Deconstructing the Digital Divide: Identifying the Supply and Demand Factors That Drive Internet Subscription Rates[1]

Michael J. R. Martin[2]
Education and Social Stratification Branch
United States Census Bureau

SEHSD Working Paper Number: 2019-15
DA Number: CBDRB-FY19-ROSS-B010

Abstract: Unequal access to the internet, better known as the digital divide, can be broken down into issues with internet supply (or the availability of broadband connections) and demand (or the rate at which those connections are utilized). This paper examines newly available data sources to highlight the role that these supply and demand factors play in driving gaps in internet subscription rate. Supply factors are captured by the FCC's Form 477 filing, which reports the availability of broadband connections. Demand factors are captured by a number of demographic and structural factors at the county level. These measures are used in models to predict the rate of internet subscription within counties, using newly available 5-year ACS data. Overall, I find that factors of internet supply, demographic and structural demand all play a role in determining county level subscription rates. I close by outlining the potential for future research and potential implications for approaches to narrow the digital divide.

[1] This paper is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The views expressed on statistical and other issues are those of the authors and not necessarily those of the U.S. Census Bureau.
[2] For questions or comments about this work or other questions related to American Community Survey data on computer and internet use, Michael can be reached at michael.j.martin@census.gov.

Introduction

In 2015, President Obama announced that his administration would focus on ensuring that all Americans would have access to high speed broadband internet (Obama 2015). Citing a lack of available broadband internet in rural communities, the president called for increasing infrastructure in rural areas. In 2017, the Trump administration announced a new program aimed at building rural broadband infrastructure (U.S. Department of Agriculture 2017). High speed internet access is increasingly required to access numerous aspects of everyday life, including medical information, employment opportunities, and educational resources. The policies proposed by both Presidents Obama and Trump are designed to help close the digital divide, or the phenomenon whereby particular areas or groups of people have lower rates of internet access. Expanding the national broadband infrastructure would address one of the major facets of the digital divide; namely, that infrastructural shortcomings can prevent certain communities from having affordable or even available internet. But to what extent does internet availability actually impact internet uptake, or the actual utilization of available internet services? And, in places where internet uptake lags, what other factors contribute to the digital divide?

Since 2013, the American Community Survey has asked respondents questions about internet subscription status. With the December release of the 5-year ACS file, county level estimates of internet subscription rates are now publicly available for each county in the U.S. for the first time. In addition to exploring sub-state patterns of internet uptake[3], pairing these survey data with administrative data on internet availability collected by the FCC allows for the first large scale, nationally representative analysis of the determinants of internet uptake at the community level. Incorporating measures of both availability and community contexts provide

[3] Within this paper, I use broadband subscription and broadband uptake interchangeably. These two terms represent the same concept: the degree to which households report connections to an available broadband service.

greater understanding of the determinants of the digital divide, and allows for differentiation of communities that are disconnected due to issues with internet availability from those that are less connected due to economic or social factors.[4]

Background

Previous Research

Early evidence for the digital divide was derived from questions in the Current Population Survey (CPS), which first asked about internet use in 1997. Using these early data, researchers identified gaps between connection rates for wealthy and poor households, particularly for households in rural or central city locations (NTIA 1998). Initially, researchers hypothesized that these shortcomings in internet use were due to "material factors," such as a lack of a computer or available internet connection at a household (van Dijk & Hacker 2003; DiMaggio and Hargittai 2001). Clearly, material factors are important, as households have traditionally required a computer and household internet connection to reliably access the internet.[5] While the wide availability of mobile broadband may also be providing an alternative for some households to get connected (Ryan 2018), an available fixed internet connection remains the necessary precursor to an internet subscription for the majority of homes.

Addressing the material factors of the digital divide has been the primary focus of government policy, with particular attention paid to the challenges and costs of connecting less dense rural counties (FCC 2010a; FCC 2010b; Obama 2015). The 2010 National Broadband Plan outlined a number of major approaches for the government to take in order to improve the national broadband environment. Broadly, these approaches focus on improving the deployment

[4] For an exploration of how internet availability is itself related to the social and economic characteristics of a community, see Martin and Lewis, 2018.
[5] Some alternatives outside the home do exist, such as internet available in public libraries, though these points of access are outside the scope of this paper.

of broadband infrastructure to reach "high cost" areas, such as rural counties where the cost to connect customers is greater than in denser urban counties (FCC 2010c, FCC 2010b). The plan also calls for increased competition between broadband providers nationwide to "maximize consumer welfare, innovation, and investment" (FCC 2010c). Government agencies are also instructed to facilitate "network upgrades and competitive entry" for internet service providers. This involves opening up the broadband spectrum to new providers, helping to finance new infrastructure developments, and expanding current broadband infrastructure (FCC2010c). These approaches address the basic necessity for internet uptake, as broadband subscriptions first require available broadband internet services.

Alongside infrastructural improvements, the National Broadband Plan also pays attention to non-infrastructural impediments to broadband uptake, specifically the costs of broadband internet. In an FCC survey of broadband uptake, the most common reason non-users cited for a lack of a home broadband connection was the potential high costs of the service, with 36% of non-users reporting this as their main barrier (Horrigan 2010). In order to combat high costs, the National Broadband Plan highlights two approaches. First, broadband markets require competition, especially in areas where broadband deployment requires relatively high levels of financial investment on the part of internet service providers (FCC 2010c). Second, the government needs to sometimes intervene to ensure that lower-income Americans can afford higher speed broadband services at their home, a step that involves not only expanding subsidies for lower-income households, but also expanding the Lifeline program which helps provide discounted mobile phones to qualifying households (FCC2010a).

Cost concerns and infrastructural barriers are not the only barriers to high speed internet connection. The same FCC survey that showed cost as a prohibitive factor also found that roughly 22% of individuals who did not use broadband did so for reasons related to digital literacy, such as a lack of familiarity with broadband internet or privacy concerns (Horrigan 2010). Individuals lacking the skills to set up and maintain a broadband connection are less likely to have a broadband connection in their home (van Dijk & Hacker 2003). Lower levels of education may also contribute to this lower levels of digital literacy, as individuals with less than a bachelor's degree report comparatively lower rates of broadband subscriptions than more highly educated individuals (Horrigan 2010). The National Broadband Plan recognized that a lack of skills in accessing the internet may diminish uptake even when affordable connections are available, and outlined a National Digital Literacy Corps to increase familiarity with digital resources (FCC 2010c). Should issues with technical skills or digital literacy represent a substantial hurdle for broadband subscription, increasing familiarity with how to set up, access, and use the internet in a safe and secure manner could help overcome these issues.

Another group of respondents to the same FCC survey indicated that the internet was not relevant to their daily life. A full 19% of FCC survey respondents felt that the internet was unnecessary or a "waste of time" (Horrigan 2010). These individuals may have the skills to connect and may do so outside the home, but for a number of reasons, they choose not to establish a household connection. Some of these individuals may have been exposed to the internet for the first time later in life and having lived a disconnected life for many years, may not see the benefit of connecting from home (van Dijk & Hacker 2003). This is particularly true for elderly Americans, a broad population who has lived without internet availability for most of their lives, and other groups who, for a variety of reasons, have not had household internet access for many years. For example, individuals raised in unconnected, lower income families may learn to get by only with internet access external to their homes, such as at school or a public library, and these individuals may not view home connections as necessities. While concerns with infrastructure, cost, and skill gaps can be addressed with specific policies, these individuals may be harder to directly impact through government programs, and may remain disconnected unless and until their views on the relative value of broadband connections change.

Even as broadband connections are rolled out to more communities, value and skill based hurdles may continue to prevent some households from fully connecting (van Dijk & Hacker 2003; DiMaggio & Hargittai 2001; DiMaggio et al 2004). Recent evidence points to a developing secondary digital divide, as variation exists in how the internet is utilized by particular groups (DiMaggio & Hargittai 2001; DiMaggio et al 2004; Mossberger et al 2012; Mossberger et al 2013). Though household penetration of internet connectivity has increased over the past few decades, improvements in broadband technology have created inequalities in the speed of access for many communities. With this in mind, the FCC recently updated its benchmark for broadband connection goals to require higher download and upload speeds (FCC 2015). Previously, the threshold for high speed internet connections that should be available to the majority of Americans was set at 4 Mbps download and 1 Mbps upload. Currently, this threshold has been increased to require 25 Mbps download and 3 Mbps upload, a drastic increase in speeds reflecting increasing usage of the internet for high-quality streamed media and telecommunications. Thus, while the overall number of connected homes continues to increase, disparities in terms of household connection speed may hinder disadvantaged households from making full use of their internet access.

Toward a Model of Broadband Uptake

Overall, a number of supply and demand factors operate to drive broadband uptake in American communities. At the most basic level, the level of broadband availability sets a baseline for whether or not a household can decide to subscribe to the internet. Communities with fewer homes that *can* be connected will invariably have lower shares of homes that are actually connected, particularly in comparison to communities with wider broadband availability. In addition to whether or not a connection is available at all, competition between service providers may result in greater uptake, as these providers may offer services at different price points or with different technologies. Given that the most common reason for not accessing the internet from home is cost, competition that drives cost down and increases consumer options may provide higher community connectivity (Horrigan 2010; FCC 2010c). The speed of the internet services available is likely a factor as well, as consumers might be unwilling to pay for broadband connections that do not allow for full utilization of the internet. As higher speeds become available, service providers may also offer slower tiers of service for less money.

Demand for broadband services comes both from the structural and demographic aspects of a community. Structural demands come largely from the potential return on investment an internet service provider can expect for providing broadband services to a county. For example, the cost-per-customer to establish broadband connections in rural communities is higher than in urban ones (FCC 2010b) and, as such, rural communities tend to have lower rates of broadband available (Martin & Lewis 2018). Counties with large shares of renters also tend to have higher rates of internet availability (Martin & Lewis 2018), an outcome likely due to the ease with which an internet service provider can tap a large number of potential clients by providing a connection to a single apartment building with multiple housing units. Similar factors may operate in counties with newer housing stock and those that are more densely populated. In addition to the potential for returns on investment for infrastructural deployment in these types of counties, service providers may see higher returns by advertising in high-value communities, instead of less densely populated rural areas.

Demographic demand is also likely to play a key role in driving broadband uptake. Considering that internet costs are one of the most commonly cited reasons why households eschew a broadband subscription, in general, counties with higher median incomes should have higher rates of broadband uptake than less affluent areas. Communities with higher “digital literacy,” such as those with higher rates of educational attainment, may have a higher demand for broadband internet services. Older communities, on the other hand, may have reduced demand for internet services as elderly individuals might lack the technical skills and motivation to pursue a household connection (Horrigan 2010; van Dijk & Hacker 2003; DiMaggio & Hargittai 2001; DiMaggio et al 2004). Counties with higher rates of black and Hispanic residents, as well as those with a greater share of households with at least one disabled member, represent additional areas where demographic demand maybe low. Similar to elderly populations, these other groups have been shown to have lower rates of access, for reasons attributable to both technical skill and value based factors (Horrigan 2010; van Dijk & Hacker 2003; DiMaggio & Hargittai 2001; DiMaggio et al 2004). Taken together, community demographics can impact the rates at which specific counties connect to the internet from home.

Current Study

In this study, I provide a clearer picture of the factors contributing to inequality in internet subscription rates resulting in the persistence of the digital divide. Recent U.S. policy has focused heavily on how supply factors influence subscription rates, attempting to both increase the supply in underserved areas, while increasing competition between internet service providers (ISPs) to lower costs for the average consumer. The analysis proposed herein will test how well these supply factors operate to reduce disparities in internet subscription.

In addition to factors related to broadband supply, I also incorporate structural demand factors, such as the share of households classified as urban and the share of residents that rent their home, and demographic demand factors, such as the median level of income within a county, in an attempt to explain internet usage gaps that persist even in areas with a relatively high supply of internet. In doing so, I leverage new data sources, foremost among them the 5-year ACS data release, to model county level subscription rates controlling for the availability and speed of internet services. Specifically, this paper addresses the following research questions:

- What are the patterns of broadband availability and subscription at the county level?
- To what extent do infrastructural or supply factors, such as internet availability and level of competition, determine the rate of internet subscription across counties? Is this relationship impacted by the level of speed available to the average resident and the number or different services from which a resident may choose?
- What demographic, economic, and geographic factors predict higher rates of internet subscription as they relate to demand for services?

Methodology

Data

The five year release of American Community Survey data provides the backbone for this research. In this survey, respondents are asked to provide the types of internet service, if any, for which they subscribe in their household (Figure 1).[6] In this analysis, I focus mainly on the "high speed" broadband category comprised of cable, DSL, and fiber-optic connections. This category is more likely than the alternatives such as satellite or terrestrial fixed wireless services to meet the FCC's benchmark for a high-speed broadband connection necessary for ordinary internet use. Though I cannot determine household internet speed from the ACS data, these technologies have the capacity to deliver higher speeds of service. As such, I refer to these connections as "high capacity" connections throughout this paper. As I am focused on the contextual determinants of subscription, I operationalize this variable as the share of persons in a county that reports connecting to the internet using high capacity broadband services.

The ACS also provides for measurement of the social and economic factors that may function in determining or impacting connection rates. Median household income, level of education, and age all may directly affect connection rates as older, poorer and less well-educated households frequently eschew connections for a number of reasons discussed above. Past research has also demonstrated that certain racial and ethnic minorities, as well as individuals with disabilities tend to have lower levels of connectivity (Carlson and Goss 2016; CEA 2015; GAO 2015; Ryan and Lewis 2017).

Data on internet availability are derived from the FCC's Form 477 data. Compliance with the form requires that ISPs identify each type of internet service they make available by technology and speed for each census block. An example of the filing for a single block is provided in Table 2.[7] Eight different services are offered in this block for each of five different transmission technologies. Though not present in this example, if a company provides two

[6] Prior to 2016, respondents were provided a slightly different set of response categories, with cable, fiber optic and DSL all being provided as individual response options. For a detailed account of the changes made, see Ryan, 2018.
[7] In the full filing, the FCC requires ISPs to report both commercial and residential internet services. For this analysis and in this example table, I reduce the services to be analyzed to just residential offerings.

services of the same transmission type at different speeds, these are reported separately. For example, if AT&T also offered a higher speed ADSL2 service at 25 Mbps down and 3 Mbps up, it would be reported as a unique entry.

The FCC provides guidelines determining availability, indicating that a service must be reported if "the provider does, or could, within a service interval that is typical for that type of connection--that is, without an extraordinary commitment of resources--provision two-way data transmission to and from the internet with *advertised speeds* exceeding 200 kbps in at least one direction to *end-user premises* in the census block." (FCC Form 477 Instructions 2016, p 17; emphasis theirs). These data include all wired connection types, such as cable, DSL, and fiber-optic internet, as well as satellite internet and terrestrial fixed wireless connections; the FCC collects data on mobile broadband separately and those connections are not covered in this paper.[8] I rely on the December 2015 version of the Form 477 data in order to best match these data to the midpoint of the 5 year ACS data file which covers from 2013 to 2017.[9]

These publically-available data have complete coverage of census blocks with each block at least having satellite coverage,[10] and the average block having over 5 different services offered.[11] Form 477 data contain detailed technology codes and download speeds, providing a thorough depiction of internet availability. For this study, I chose to focus only on whether or not individuals have access to a cable, DSL or fiber-optic service, as this matches the high

[8] Mobile broadband coverage is provided with slightly less detail and as coverage area estimates in the form of a shapefile rather than linked specifically to an underlying census geography.
[9] American Community Survey data are subject to sampling error. These data represent estimates over the 2013-2017 time period, and do not reflect the estimate at any specific point within the time period (including the midpoint). For more information on sampling and non-sampling error, confidentiality protection, and definitions in the ACS, visit https://www.census.gov/programs-surveys/acs/technical-documentation/code-lists.html.
[10] ISPs are able to report satellite coverage for an entire state if they believe that they could, within reason, make their service available anywhere within the state.
[11] This number is inflated partially due to the broad availability of satellite internet. The three major satellite providers (Hughes net, ViaSat and VSAT Systems) each offer at least one satellite service in every census block in the US.

capacity category available in ACS data. To accomplish this, I condense these data to a block-level file indicating the presence or absence of internet service from these specific technologies. I also consider the impact internet speed may have on the likelihood of subscription. To accomplish this, I identify the average maximum advertised download speed for each block. Lastly, I obtain a measure of competition by identifying the total number of high capacity services for each block.

In order to merge these data with county level ACS data, I build county level profiles, discussed in more detail below. To do so, I also rely on 2010 block level Census population counts. While these population estimates are now somewhat dated, they provide the closest reliable estimate of population that I can merge with the block level Form 477 data. I explain my process for using these figures to calculate availability at the county level in the next section. However, it is beneficial to note that all estimates using these population counts and the 477 data are just our best estimate for what availability actually looks like in a given county. Population counts at the block level have certainly changed between 2010 and 2013-2017, the years from which we source our ACS data, and further broadband deployment has likely occurred following the filing of Form 477 in 2015. The estimates generated from these figures should be treated as just that, with readers acknowledging that the true values are likely different but impossible to estimate more accurately without additional data sources.

Method of Analysis

While ACS data is freely available at the county level, more work must be done to prepare the Form 477 data. After merging these block level data with 2010 population counts at the block level, I use a two factor approach to obtain a county level estimate. First, for the share of a county that has access to high capacity broadband services, I create a binary indicator of whether or not a block has at least one high capacity service available, and then obtain the total population living in blocks that have these high capacity services available within a county. From there, I estimate the share of the total county population that should have access to these services. This figure represents the population in a county that likely has high capacity services available, at least somewhere within their block. Next, for the average number of high capacity broadband services and the average advertised maximum download speed, I obtain an overall average value for the county, weighted by the population in each block. Thus, these numbers can be interpreted as the potential broadband market place to which the average resident is exposed when they shop for broadband services.

My analysis unfolds in two stages. The preliminary descriptive analysis estimates the overall rate of subscription and internet availability across the United States, while also paying attention to the variation present across counties. After establishing these county level patterns, I turn to a set of OLS regression models that estimate the share of a county reporting subscription to a high capacity broadband service, conditional on availability, average speed, and a number of other contextual social and economic factors. I also make use of standardized coefficients to compare the relative impact each of these variables has on broadband uptake.

Results

Internet Availability

High capacity broadband availability is, all things considered, fairly high. I present the descriptive statistics for measures derived from Form 477 data in Table 1. [12] In the average county,[13] 86.7% of households had some form of DSL, cable, or fiber-optic internet service

[12] These estimates are drawn from combining administrative 477 service availability data with 2010 Census complete count population data. As a result, no sampling error is introduced into the analysis. However, as noted previously, the five-year gap between the Form 477 and 2010 Census data may pose limitations for analysis.
[13] This refers to the mean county level estimate.

available for purchase, and in roughly three quarter of counties at least 81.1% of households had access to a high capacity broadband service. While most people had access to some sort of high capacity broadband service, the average household had only 1.9 different high-capacity broadband services from which to choose.

Speed of connection for most counties tends to be high as well. In the average county, the average maximum advertised download speed is 58.2 Mbps. This rate is more than double the FCC's "high speed" benchmark of 25 Mbps, though it is important to reiterate that this is just the maximum *advertised* speed. Internet service providers may advertise one maximum speed but due to technical limitations, actually deliver speeds considerably lower than the theoretical maximum. While the average download speed is high, residents in a number of counties have lower speeds available to them. In roughly 25% of counties, the average download speed is 20.6 Mbps or lower. Roughly 30% of counties fall below the FCC's "high speed" download benchmark.

In Figure 2, I provide a descriptive map of the availability of at least one high capacity broadband service. Counties with the darkest shade had the highest availability of a cable, DSL, or fiber optic connection, while those with the lightest shade had these technologies available for under 50% of the population. While some states are fairly uniform in their service availability (like Massachusetts and Connecticut, where every county falls in the top category), others contain service availability shares that run the full range of the scale (like Arizona and New Mexico).

The number of services available also varies county by county as shown in Figure 3. Unlike Figure 2, each state has multiple counties from different categories. Counties with an average of less than 1 high capacity service are found throughout the West, Midwest, and South, with one such county present in upstate New York. There are relatively few counties that fall in the highest category, with only 27 counties having an average number of services available higher than 4.

Figure 4 provides a map of the average speed of high capacity services by county. While counties with higher average speed are found in areas one might expect, such as the counties in the Los Angeles and Seattle metropolitan areas, they are also found in a number of states one might not expect. For example, a number of counties in North Dakota fall into the highest category. These figures may be dragged up by counties in which gigabit (1000 Mbps) services are available, particularly if there are only a few other lower speed services present. In a number of rural counties, gigabit fiber connections have become more prevalent, especially when driven by investments by a particular internet service provider. In North Dakota, for example, the Midco telecommunications company has rolled out gigabit fiber in a number of communities. Though these connections exist, cost considerations and other factors may impact broadband uptake even in the face of extremely fast speeds. Counties that fall below the FCC's designation of high speed tend to fall outside major urban areas, and can be found throughout much of the United States.

Internet Subscription

The geographic patterns present in the Form 477 availability data suggest that unequal access may remain an issue in determining broadband uptake. While some counties display overall broad patterns of availability, others have limited availability of high capacity services, either in the share of the population for which any of these services are available or in the average number of services offered. Figure 5 offers our first glimpse into patterns of household high capacity broadband subscriptions throughout counties. As with Figure 2, high subscription rates can be found throughout the country. In some states, like Massachusetts, Connecticut and Vermont, the majority of counties have over 70% subscription rates. In others, like Mississippi, the majority of counties fall below a 40% subscription threshold. In most states, however, counties fall into a number of different subscription categories.

Table 3 presents the descriptive statistics for high capacity broadband subscription rates, along with the independent variables to be used in the regression models that follow. In the average county, roughly 53.6% of households report having a subscription to a high capacity broadband service. In the most connected counties, subscription rates are no higher than 88.1%. Though the penetration of available services is relatively high as seen in Table 1, it is apparent that subscriptions lag somewhat behind available connections. This is an early indicator that supply is not sufficient to determine broadband uptake, a relationship I examine in more detail in the regression models below

Modeling Broadband Uptake

Table 4 presents a series of regression models exploring the relationship between broadband uptake and a number of supply and demand factors. These models are weighted by county populations to better represent the relationship for the average household. Model 1 presents a simple bivariate regression wherein the share of households with a high capacity broadband subscription is regressed on the share of households for which those services are available. As theorized, these two measures are closely linked. For every 1% increase in the share of households for which high capacity services are available, household subscription rates increase by roughly .988%. In this model, the negative value for the intercept indicates the lag evident in the descriptive analyses above, whereby broadband uptake trails availability by roughly 28 percentage points. While expected, this finding also supports the use of this approach of combining survey data from the American Community Survey and availability data from Form 477. Though they measure two different aspects of internet penetration in America, the fact that they present results that are closely correlated is a good sign that the underlying phenomena are closely linked.

In model 2, I introduce additional measures of supply. I capture ISP competition at the county level by using the average number of high capacity services available, as well as the average maximum advertised download speed.[14] These variables both operate as expected. More competition among services results in higher rates of uptake, with each additional service resulting in a 5.6% increase in broadband subscription. As consumers are afforded a higher degree of choice in pursuing broadband services, they may be more likely to find a service plan or level of cost that suits their needs, and are therefore more likely to subscribe. Speed also increases broadband uptake, with an increase of 10 Mbps resulting in a roughly .7% increase in the broadband subscription rate.

Model 3 introduces a number of structural demand factors. Earlier I theorized that the share of persons classified as living in an urban environment, the share or residents renting their home, the number of persons per square mile[15], and the share of households built after the year 2000 might all indicate communities with higher potential returns on investment for internet service providers. With the exception of the share of the population renting their homes, these variables are all linked to increases in broadband subscription rate as well. As I am already controlling for availability and competition, this finding may indicate a higher degree of non-

[14] For each of these variables, I rely on a top coded transformation of the variable, with valid codes encompassing 93% of counties or more. The average number of high capacity services is top coded at 3, as only 7% of counties fall beyond this mark, with 4 counties falling above 10. The average download speed is likewise capped at 150 Mbps, with 7% of counties falling above this mark.

[15] Estimate obtained through 2010 Decennial Census data

infrastructural investment in these areas. For example, service providers may be more heavily advertising in particularly dense counties either through media buys or low-tech options such as circulars in newspapers or fliers. Newer housing developments may also provide infrastructure within the home for broadband connections, thus reducing the barriers to establishing an in home subscription in the first place. In older homes, wiring for broadband may be more burdensome, while new homes may come pre-wired for connections.

Model 4 introduces demographic demand variables. Income (expressed as the log of the median household income) and the share of the population with a bachelor's degree or higher were both theorized as having a positive impact on broadband subscriptions rates, a finding that is confirmed by the model. Wealthier households and those with higher levels of education may view internet access as more of a necessity, and are thus likely to subscribe conditional on supply factors. As the median age in the county increases, the high capacity broadband subscription rate declines, though this relationship eventually plateaus at high ages. Elderly populations may be making a value or skill based decision as the prior literature suggests, and choosing to view a household subscription as unnecessary or unwanted. Similarly, the share of the county that is black or Hispanic is negatively associated with broadband subscription rate, as well as for the share of residents with a disability. Elderly, disabled, and racial minority populations may lack the technical skills required to establish home connections even in counties where these connections are widely available. As this is just a county level model, I cannot specifically determine which factors result in these populations being linked to lower broadband uptake. It is possible that a household level model could explain away some of this effect by looking at individual household income and education levels, both of which could correlate with race and disability status. Lastly, in this model the sign for share of renters in a county reverses direction, indicating that this relationship is moderated by on one or more demographic demand variables.

Relative Contribution of Supply and Demand Forces

Model 5 in table 4 presents the standardized coefficients for the final model of variables, which allows comparisons to be drawn in the relative contribution each variable makes on broadband uptake. The most influential variable is median income, with a one unit increase in standardized logged median income associated with an increase of .4 in standardized broadband subscription rate. Given that the most likely reason provided for a lack of subscription was cost (Horrigan 2010), this makes a great deal of sense; broadband subscriptions can be expensive, and they are more likely to be found in places where many people can reasonably afford them. Income represents the most sizeable demographic demand factor at play in determining broadband uptake, followed closely by age, race, and education, while share of residents with a disability makes a relatively small contribution.

The share of residents living in an urban area also makes a substantial contribution out of the structural demand factors. With a standardized coefficient about 0.3, a 1 unit increase in standardized percent urban represents a sizeable shift in the standardized percent with broadband subscriptions. The contributions of the share households built after 2000, the share of renters, and the number of persons per square mile all have standardized coefficients below .1.

Lastly, measures of supply also substantially contribute to broadband subscription rates. A one unit increase in the share of households with a high capacity service available is linked to a .231 increase in broadband subscription rate. Compared to overall supply, speed (standardized coefficient of .036) and competition (standardized beta of .029) are less impactful, and represent the smallest standardized betas out of all variables included in the model.

Testing with an alternative measure of availability

Given the highly skewed nature of one of my main explanatory variables, I tested my models with an alternative measure of broadband availability[16]. Rather than looking at the share of the population that has at least one high capacity broadband service available, I calculated the share of the population with 2 or more services available. In the average county, only 62.4% of households have 2 or more broadband services available, compared to 86.7% of households that have access to one or more services. Fewer counties also exhibit full or near full coverage. In models containing this alternative measure of availability, all relationships remain significant and coefficients are in the same direction as in my primary models. Predicted values obtained from the model show a high degree of correlation with each other, indicating that the results are comparable using either measure of availability (r = .95 for both Pearson's *r* and Spearman's rank ordered correlation).

Discussion and Conclusion

Despite continued investment by policymakers in increasing broadband connectivity throughout the U.S., the digital divide has not completely vanished. Lack of an available high capacity broadband option represents the most immediate barrier to access, and this barrier is one that continues to receive attention by policy makers (Mossberger et al 2013). In places where at least some high capacity service is available, a lack of competition between providers may result in higher prices, or speeds may not be high enough to attract customers in an environment where higher and higher speeds are required for full internet utilization. These components of broadband supply only explain a portion of the variation in broadband subscription rates and are insufficient to explain the entirety of the digital divide.

[16] Not shown here, but available upon request.

A better understanding in her determinants of broadband uptake can be found by incorporating measures of both structural and demographic demand, with median income, age, percent urban, race, and education all playing substantial roles in predicting county broadband subscription rates. Even in counties with a high broadband availability rate, broadband uptake rate may be low as particular populations have lower levels of technical skills or place a lower level of importance on establishing a home broadband connection. It would be inappropriate to address lagging broadband subscription rates in these communities solely by increasing competition or the average speed of service, though for cases where income is a factor this approach may have the desired impact. Instead, investments in digital literacy may help to raise the rate of connection in these communities. Whether this occurs through investments in digital literacy via early education or through outreach programs to educate individuals later in life is beyond the scope of this paper and likely requires a customized approach tailored to different target populations. Overall, these models show that while investments in infrastructure are fundamental and important in improving broadband subscription rates, it will be impossible to close the digital divide without attention to cost concerns, gaps in community technical skills and differences in judgements regarding the value of having broadband internet at home.

Despite increases in national broadband subscription rates, continuing to track the digital divide remains important for a number of reasons. First, though the overall share of households subscribing to a broadband internet service is increasing (Ryan 2018), households in particular parts of the country may still face infrastructural burdens with regard to the availability of services. The costs to overcome these burdens may be greater than previous investments as these disconnected communities are increasingly rural and less densely populated (FCC 2010b; Martin and Lewis 2018). Even though these costs may be substantial, these communities may represent some of our most vulnerable populations, such as Native Americans living on reservations, lower income families, and elderly individuals. Second, though overall subscriptions are increasing nationwide, not all subscriptions are created equal. The FCC has recognized this by increasing the benchmark for a service to be considered "high speed broadband" from 4 Mbps download and 1 Mbps upload to 25 Mbps down and 3 Mbps up. In the near future, even higher speeds may be required for regular internet usage, requiring an upward revision of this benchmark. Were this to occur, particular communities and groups of people may again find themselves underserved digitally as they again lag behind wealthy, well-educated communities. Understanding the factors that contribute to our current digital inequalities will better prepare us to respond to these challenges in the future.

Next Research Steps

This paper lays the groundwork for a number of additional projects using these data. In this paper, I have presented a basic community level model of the factors that drive broadband subscription. While it provides a unique viewpoint heretofore absent from the literature, it can only go so far to describe the elements of the digital divide. As a county level model, it provides a good overview and starting point for this research, but many worthwhile analyses remain.

First, additional analyses should investigate the impact of spatial structure on these models. Given that the deployment of fixed broadband infrastructure is inherently spatial as cable connections need to tie consecutive blocks to one another, it makes sense to treat these models with some degree of spatial scrutiny. The maps provided in this report indicate that in a number of cases, spatial clustering may be occurring. I also mapped the residuals of my model and found evidence of moderate clustering, particularly throughout the south and Midwest, resulting in a Moran's I of 0.35. Incorporating spatial structure using a spatial lag or error model may correct for any spatial autocorrelation in these residuals and improve model fit. Additionally, use of a Geographically Weighted Regression model may allow for identification of counties where internet supply is high but less impactful in driving broadband uptake than structural or demographic demand factors.

I also fully intend to disaggregate these relationships in the future. Working with restricted household level American Community Service data, I will be able to test whether or not household subscription is conditional both on availability and contextual factors in a series of hierarchical models. These models should provide a greater deal of insight into which factors drive the digital divide for individual households, rather than for aggregated counties.

Lastly, these analyses only look at high capacity fixed broadband subscriptions. This was done in part due to the fact that these connections tend to meet the FCC's benchmark for high speed connections. These categories also neatly align with the estimates available in the American Community Survey. It would be possible to reproduce these analyses for other, non-satellite[17] connections. For example, researchers may be interested in the impact of Terrestrial Fixed Wireless connections on internet subscription rates, which are reported in the Form 477 dataset. This connection type is frequent in rural areas, offering generally slower speeds, but without costly broadband infrastructural investments on the part of service providers. Form 477 also reports mobile broadband, though unlinked from census geographies. Instead, these data are linked to shapefiles, which an interested researcher would need to re-merge with underlying census data in order to obtain a useful comparison. Regardless, both the Form 477 dataset and the American Community Survey present ample opportunity for detailed research into both community and individual determinants of the digital divide.

[17] Satellite availability is 100% throughout the country due to the nature of the technology.

Works Cited

Carlson, Edward and Justin Goss. 2016. "The State of the Urban/Rural Digital Divide." Washington, DC: National Telecommunications and Information Administration. Retrieved September 7, 2017 (https://www.ntia.doc.gov/blog/2016/state-urbanrural-digital-divide).

Council of Economic Advisors. 2015. *Issue Brief: Mapping the Digital Divide*. Washington, DC: Council of Economic Advisors. Retrieved September 7, 2017 (https://obamawhitehouse.archives.gov/sites/default/files/wh_digital_divide_issue_brief.pdf).

DiMaggio, Paul and Hargittai, Eszter. 2001. *From the 'Digital Divide' to 'Digital Inequality': Studying Internet Use as Penetration Increases*. Princeton University, Woodrow Wilson School of Public and International Affairs, Center for Arts and Cultural Policy Studies. Working Papers.

DiMaggio, Paul, Hargittai, Eszter, Celeste, Coral, and Shafer, Steven. 2004. "From Unequal Access to Differentiated Use: A Literature Review and Agenda for Research on Digital Inequality". *Social Inequality*.

Federal Communications Commission. 2010a. "Broadband Action Agenda." Retrieved February 12, 2019 (https://transition.fcc.gov/national-broadband-plan/national-broadband-plan-action-agenda.pdf).

Federal Communications Commission. 2010b. "The Broadband Availability Gap." OBI Technical Paper No. 1. Retrieved August 20, 2018 (https://transition.fcc.gov/national-broadband-plan/broadband-availability-gap-paper.pdf).

Federal Communications Commission. 2010c. "Connecting America: The National Broadband Plan". Retrieved February 12, 2019 (https://transition.fcc.gov/national-broadband-plan/national-broadband-plan.pdf)

Federal Communications Commission. 2015. "2015 Broadband Progress Report and Notice of Inquiry on Immediate Action to Accelerate Deployment." FCC 15-10. Retrieved August 20, 2018 (https://docs.fcc.gov/public/attachments/FCC-15-10A1.pdf).

Federal Communications Commission. 2016. "FCC Form 477: Local Telephone Competition and Broadband Reporting." OMB-3060-0816. Retrieved August 20, 2018 (https://transition.fcc.gov/form477/477inst.pdf).

Government Accountability Office. 2015. *Broadband: Intended Outcomes and Effectiveness of Efforts to Address Adoption Barriers are Unclear*. GAO-15-473. Retrieved September 7, 2017 (http://www.gao.gov/assets/680/670588.pdf).

Horrigan, John. 2010. *Broadband Adoption and Use in America*. OBI Working Paper Series No. 1. Washington, DC: Federal Communications Commission. Retrieved September 7, 2017 (https://transition.fcc.gov/national-broadband-plan/broadband-adoption-in-america-paper.pdf).

Martin, Michael J.R. & Jamie Lewis. 2018. "What Encourages Internet Service Providers to Provide Service? A Look at Merged Administrative and Survey Data." Poster presented at Population Association of America meetings; Denver, CO.

Mossberger, Karen, Caroline J. Tolbert, & Allison Hamilton. 2012. "Measuring Digital Citizenship: Mobile Access and Broadband". *International Journal of Communication*, 6, 2492–2528.

Mossberger, Karen, Caroline J. Tolbert and William Franko. 2013. *Digital Cities: The Internet and the Geography of Opportunity*. New York: Oxford University Press.

NTIA (National Telecommunications and Information Administration) 1998. *Falling Through the Net II: New Data on the Digital Divide*. July.

Obama, Barack. 2015. "Remarks by the President Promoting Community Broadband." Cedar Falls, Iowa. January 14th. Retrieved August 20, 2018 (https://obamawhitehouse.archives.gov/the-press-office/2015/01/14/remarks-president-promoting-community-broadband).

Ryan, Camille. 2108. "Computer and Internet Use in the United States: 2016." *American Community Survey Reports*, ACS-39. Washington, DC: U.S. Census Bureau. Retrieved August 20, 2018 (https://www.census.gov/library/publications/2018/acs/acs-39.html).

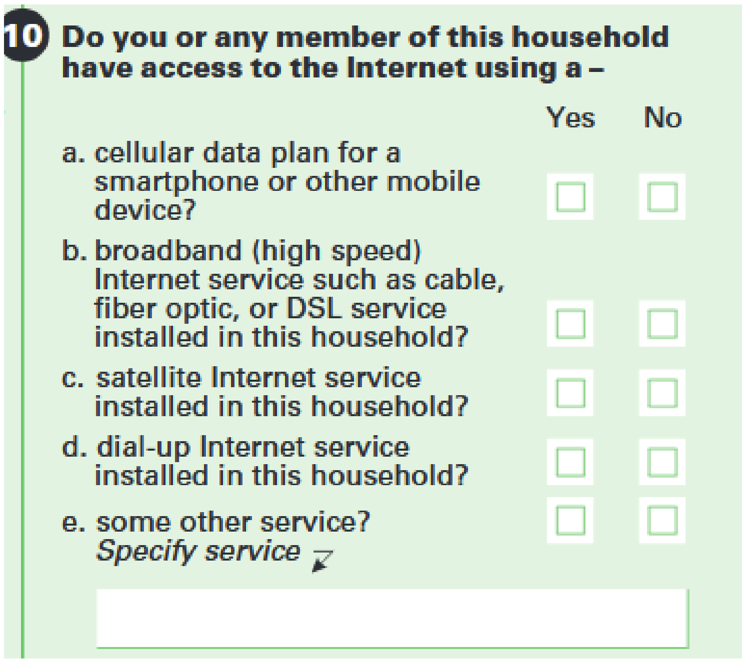
Ryan, Camille and Jamie Lewis. 2107. "Computer and Internet Use in the United States: 2015." *American Community Survey Reports*, ACS-37. Washington, DC: U.S. Census Bureau. Retrieved August 20, 2018 (https://www.census.gov/library/publications/2017/acs/acs-37.html).

van Dijk, Johannes A. G. M., & Kenneth Hacker. 2003. "The digital divide as a complex and dynamic phenomenon." *Information society*, 19:4, 315-326.

U.S. Department of Agriculture. 2017. "Report to the President of the United States from the Task Force on Agriculture and Rural Prosperity" U.S. Department of Agriculture. October. Retrieved August 23, 2018. (https://www.usda.gov/sites/default/files/documents/rural-prosperity-report.pdf).

Tables and Figures

Figure 1. 2017 ACS Internet Subscription Question


10 Do you or any member of this household have access to the Internet using a –

	Yes	No
a. cellular data plan for a smartphone or other mobile device?	☐	☐
b. broadband (high speed) Internet service such as cable, fiber optic, or DSL service installed in this household?	☐	☐
c. satellite Internet service installed in this household?	☐	☐
d. dial-up Internet service installed in this household?	☐	☐
e. some other service? *Specify service*	☐	☐

Table 1. Descriptive Statistics of Internet Availability data from Form 477 Filings, 2015

Variable	Mean	Std. Dev	Minimum	1st Quartile	Median	3rd Quartile	Maximum
% Households with High Capacity Service Available	86.7	15.0	0.0	81.1	91.7	97.5	100.0
Average *N* High Capacity Services Available	1.9	0.8	0.0	1.4	1.9	2.4	12.0
Average Download Speed for High Capacity Services	58.2	78.9	0.0	20.6	39.2	66.0	996.2

Source: December 2015 FCC Form 477 Fixed Broadband Deployment Data; U.S. Census Bureau, 2010 Decennial Census

Table 2. Example of FCC Form 477 Filing Data, 2015

HoCo #	Holding Company Name	Transmission Technology	Max. Advertised Download Speed	Max. Advertised Upload Speed
130077	AT&T Inc.	AxDSL	1.5	0.384
130077	AT&T Inc.	ADSL2	12	1
180057	The T1 Company LLC	Other Copper	12	12
130235	Charter Communications	Cable 3.0	100	5
131480	WideOpenWest	Cable 3.0	50	5
290111	ViaSat, Inc.	Satellite	12	3
130627	Hughes Network Systems	Satellite	15	2
300167	VSAT Systems, LLC	Satellite	2	1.3

Source: December 2015 FCC Form 477 Fixed Broadband Deployment Data

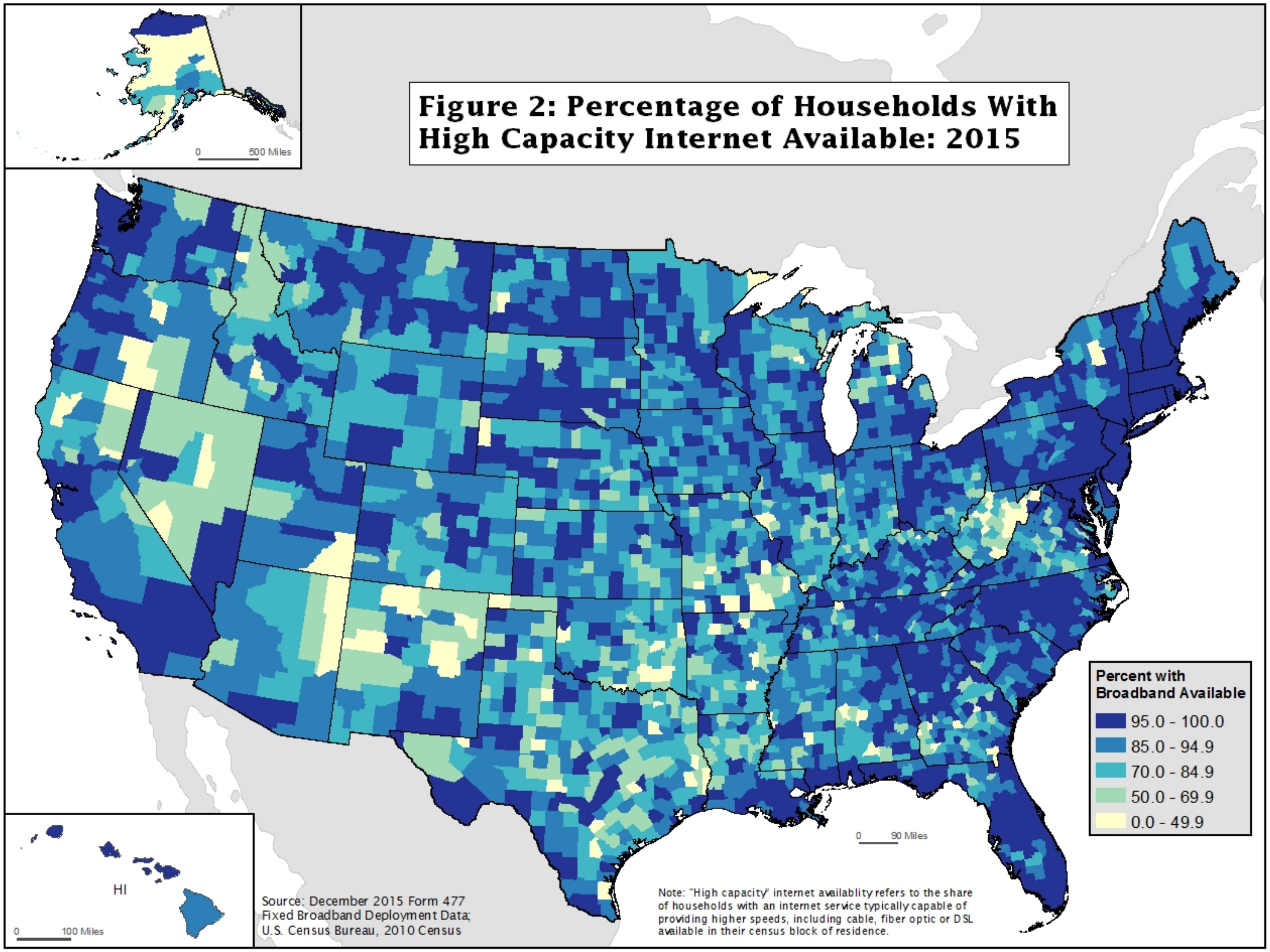


Figure 2: Percentage of Households With
High Capacity Internet Available: 2015
Percent with
Broadband Available
95.0 - 100.0
85.0 - 94.9
70.0 - 84.9
50.0 - 69.9
0.0 - 49.9
0 500 Miles
HI
0 100 Miles
0 90 Miles
Source: December 2015 Form 477
Fixed Broadband Deployment Data;
U.S. Census Bureau, 2010 Census
Note: "High capacity" internet availablity refers to the share
of households with an internet service typically capable of
providing higher speeds, including cable, fiber optic or DSL
available in their census block of residence.

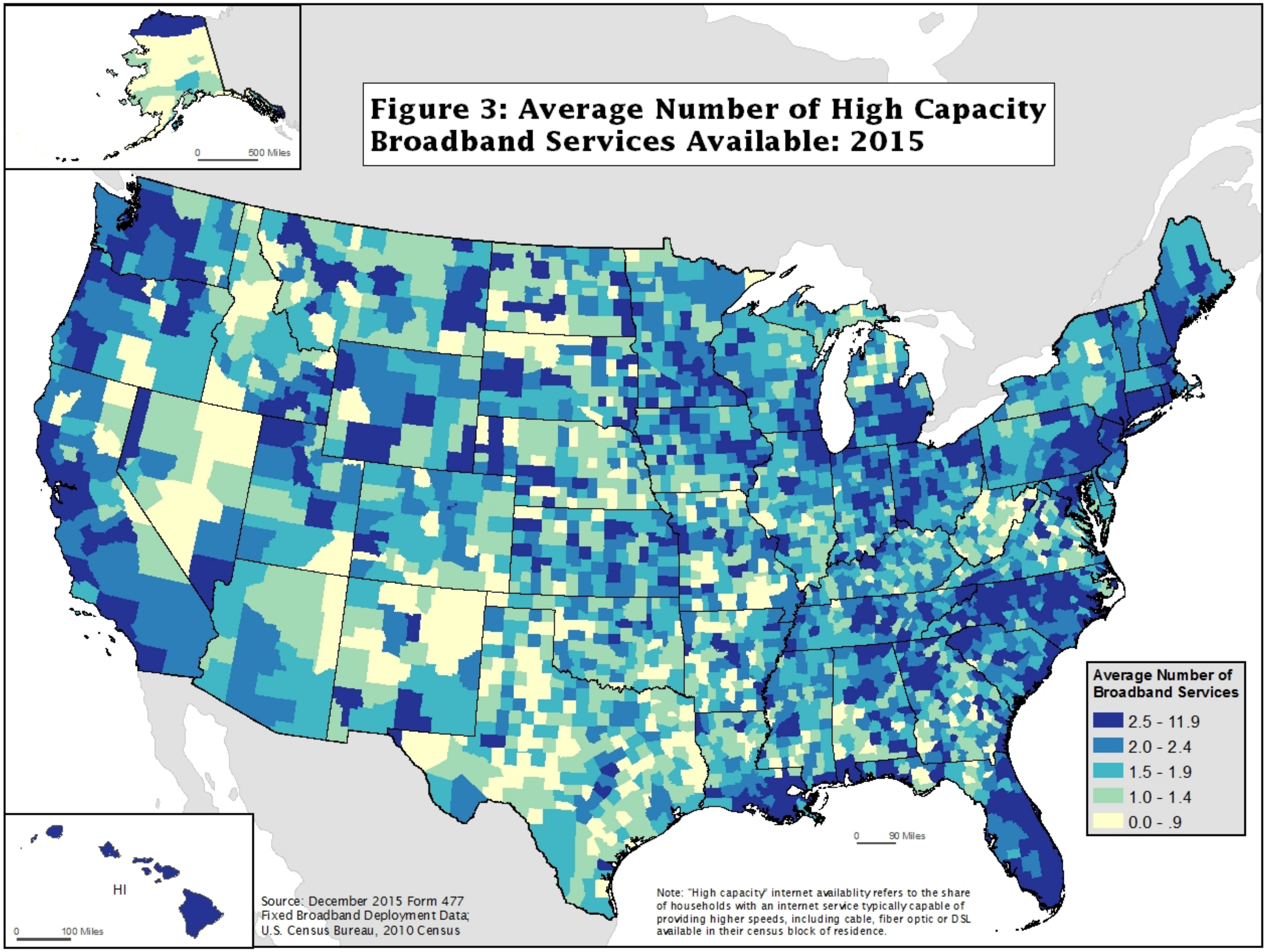


Figure 3: Average Number of High Capacity
Broadband Services Available: 2015
0 500 Miles
Average Number of
Broadband Services
2.5 - 11.9
2.0 - 2.4
1.5 - 1.9
1.0 - 1.4
0.0 - .9
0 90 Miles
HI
0 100 Miles
Source: December 2015 Form 477
Fixed Broadband Deployment Data;
U.S. Census Bureau, 2010 Census
Note: "High capacity" internet availablity refers to the share
of households with an internet service typically capable of
providing higher speeds, including cable, fiber optic or DSL
available in their census block of residence.

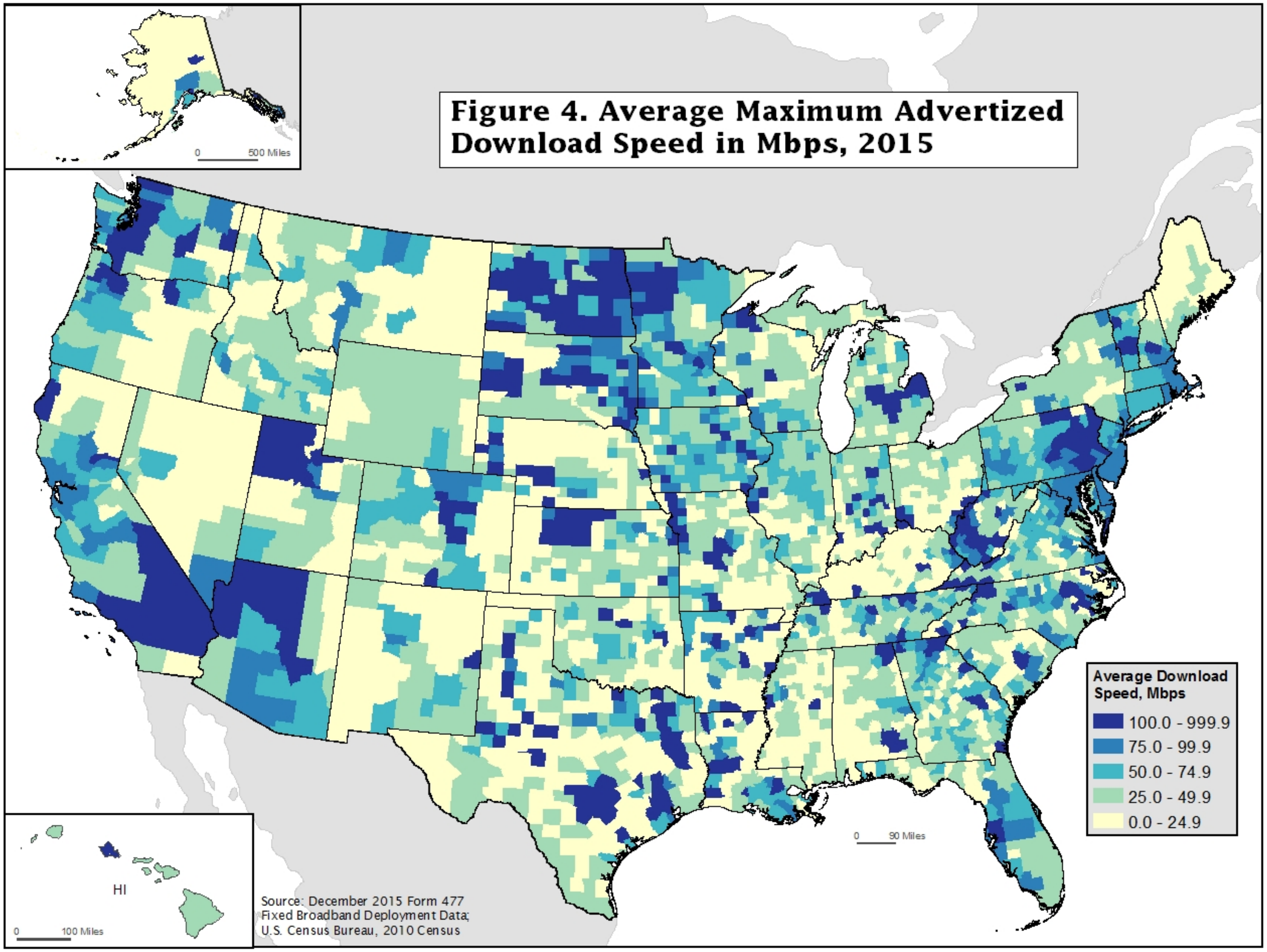


Figure 4. Average Maximum Advertized Download Speed in Mbps, 2015
0 500 Miles
HI
0 100 Miles
0 90 Miles
Average Download Speed, Mbps
100.0 - 999.9
75.0 - 99.9
50.0 - 74.9
25.0 - 49.9
0.0 - 24.9
Source: December 2015 Form 477 Fixed Broadband Deployment Data; U.S. Census Bureau, 2010 Census

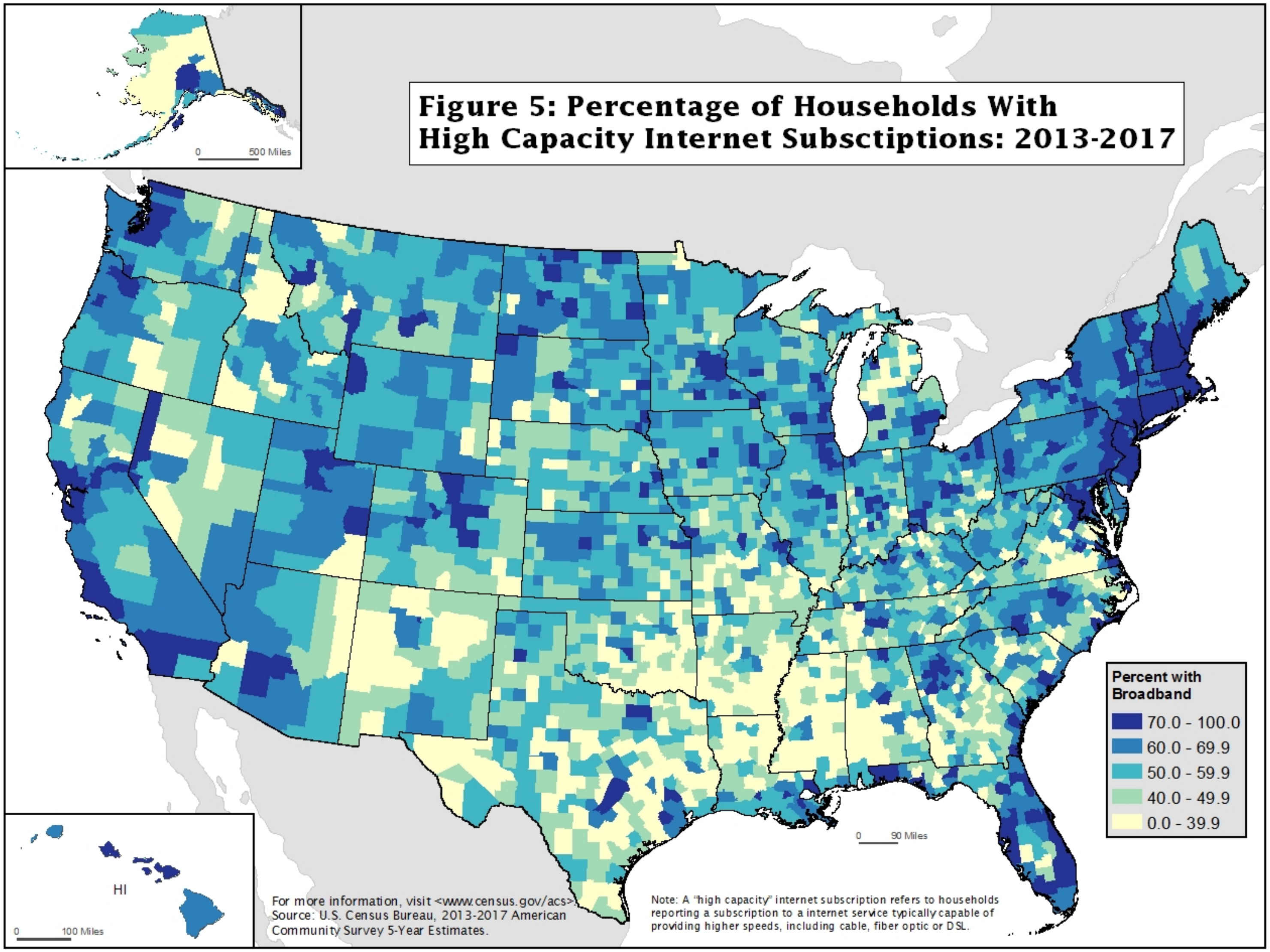


Figure 5: Percentage of Households With
High Capacity Internet Subsctiptions: 2013-2017
0 500 Miles
HI
0 100 Miles
0 90 Miles
Percent with Broadband
70.0 - 100.0
60.0 - 69.9
50.0 - 59.9
40.0 - 49.9
0.0 - 39.9
For more information, visit <www.census.gov/acs>.
Source: U.S. Census Bureau, 2013-2017 American Community Survey 5-Year Estimates.
Note: A "high capacity" internet subscription refers to households reporting a subscription to a internet service typically capable of providing higher speeds, including cable, fiber optic or DSL.

Table 3. Descriptive Statistics for Broadband Subscription and ACS Correlates, 2013-2017

Variable	Mean	Std Dev	Minimum	1st Quartile	Median	3rd Quartile	Maximum
% Households with High Capacity Subscriptions	53.6	13.6	9.4	44.4	54.6	63.2	88.1
% Urban	40.5	31.5	0.0	10.6	39.2	65.6	100.0
% Renters	28.7	8.3	6.5	23.1	27.4	32.5	96.2
Persons per square mile	259.1	1724.4	0.0	16.9	45.2	113.6	69468.4
% Households built after 2000	16.4	8.2	1.1	10.5	15.0	20.5	63.9
Median Income	49,754	13,154	19,264	41,122	48,065	55,764	129,588
Median Age	41.2	5.4	21.6	37.9	41.2	44.2	66.4
% Bachelors or Higher	21.2	9.3	4.7	14.7	19.0	25.3	78.1
% Black or Hispanic	18.0	18.8	0.0	3.8	10.0	27.9	99.2
% with a Disability	15.9	4.4	3.8	12.8	15.4	18.6	34.2

Source: U.S. Census Bureau, 2013-2017 American Community Survey 5-Year Estimates; U.S. Census Bureau, 2010 Decennial Census

Table 4. Ordinary Least Squares Regression Model of Broadband Subscription*

	1	2	3	4	5[1]
% HHs with High Capacity Service Available	0.988*** (0.021)	0.629*** (0.028)	0.368*** (0.025)	0.360*** (0.016)	0.231*** (0.016)
Average *N* High Capacity Services Available		5.575*** (0.411)	3.023*** (0.380)	0.661** (0.241)	0.029** (0.241)
Average Download Speed for High Capacity Services		0.068*** (0.004)	0.055*** (0.003)	0.010*** (0.002)	0.036*** (0.002)
% Urban			0.148*** (0.010)	0.134*** (0.007)	0.299*** (0.007)
% Renters			-0.280*** (0.014)	0.074*** (0.012)	0.073*** (0.012)
Persons per square mile			0.014*** (0.001)	0.004*** (0.001)	0.062*** (0.001)
% Households built after 2000			0.051*** (0.013)	0.092*** (0.009)	0.079*** (0.009)
% Bachelors or Higher				0.106*** (0.013)	0.105*** (0.013)
Median Income (log)				17.646*** (0.643)	0.411*** (0.643)
Median Age				-0.553** (0.192)	-0.214** (0.192)
Median Age (squared)				0.010*** (0.002)	0.316*** (0.002)
% Black or Hispanic				-0.115*** (0.006)	-0.208*** (0.006)
% with a Disability				-0.125** (0.045)	-0.039** (0.045)
Intercept	-27.898*** (2.068)	-12.374*** (2.094)	12.339*** (1.941)	-172.433*** (7.567)	0.000*** (7.567)
Adjusted R^2	0.403	0.487	0.634	0.858	0.858

*(Standard Errors in parentheses; *** = p<.001, ** = p<.01, * = p<.05)

[1] Model 5 contains standardized beta coefficients