977). Comparative Static Properties of Regulated Airline Oligopolies. *Bell Journal of Economics*, *8*, 565-576.

Schmalensee, R. (1979). Market Structure, Durability, and Quality: A Selective Survey. *Economic Inquiry*, *17*, 177-196.

Seim, K. (2005). An Empirical Model of Firm Entry with Endogenous Product-Type Choices. *Rand Journal of Economics 37*(3), 619-640.

Shin, R., & Ying, J. (1992). Unnatural Monopolies in Local Telephone. *Rand Journal of Economics*, 23, 171-183.

Stenberg, P., Morehart, M., Vogel, S., Cromartie, J., Breneman, V., & Brown, D. (2009). Broadband Internet's value for rural America. *U.S. Department of Agriculture, Economic Research Service, Economic Research Report No. 78.*

Sundaresan, S., Donato, de W., Feamster, N., Teixeira, R., and Pescapè, A. (2011). Broadband Internet Performance: A View From the Gateway. In *Proceedings of the ACM SIGCOMM 2011 Conference*, Toronto, Canada.

Swan, P. (1970). Durability of Consumer Goods. *American Economic Review*, *60*, 884-894.

Swan, P. (1971). The Durability of Goods and the Regulation of Monopoly. *Bell Journal of Economics and Management Science*, *2*, 347-357.

Tamer, E. (2003). Incomplete simultaneous discrete response model with multiple equilibria. *Review of Economic Studies 70*(1), 147–167.

Tax Foundation (2013). Property Tax Data by County, Retrieved July 17, 2013, from interactive.taxfoundation.org/propertytax/.

Telogical Systems (2012). *Pricing Trends Data*. Telogical Systems, McLean, VA.

US Census Bureau. (2010). *US Census 2010*. Retrieved May 31, 2013, from www.factfinder2.census.gov.

US Census Bureau. (2012). *American FactFinder*. Available at factfinder2.census.gov.

US Census Bureau. (2013a). American Community Survey 2006-2010. *Social Explorer*. Retrieved May 31, 2013, from www.socialexplorer.com.

US Census Bureau. (2013b). County Business Patterns 2011. *American FactFinder*. Retrieved June 12, 2013, from factfinder2.census.gov.

US Census Bureau. (2013c). *TIGER/Line Shapefiles*. Retrieved May 31, 2013, from www.socialexplorer.com.

US Department of Agriculture. (2013). *Web Soil Survey*. Natural Resources Conservation Service. Retrieved May 15, 2014, from websoilsurvey.nrcs.usda.gov.

Vuong, A. (2014). CenturyLink to turn Denver into a Gig City with 1 gigabit per second. *Denver Post*, [online] 5 Aug, Accessed on 10-11-14 at http://www.denverpost.com/business/ci_26275048/centurylink-turn-denver-into-gig-city-1-gigabit

Wallsten, S., & 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.

Xiao, M., & Orazem, P. (2011). Does the Fourth Entrant Make Any Difference? Entry and Competition in the Early U.S. Broadband Market. *International Journal of Industrial Organization*, *29*(5), 547-561.

Zager, M. (2011). Modeling the Cost Of Rural Fiber Deployment. *Broadband Communities Magazine*, [online] March, Accessed on 10-10-14 at http://bbpmag.com/2011mags/marchapril11/BBP_MarApr_CostOfFiber.pdf