Digital Infrastructure

By Shane Greenstein[1]

Comments welcome.

For NBER

Economics of Infrastructure Investment

November 15-16, 2019

Abstract

What determines variance in the supply of innovative digital infrastructure, and how does it shape economic outcomes? The first half of the essay covers the economic impact of deployment and adoption of access services, while the latter part of the essay covers complementary activities that enables internet access to deliver better performance. The latter discussion uses examples to illustrate broad observations and issues, especially where statistical research lags business practices. This essay emphasizes economic research about the United States and covers the global experience when possible. It stresses the large number of unanswered policy-relevant research questions.

[1] Harvard Business School and NBER. Contact at sgreenstein@hbs.edu. The author would like to thank Klaus Ackermann, Tim Bresnahan, David Clark, kc Claffy, Chris Forman, Avi Goldfarb, Maggie Kelleher, John Laprise, Kristina McElheran, Abhishek Nagaraj, Frank Nagle, Tommy Pan Fang, Bill Rogerson, Greg Rosston, Alicia Shems, Scott Stern, Catherine Tucker, and Pai-Ling Yin for conversations and comments. All errors are mine alone.

I. Introduction

The deployment and adoption of the commercial internet in the 1990s brought about a major restructuring of digital infrastructure, and it supported growth of many visible and innovative digital services. Today that includes a range of innovative businesses in the “sharing economy” (e.g., Uber, Lyft, Airbnb), social media (e.g., Facebook, Twitter), mobile information services (e.g., ticketing, instant messaging), and many other facets of electronic commerce (e.g., electronic retailing, and ad-supported media). In 2017, electronic retailing reached over $545 billion for “Electronic Shopping and Mail Order Houses (NAICS 4541).” This category grew 65% from 2012. In 2017, online advertising contributed $105.9 billion in revenue among “Internet Publishing and Broadcasting and Web Search Portals (NAICS 519130).” This category grew 250% since 2012. All were much smaller in the 1990s, and their operations has changed dramatically in a few decades.

This essay addresses gaps in economic research about the less visible parts of this experience, namely, the digital infrastructure. Perceived as activities below the “application layer,” these are the activities associated with enabling the “transport layer” and “networking layer” of the internet. More concretely, the review analyzes the supply of many special types of equipment for enabling the commercial internet —root servers, fiber, broadband lines, networking switches and routers, content delivery networks, cloud facilities, and cellular towers. Interconnected at an extraordinary and unprecedented scale across the globe, this equipment works directly with the privately owned investments of millions of content providers and billions of user applications. Many firms and organizations involved in digital infrastructure—e.g., data center operators, name-server providers, content delivery network specialists, data carriers and access providers—supply these services.

What determines variance in the supply of innovative digital infrastructure, and how does that variance shape the performance of digital services? The essay reviews studies from a number of areas, with an emphasis on research in innovation economics, industrial economics, growth accounting, and urban economics. As illustration, why did some frontier applications – for example, in streaming – blossom only recently, and what role did improvements in digital infrastructure play in that rollout? Is the contribution of digital infrastructure to growth correctly measured, and if not, what lessons carry over to measuring pricing and qualitative improvement? Does statistical evidence support the view that digital infrastructure contributes to better performance of frontier data-intensive applications in some locations and not others, and if so, are these differences temporary or becoming greater over time?

The essay presents research in two related groupings. Section II asks questions about the determinants of, and impact from, access networks. The spread of dial-up networks accompanied the commercialization of the internet and satisfied the earliest demand for access at households and business. Soon after privatization, a variety of entrants aimed at providing faster speeds than delivered by dial-up. Today these access technologies go by the label *broadband*.[2] Today broadband providers dominate supply of access services to the internet for both households and business establishments. The economics of broadband have received more attention than any other topic. It also provides the starting point into a larger conversation about the impact of digital infrastructure, because most applications of digital infrastructure require users to access the internet at high bandwidth.

Section III discusses the less visible digital infrastructure that enables wider access or broadly enhances the quality of the performance of internet. There are many such activities, and considerable innovative improvements to them, but the economics of these activities have received less attention than that given to access. This discussion uses examples to illustrate broad observations and issues, especially where there is a lack of statistical research, and many open questions remain. As the review shows, the changing features of the landscape generate many challenges for statistical analysis. The discussion divides its attention into two sections. Section III covers a range of privately supplied activities that help the internet deliver better performance. Section IV discusses protocol development and issues in the governance of many types of digital infrastructure.

[2] The definition of broadband has undergone changes over time, as regulatory expectations change. For purposes of this discussion, the definition will be loose, and encompass any wireline technology faster than the data rates of 56k dial-up, including ISDN, DSL, and cable modem service. Among wireless technologies, all Wi-Fi (IEEE 802.11b, g, n and more), 3G, 4G and 5G cellular services, as well as modern satellite service, are broadband.

It is important to state the scope of the review. This review emphasizes economic research about the United States, and covers the global experience when possible. Its principal goal is to inform research and policy analysis, not to advocate choices over policy. For example, while the review informs analysis of "net neutrality" – i.e., questions about alternative proposals for regulating access – it does not discuss tradeoffs in specific proposals. A curious reader can go to many other sources.[3] Additionally, the review does not provide a comprehensive and detailed explanation of the routine operations of the internet, nor does it explain much of the minutiae behind regional operations of digital infrastructure. Again, this is available from other sources.[4]

II. Access Networks

Dial-up internet access blossomed in the 1990s. It built upon the existing telephone network, which was nearly ubiquitous prior to the diffusion of the internet. Along with regulatory approaches that encouraged a competitive supply of services that used dial-up, the first generation of dial-up access was available almost everywhere in the United States within a few years after privatization of the network (Downes and Greenstein, 2002). Broadband networks did not replicate this achievement. Broadband networks required upgraded physical investments. In low-density areas, the costs of building such lines were high, and discouraged any buildout.

In 2001, only about one-half of U.S. households had access to the internet, and virtually all access occurred over dial-up; whereas today, approximately three-quarters of U.S. households have broadband internet access in their homes.[5] The vast majority of business users in urban and suburban areas also contract for broadband internet access. In its most common form, firms

[3] As of this writing, this literature has focused on theoretical and legal issues, sometimes supported by examples, and not econometric measurement. For a review, see e.g., Neuchterlein and Weiser (2005) for a thorough review of the origins of many regulatory rules at the outset of the internet, and Greenstein, Peitz, and Valleti (2016) for a review of research economic research on the topic.

[4] For a thorough description of the architecture and its evolution, see Clark (2018). An economic history of the growth of the commercial internet from its government roots, see Greenstein (2015). For description of the basic operations of the transport and internet layer today, and the origins of congestion, see Clark et al (2014). For a description of its basic economics of data networks, see Greenstein (2020).

[5] Greenstein and McDevitt (2011) provide data on the replacement of dial up with broadband. For the diffusion of broadband, see https://www.census.gov/content/dam/Census/library/publications/2017/acs/acs-37.pdf, and https://www.pewInternet.org/fact-sheet/Internet-broadband/.

offer broadband as either DSL (digital subscriber line) over a phone line or through cable modems retrofitted to cable television systems. More recently, broadband over fiber has become more widely available.

Economies of density shape the cost of supplying broadband, both in the United States and across the globe. In developed countries, most suburban locations saw options from one or two wireline providers, with occasional overbuilding leading to more. Downtown locations in high-density settings could support greater entry, aimed at business customers and/or multi-occupation residences.[6] Beyond those simple statements, the actual experience with entry depended on a host of factors, such as regulatory rules for pole attachments and ease of interconnection

II.1. Adoption and use

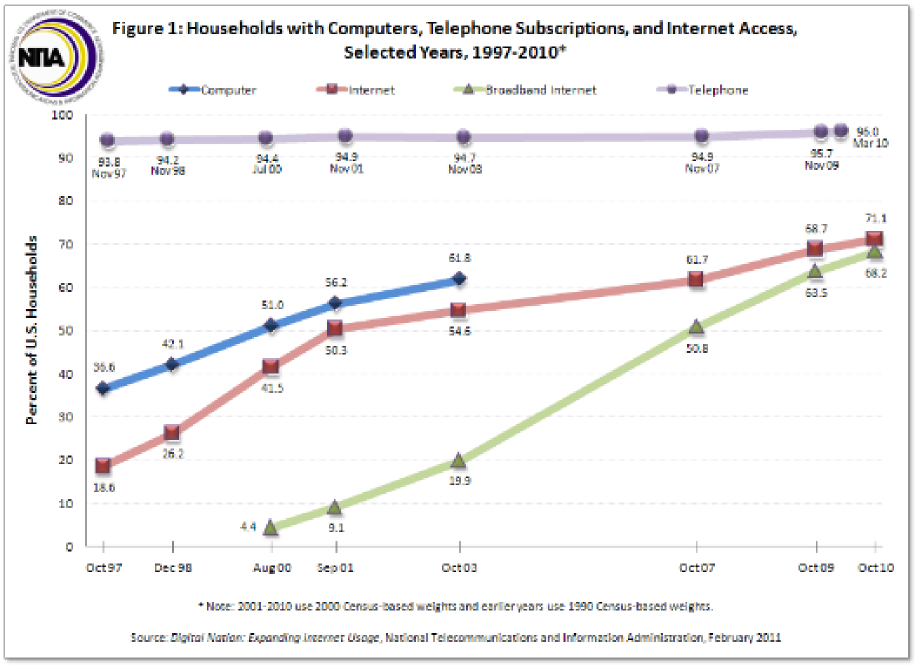
Adoption of broadband to households followed a standard "s-curve," with one-half of U.S. households moving to broadband in 2007, continuing to growth thereafter. (See Figure 1, from a publication in 2011,[7] for an illustration.) Eventually diffusion fell short of universal adoption due to lack of interest, lack of income, and occasionally lack of availability (i.e., in low-density areas). Today, adoption of broadband internet in the US hovers just below 80% of households. In most developed countries, the percentage is higher. Worldwide, close to half of the global population uses the internet.[8]

The growing importance of digital infrastructure is also visible in other access markets. For example, users of cellular telephony migrated from 3G to 4G, the latter entirely supporting digital communications.[9]Presently, and more visibly, more than three-quarters of U.S. households own at least one smartphone, rising from virtually none in 2007.[10] In addition, Wi-Fi

[6] See, e.g., Connolly and Preiger (2013), and Wallsten and Mallahan (2013).
[7] https://www.ntia.doc.gov/report/2011/digital-nation-expanding-Internet-usage-ntia-research-preview
[8] World Bank statistics. https://data.worldbank.org/indicator/IT.NET.USER.ZS?view=chart
[9] 4G is the fourth generation of broadband cellular technology, succeeding 3G. 4G uses only packet switching technology, unlike 3G, which used both packet-switching and (in parallel) the (old) circuit-switching technology. As of this writing 5G contains much more capacity than 4G, and has only just begun to deploy in developed countries.
[10] https://www.census.gov/content/dam/Census/library/publications/2017/acs/acs-37.pdf, accessed April 16, 2019.

technology also has improved and spread over the last two decades.[11] More than 86% of homes with access to broadband now employ some form of Wi-Fi for accessing applications, a usage that, like smartphones, started from nothing in the mid-1990s.[12]

Figure 1: Internet adoption over time



To determine what economic benefits accrued from switching to broadband requires focusing on one set of providers. Doing so not only offers a start to understanding the benefits derived from government subsidies of communications networks, it can also motivate to specific policy proposals. Government subsidies for high-speed access networks arise partly due to analogies with local telephony, in which many providers received building and operational subsidies from universal service programs.[13]

With this experience as motivation, some research inquires about non-adopters of access, and infers lessons from understanding the determinants of demand. For example, Rosston and

---

[11] Wi-fi is a standard defined by IEEE committee 802.11, operating over the 2.4 GHz and 5 GHz bands of spectrum.

[12] https://www.ncta.com/whats-new/wi-fi-how-broadband-households-experience-the-Internet, accessed April 16, 2019.

[13] There were many proposals for rural subsidies of broadband as part of the 2009 stimulus package, and they built on a previous set of subsidies in the E-rate program, which were established by the 1996 Telecommunications Act.

Wallsten (2019) examine the impact of the Internet Essentials program, sponsored by Comcast to foster adoption of broadband by qualified low-income households in the parts of the US where Comcast provides service. This program provides service at a heavily discounted price, and thus sets up an experiment to test the proposition that low-income households are reluctant to adopt due to high prices. The statistical challenge involves estimation of *new* adoption – i.e., some households who qualify for the program would have adopted at the regular higher price. By comparison with adoption among similar households in similar areas that lack such programs, Rosston and Wallsten show that demand does grow among the target population in areas where the Internet Essentials program operates, suggesting the program supports hundreds of thousands of users who would not have adopted otherwise. They also show that a substantial fraction of current non-adopters (potential users) are insensitive to price—in the sense that a large numbers of non-adopters do not change their behavior in spite of these massive price reductions. The latter result suggests non-adoption among the laggards does not have economic roots, and, if policy seeks near-universal adoption, requires non-economic focus.

Following a long tradition in communications studies, another set of studies focuses on rural broadband, and uses variance between rural areas for econometric identification of demand.[14] An interesting fact complicates inference: broadband satellite has been available in virtually every location, and for many years. For many uses, such as email, browsing, and non-interactive internet services, satellite broadband is technically sufficient, albeit more expensive than broadband in a typical suburban location. With such facts as motivation, Boik (2017) investigates user behavior in low-density North Carolina, and examines willingness to pay for satellite broadband versus willingness to substitute for wireline broadband. He finds considerable willingness to pay for wireless access, which, in turn, limits the potential welfare gains from subsidies for building out wireline access. Indeed, it renders uneconomic most subsidies for wireline services in low-density locations.

Another line of research measures heterogeneity in user demand for access. For example, Rosston, Savage, and Waldman (2010) examine the demand for more speed as one of many attributes consumers choose to pay for. They find that a small set of users will pay for higher

[14] See e.g., Whitacre, Gallardo and Strover (2014).

speeds at any point in time. The result suggests that user (un)willingness to pay for more speed acts as a fundamental brake on investment in upgrades in the short run, and as households migrate to higher speeds it is consistent with gradual unmeasured improvements.

Building on these estimates, Greenstein and McDevitt (2011) analyze the returns to households from upgrading to broadband from dial-up access. Despite the low valuations for frontier speeds, they show that the broadband upgrade over dial-up conferred large consumer surplus on the economy. More provocative, the consumer price index for access underestimated those gains, which would have generated at least a 2-3% decline in prices each year. They conclude that even the most conservative estimates of quality adjustment suggest there is an underestimate in standard economic measurement of pricing.

II.2. Pricing

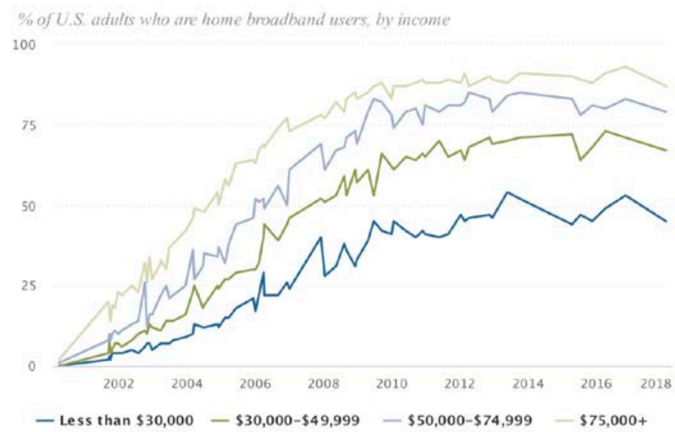
On the surface, internet access adoption would seem to follow the classic model of adoption, whereby those with the greatest willingness to pay adopt earliest, and those with lower willingness to pay adopt later—due to declines in price, increases in quality, or both. In this model, the value to consumers provides the value of access in terms of consumer surplus. This model has considerable appeal because it also provides a path towards valuing improvements from access infrastructure.

On the surface, the classic model would appear to be a good approach for measurement. After all, just contrast 2001 with 2016. Around 2001, dial-up dominated access to the internet, and approximately half of the U.S. households were online. Web traffic dominated the internet, and wireless access had just entered a new era with the introduction of Wi-Fi and 3G cellular service, which ran a data service in parallel with voice services on cellular towers and handsets. By 2016, broadband access dominated all modes of access, and three-quarters of U.S. households maintained connections online. In 2016, the predominant applications leading to data traffic were streaming, video, and gaming;Wi-Fi5 and 4G served as the predominant wireless modes of transmission. This fifteen-year history suggests a large and valuable increase in access networks that should manifest in price declines, quantity increases, and qualitative improvement.

One positive symptom of improvement shows up in GDP (especially after the Census reclassified activities to help with tracking such activity). To date, from 2012 to 2017, payments for access to wireline forms of internet access reached $88.7 billion, growing more than 30% in those five years. In addition, payments for access fees to wireless service reached over $90.0 billion, an increase of 57%.[15]

An estimate for user adoption by income, shown in Figure 2, also seems to fit the model.[16] While adoption grows across all demographic groups, the variance in adoption across income is readily visible. The persistent pattern—with lower income groups adopting less frequently—motivates hypotheses that high prices deter low-income households from purchasing internet access.

Figure 2. Broadband adoption by income level[17]



Note: The Center has used several different question wordings to identify broadband users in recent years, which may account for some variance in broadband adoption figures between 2015 and 2018. Our survey conducted in July 2015 used a directly comparable question wording to the one conducted in January 2018.
Source: Surveys conducted 2000-2018.
PEW RESEARCH CENTER

Other parts of the measured record contains a somewhat more ambiguous record. The growth displayed in Figure 2 (and Figure 1) ought to arise from either declines in prices or increases in quality, or both. Yet, the consumer price index (CPI) for access, which since the mid-2000s has

[15] Statistics of U.S. Business, U.S. Census.
[16] These graphs aggregate periodic surveys conducted by the Pew Internet and American Life Project (the data are not smoothed).
[17] See https://www.pewInternet.org/fact-sheet/Internet-broadband/, under "Who has Home Broadband?" with income as the primary sorting variable. Accessed August, 2019.

de facto covered broadband delivery, has remained flat for an extended period of time after a one-time drop in the middle of the decade. For example, in 2007, the CPI was at 73.2, and in 2018 it was at 76.0.[18] In other words, after a dozen years, the consumer price of broadband has *increased* by 3.8%. The puzzle does not disappear when this pattern compares with other indices. The closest comparable CPI—that for wireless services, which also includes the price of telephone calls—displays a drop from 64 to 46 (a 28% decline in prices).[19]

Simple alternative explanations do not provide an answer. Increased adoption cannot account for the rise in GDP in the face of no price change. From 2011 to 2018, approximately 3% to 5% of U.S. households first began using broadband internet, depending on the survey. That is far too small to account for a 30% growth in revenue.[20] Expenditure per household must have gone up, but how did that happen without a nominal price decline?

One explanation stresses that quality must have improved, and not been measured. Some evidence suggests this is the case. For example, there is evidence of increasing speeds over time for all the major wireline networks.[21]

There are several potential reasons for why that speed increase goes unmeasured. First, as with many other consumer services, the CPI for broadband compares the prices of contracts for a given service.[22] The procedure reduces or eliminates measuring qualitative improvement if contracts do not reflect those improvements, and CPI employees do not impute them. Better caching, buffering, and other intermediate features do not factor directly into pricing in contracts, and is consistent with this concern.

---

[18] See the series for Internet services and electronic information providers in U.S. city average, all urban consumers. https://data.bls.gov/PDQWeb/cu.

[19] See Wireless telephone services in U.S. city average, all urban consumers. https://data.bls.gov/PDQWeb/cu.

[20] Relatedly, the revenue increase did not largely arise from more adoption. See https://www.pewInternet.org/fact-sheet/Internet-broadband/, and https://www.census.gov/content/dam/Census/library/publications/2017/acs/acs-37.pdf.

[21] See, e.g., the Netflix comparisons of measured speeds over 2012-2018 yields a doubling of realized speeds for most networks (https://ispspeedindex.netflix.com/country/us/, accessed April, 2019).

[22] The CPI is constructed from a weighted average of contracts for ostensibly similar services, where the weights come from household surveys and the contracts come from suppliers. This necessarily underestimates the introduction of new goods—here, experienced as higher speeds—and qualitative improvements not reflected in common measures, such as bandwidth.

More subtle, contracts measure bandwidth, and increasingly appear as part of a tiered menu of quality and price. If households do not use the same contract over the entire period, it is possible for them to increase expenditure on access without any measured price increase in a CPI. Hence, existing procedures also creates an upward bias in the price index that fails to account for users switching to better contracts, which is a variation on an "outlet bias."

Another explanation stresses issues with definitional boundaries. The price measurement system treats access as a distinct service from content. Changes in the quality of content play no role in the price index for access. Said another way, the standard measurement framework focuses on transactions for access, not freely available services that users obtain along with their access. Yet it is undeniable that content, broadly construed, has improved over the long run—from static email and browsing to search to social media to streaming and interactivity. In short, no economic account measures improvements in the average quality of those complementary services. Those could motivate more adoption over time, as well as purchases of more bandwidth. Yet we only see the result, more expenditure; not the cause, more quality.

Such inadequacies frame numerous research questions. For example, Wallsten and Mallahan (2010) analyze the determinants of the quality of broadband firms. They examine the effect of more entry on the quality of the broadband provided, measured by its bandwidth, and then exclude suppliers from the count unless those suppliers provide to a minimum threshold of customers. They found that the typical zip code contained one or two suppliers of broadband, and a small number had three or more. Their analysis shows that the third entrant does not change pricing, but does generate competitive pressures for qualitative improvement.[23]

A related approach goes to industry data and confines attention to small cities with no more than two broadband suppliers. Chen and Savage (2011) focus on the role of competition in shaping pricing. They match cable and DSL internet access providers in all the western states, and

[23] A related approach focuses on the effect of city governance on broadband upgrades. Note the following two examples: Seamans (2012) examines whether perceived threats of municipal entry generate faster upgrades and finds that it does. Using micro-evidence, Skiti (2019) examines whether potential competitive entry generates any response and finds that it does.

compare pricing differences between monopolies or duopolies. They find that demand for variety mediates pricing and sometimes prevents price decline from the entry of an additional supplier.

Another promising approach analyzes both the contracts between users and access firms, and subsequent user behavior. Today usage-based pricing and data caps in wireline access contracts are more common, but only a little research has modeled the adoption and use decision in the presence of these contracting constraints. Nevo, Turner, and Williams (2016) provide a framework for doing so. They analyze usage data for a set of customers of a single ISP (Internet Service Provider). These users face distinctive three-part tariffs, which impose a shadow value on the price of data as users approach their monthly allowances.[24] Users are sensitive to the charges affiliated with reaching a data cap, but they also endogenously select into capacity consistent with their own use. Variation in user behavior permits an analyst to recover variation in the willingness to pay for broadband. Such willingness-to-pay estimates provide insight into the gaps between private and social incentives to build or upgrade broadband. Nevo, Turner, and Williams (2016) estimates suggest the gap is substantial.[25]

Byrne and Corrado (2019) approach the valuation issues by focusing on valuing the missing free complements. Borrowing insights from measurement of capacity utilization, they argue that IT (information technology) capital is used more intensively by consumers who use access technologies for free goods.[26] The complementarity between paid access services and network use leads to a quality adjustment for the price of access services. Looking across all consumer network services, namely, cable television, cellular telephony, and the internet, they calculate a nearly $1800 boost to consumer surplus per connected user, which amounts to a one-half percentage point addition to U.S. real GDP for 2007–17. That suggests the derived demand for access infrastructure is large, and so is its underlying value. None of the free services could

[24] See Burnham et al. (2013) for an early census of the use of tiered pricing and caps based on the usage of data in wireline and wireless forms.

[25] Malone, Nevo, and Williams (2017) also examine the willingness to pay for more bandwidth, based on usage data from one ISP. They focus on the tradeoffs for different ways to approach congestion of networks. They show that peak load pricing along with caching more effectively deals with congestion than does throttling of traffic

[26] This approach follows numerous studies that examine the time spent online as a possible avenue for valuing digital goods. See, for example, Goolsbee and Klenow (2006), Hitt and Tambe (2007), Goldfarb and Prince (2008), Brynjolfsson and Oh (2012), Brynjolfsson, Collis, Eggers (2019), Boik, Greenstein and Prince (2019).

provide such satisfaction without employing digital access and relying on nearly ubiquitous access.

The progress in understanding the experience outside the US has tended to take advantage of data and idiosyncratic institutional details that create opportunities for insights. An early, important paper in moving this topic to international comparisons is Wallsten and Riso (2010), which shows there is wide variance in access prices and availability across countries. Yet, that has not stopped naïve approaches that reduce the nuances in prices to one statistic to facilitate comparisons.[27] Wallsten and Riso's findings suggest that single statistics heightens the potential for unobservable factors in cross-country regressions.

Another set of studies examines the deployment of broadband in the United Kingdom. Due to the makeup of the U.K.'s underlying switch network, broadband deployed in a somewhat random geographic pattern, creating similar neighboring areas with different broadband experiences. This quasi-randomness creates plausible exogeneity and can identify the consequence of speed on productivity. One line of research, for example, looks at the consequences of uneven broadband deployment on property prices for homes (Ahfeldt, Koutroumpis, Valleti, 2014) and finds that the large impact on local prices for housing suggests that consumer valuation of broadband effects pricing of real estate.

II.3. Business use

Diffusion of the internet created two investment trajectories at business establishments, a framework first proposed for a sample of adopters by Forman (2005), and developed into a framework for analyzing the US economy by Forman et al (2005). One type of investment involved developing access to support email and browsing for employees. Additionally many businesses had to change their business processes. This involved investments to supply services for customers and to receive inputs from suppliers. These tended to involve distinct investment

[27] Most commonly used are OECD, 2014, or WEF, 2016, which get their broadband prices from the same source, data from the World Bank.

due to the different costs coordinating with partners, and due to the distinct complementary investments associated with adapting to the constraints of partners (McElheran, 2015).

Forman et al (2005) compare use of *basic* with *advanced* internet technologies at U.S. businesses near the end of the first wave of investment after the commercialization of the internet. They show that almost every establishment (approximately 90%) adopted the basic internet, while advanced internet showed up more prevalently in some cities. Several factors played a role in the deployment of advanced internet technologies. Some locations contained data-intensive industries that had recently made capital investments in computing and business equipment, which raised the returns to complementary investments in digital infrastructure. Some locations contained businesses with a distinct type of labor. More educated and more skilled labor could take advantage of digital infrastructure, again, raising the returns. Finally, some businesses were more productive and more profitable than other firms, and, thus, could make bigger investments in all capital equipment, including digital infrastructure.

The experience within business in the 1990s creates difficult challenges for measurement. The deployment of email and browsing cannot generate insight into whether adoption of novel digital technologies had an impact, because these basic internet technologies became available *and adopted* almost everywhere in the US within a few years, leaving no variance from which to infer the gains. What can be inferred? It is possible to examine changes consistent with adoption of advanced internet technologies, which required broadband and complementary assets, and for which there is tremendous measured variance across regions and industries.

Forman et al (2012) consider this variance for the question: Did investment in advanced internet technologies become associated with alleviating or acerbating regional inequality? Building on research linking information technology use to productivity gains,[28] an optimistic view forecasts that digital infrastructure potentially could reduce distances and aid those who lived at a distance from areas with higher incomes. They find, in contrast, that the first wave of the investment boom exacerbated regional inequality. Using an instrumental variable approach and a battery of additional tests to account for the endogeneity of the investment, they relate wage growth to

[28] For a recent review see e.g., Cardona, Kretschmer and Stobel (2013).

investment in advanced internet technologies. They find that business adoption of the internet makes regions with higher income richer in some places, but not everywhere. The largest divergences occur in major urban areas with skilled workforces and considerable prior investment in IT. In short, the high-income locations experienced the most wage growth. Note the large open question this raises: there has been no follow-up since to observe whether two decades of subsequent growth have altered the pattern.

Inadequacies in data make it challenging to infer the productivity effects of broadband. A researcher typically has access only to either (a) available supply of broadband or (b) purchased supply of broadband. Each suffers from a distinct form of endogeneity bias. There are additional challenges of measurement. At one time, the Federal Communication Commission (FCC) ostensibly tracked the former at the geographic level of the zip code but counted any firm as a supplier if it had one customer in that zip code. By including satellite suppliers, it came to numbers that reached maximal levels. It is challenging to find an econometric escape from such limited data. Today the federal government provides a broadband map of availability; availability does not tell us about adoption or use.[29]

Researchers have faced these challenges with varying approaches. One approach, taken by Kolko (2012), examines different indicators of economic change affiliated with broadband—growth in information industries, wages, employment, telecommuting, and home-based work—and focuses the investigation on medium-sized cities where exogenous instruments might be plausible, such as the topography of an area. He finds that many indicators of infrastructure do improve economic activity, but not the key ones affiliated with taxes, such as wages and employment.

The experience outside the US has some parallels, but also generates new situations which provide additional insight. DeStafano et al (2018) take advantage of the uneven rollout of DSL in the UK, and link that detailed information about firm productivity. They find that the impact of broadband on business productivity is modest at best. They do see, however, that broadband is associated with restructuring the location and scale of activity. These results suggest

---

[29] https://broadbandmap.fcc.gov/#/

complementary investments can play a significant role in fostering restructuring organizations, even when no short-term productivity improvement is visible.[30]

Using extraordinarily details data on wages, Poliquin (2018) finds another parallel experience with business adoption of broadband at Brazilian firms. Overall, wages increased 2.3% on average at establishments following the establishment's adoption of broadband. Consistent with the theory of biased technological change, wages increased the most for workers engaged in non-routine cognitive tasks, while returns were negative for routine cognitive tasks. There was no effect of broadband adoption on wages for either routine or non-routine manual tasks. He also finds that better recruitment of new employees did not shape broadband's effects, suggesting the skill bias arises from changes within an existing labor force, not additions to it.

The search for evidence about the economic impact of access has started to extend to many topics around the globe. For example, one set of studies examine the impact the global spread of digital infrastructure spread had on trade, such as Fernandes et al (2019). They examine export behavior in China during the period 1999–2007, and links firm participation in export markets to the rollout of the internet. They combine firm-level production data with province-level data on internet penetration and availability. Manufacturing rose during this period, and they look for evidence to explain how the internet contributed, if at all. They find evidence of improvements in communication with buyers and input suppliers coincident with a more visible virtual presence. Like other studies, they find that improvements depend on the availability of broadband, but broadband alone is insufficient to explain all the increases in manufacturing. They stress the role of numerous complementary investments to implement productive uses for broadband in firm processes.

Deployment of broadband access also generates symptoms of economic growth, especially in locations that previously lacked any wireline access. Hjort and Poulsen (2017) examine the gradual laying of fiber along the African coast, which enables wireline access where it

[30] This is in line with other work on the impact of broadband on the productivity of business, which also find modest effects on productivity, but measurable changes in other firm attributes in the presence of complementary investments (see, also, Destafano et al., 2019, and Haller and Lyons, 2015, 2019).

previously was only possible by satellite. This experiment creates many points of comparison between the served and underserved areas, including improved availability and adoption.[31] Their estimates show large positive effects on employment rates, with especially large increases in high-skill occupations. Remarkably, they also find gains in employment for less educated workers, though it is smaller. Firm-level data indicate that increased firm entry, productivity, and exporting contribute to higher net job creation, though such data are not complete.

II.4. Open questions

It is important to recognize gaps in the literature on access. To date a disproportionate amount of economic research focuses on the impact of *wireline* broadband access, and has not been matched by similar work about *wireless* access, where there has been an explosion in recent adoption **–** i.e., particularly in smart phones and tablets (usually tethered to Wifi routers). All evidence suggests mobile devices have grabbed increasingly higher proportions of user time, and involve increasingly higher fractions of traffic, consistent with high economic value to users. New research could inform questions related to policies for wireline and wireless access.

Questions about the spectrum for 5G and Wi-Fi provides another illustration of open questions. Spectrum is part of the electromagnetic waves that "wireless" digital signals traverse. Under the rationale that ownership motivates investment that supports rapid deployment, in recent decades many governments in the most advanced economies have auctioned exclusive rights to bands of spectrum, and done so with few or minimal obligations.[32] While auctions allocated most of the spectrum for 3G and 4G applications in many developed countries, another common practice allows shared use of spectrum, which removes licensing altogether. The popularity of Wi-Fi illustrates the potential for success with such non-exclusive allocations. Exclusive ownership and shared spectrum support different models of investment and benefits. Today there are many open questions over the merits of these different arrangements, and the value produced by each.

[31] See, also, Cariolle (2019).

[32] Typical auctions today are not entirely free of obligations. For example, they often contain buildout requirements to prevent hoarding, as well as limitations on power and range. However, they often lack explicit directives over the application.

There is also pressing need for analysis of the experience outside of developed economies. For example, Bjorkegren (2019) illustrates an innovative study about wireless infrastructure examining the demand for and benefits from mobile digital infrastructure in Rwanda. Using information about a year's worth of phone calls, he provides direct estimates of the value of belonging to a network within a country, as well as of the value of the infrastructure that supports it. Here he finds considerable evidence of network effects in demand, which suggests the externalities from digital infrastructure can be substantial in a developing economy.

It is important to develop further estimates of demand for wireless access in developing countries. Unlike many other innovative products in the modern economy, all innovative digital services do not first arise in developed countries before migrating to the markets of developing countries. A set of innovative services – and new to the world! – have begun to appear in the developed world where wireless devices have become the primary tool for accessing the internet. Many fundamental economic activities, such as payments and banking, have developed atop this ubiquitous wireless infrastructure. For example, China's most popular payments application for wireless devices, WeChat, has over 1 billion users, and, remarkably, it has become the most common electronic substitute for cash. Such innovation has become essential for economic development, and merits further understanding.

III. Network Components

The internet emerged with an architecture for internetworking that had evolved with several decades of experience in its use by DARPA and the NSF. The architecture for the internet today still bears some resemblance to its predecessor. One set of firms provides access, while another provides long-distance data connectivity (i.e., backbone), and still others provide a range of additional services, such as name-serving, domain registration, and routing. Exchange of data still largely follows the models established in the late 1990s, albeit, today at a much higher level and volume of traffic. In a range of activities, however, the architecture has evolved. The number of interconnections is much larger today, and the type of files supports a set of applications that look quite different from those dominating the network traffic in the early 1990s. Succinctly, while interconnection between commercial entities follows commercial practices that emerged after privatization, a plethora of specialists perform activities with no precedent from the Internet of the 1980s, and connectivity emerges from sophisticated negotiated agreements (Huston, 2017, Norton, 2014). In a network where email and file-transfer once dominated traffic between users with fixed locations, the network evolved to add data-intensive applications – namely, high volume web browsing, video delivery, streaming, and interactive gaming, and, most recently, many applications accommodate a mobile user.

No simple model can explain this restructuring. It came about from a variety of initiatives. Broadly stated, the supply of specialized services changed in response to newly perceived opportunities, competitive incentives to improve, and new perspectives on how to organize commercial services. Unrestricted entry permitted a wide range of experiments, and firms learned from their experience, and adapted their businesses to what they learned about demand and operations. A few large firms integrated into infrastructure, while more and different specialists offered new activities. All of that restructuring supported new application development that encouraged more adoption, and more adoption encouraged restructuring to support new application development.

This restructuring is essential to the economics of digital infrastructure, and for several reasons. One, many complements (other than access) are essential to the operations of the internet and the quality of delivered services, and, by extension, their operations shapes the potential for economic growth. Two, the supply of these complements is not ubiquitous. Within the United States, the uneven supply of complements creates regional variance in the quantity and quality of digital infrastructure. Across the globe, uneven investment creates even larger variance in quantity and quality across countries.

That just begs questions: Why does such unevenness arise? More to the point, what mechanisms generate unequal supply of network components? Are these similar to the mechanisms shaping supply of infrastructure in other areas of the economy? As it turns out, many components come from private firms in (largely) (un)regulated markets. Insights into the causes of variance in outcomes from restructuring potentially leads to lessons today about how policy can shape the variance of digital infrastructure quality in the future.

Before reviewing research about network components, the discussion begins with examples as illustration. These are from supply of content delivery networks, data centers, and the cloud. All of these commercial services emerged after privatization and grew dramatically, and enabled a range of new applications. What economic factors shaped the emergence of these innovations?

III.1. Content Delivery Networks

The experience with content delivery networks (CDNs) provides an illustration of externalities that arise within an interconnected network of complementary inputs in digital infrastructure. CDNs first became available in the late 1990s, and began spreading after the millennium as firms adopted them for use. Geographically distributed networks of servers located close to users, CDNs (1) reduce data delay and response time by rerouting user requests and (2) provide a layer of reliability and security.[33] Experience taught lessons about optimal deployment. Today all but the smallest commercial participant uses them. They are, in short, an essential layer of digital infrastructure, and importantly, they arose in response to private and commercial incentives.

In the most common arrangement, a third-party commercial CDN negotiates interconnection with an ISP or wireless access provider for the right to "collocate" a server close to users. The ISP or network also may charge a "transit" fee to the CDN to take data over its network lines (i.e., from the content firm's servers to the equipment installed by the CDN). Content providers pay the CDN provider to redistribute content to users from the CDN's servers, which the content provider "updates" at an arranged schedule over the course of the day. Due to the expense, many content providers choose to update only the most timely and popular content.

While the market for CDN services has blossomed, vertical integration into use of CDNs is also common among large application firms. Some content providers, such as Google, Apple, and Facebook, operate their own CDNs and tailor the technical features of the CDN to their own needs. Again, they negotiate a price with ISPs for "collocation" in the network, and sometimes

[33] Even when servers have gone down, the cached content in a CDN may keep a firm's content available for users. In addition, CDNs can buffer content from a denial of service attack.

pay fees for data transit. In practice, only large firms opt for this action, because it is usually less expensive to contract with a third-party CDN for small to medium volumes of traffic.

The economic value of CDNs is an open topic, and observation suggests the quantitative importance should be large because the investments are large. At this point, companies have deployed thousands of servers as CDNs in the just the US alone, and long experience has taught buyers what type of improvements to expect from purchasing services. Suppliers have learned from experience how to add new services to CDNs, such as security against DDOS[34] attacks. Such investments have occurred all across the globe.

The consequences also should be large. The growth of CDNs coincided with the improvements in consumer experience, especially in lowering latency for the large data flows supporting video. When the data packets travelled to users over dial-up in the mid-1990s, email, file transfer, and web applications comprised the vast majority of traffic; and users typically could tolerate delays. Later, not only did data traffic reach users primarily through broadband lines and change to become comprised of mostly streaming, video, and gaming applications,[35] but users also no longer experienced delays. The symbiotic relationship between improvement in broadband and wireless access and applications took advantage of those improvements, notably applications on smart phones and streaming. Many new applications would have been infeasible without CDNs, such as “Over-The-Top” streaming services like YouTube, Netflix, Sling, or HBO Go. The results are visible in the changing composition of applications and the traffic that supports them.

The spread of CDNs frames provocative economic questions about the economic impact of innovation, and its links to the motivation for undertaking it. The gains appear to distribute widely, while most of the investment expenses concentrate on a few suppliers. How do the gains distribute between CDN providers that operate the servers, the content providers that experience faster delivery, the users who enjoy previously unobtainable content, and the ISPs that charge colocation fees and gain revenue from users for better services? Because only CDN providers

[34] A distributed denial of service (DDOS) attack involves large numbers of queries to a server in a short time, exceeding its capacity and rendering it unable to provide any service. CDNs are one of several instruments that can provide buffers against such attacks.
[35] See e.g., the usage statistics in Nevo et al. (2016), McManus et al. (2018), and Huston (2017).

incur the private costs and commercial risks, understanding these gains and related externalities help to understand the incentives of improving the network. As of this writing, no economic research has approached these questions.

III.2. Data Centers and the Cloud

Data centers are a specialized form of digital infrastructure. These buildings contain rows of servers on racks, which perform computation or storage. These buildings optimize for low-energy use and optimal cooling, and they may contain expensive backup generators and structures to prevent flooding or reinforcements in floors to lower vibrations from passing vehicles. The inside wiring also may support a specific set of activities,[36] especially in critical functions that support transactions with sensitive customer data. Similar to the questions that arise with CDNs, developments in data center infrastructure raise questions about externalities in an interdependent system. Because they support the rise of "cloud" services, their economics also motivates additional questions.

The origins of data centers started innocuously enough. At the outset of the commercial internet, most firms housed their servers on company premises. That changed gradually, and as it did, businesses began to support data centers.[37] Today third-party suppliers of data centers in the United States allocate assets worth at least several hundred billion dollars.[38] Data centers lower latencies for business users, enable large-scale computing and innovative uses for that scale, consolidate managerial challenges and reap efficiencies from solutions to those challenges, enable flexible uses that previously were not possible, and remove frictions to accessing big-data applications. These abilities reduce frictions to supporting applications for a mobile labor force.[39]

Contracts for data centers cover every conceivable arrangement and option between ownership and rental markets. At one extreme, rental markets arise for just about any arrangement a buyer

---

[36] The data center for the New York Stock Exchange, for example, permits many firms to access trading services at especially fast rates. As another example, a segment of business users in health, finance, and transportation require high security and high reliability—often referred to as the *five nines* of reliability, namely, 99.999% uptime.
[37] See, e.g., Jin and McElheran (2017), or Byrne et al. (2018).
[38] Greenstein and Pan Fang (2019).
[39] See e.g., Ewens et al. (2019), and DeStafano, Kneller, and Timmins, 2019.

could want. There are plenty of firms that will take full responsibility for the operations of the building and electronic equipment for a service fee. Many buyers with generic needs—such as storage for backup—rent space in data centers at various time intervals (e.g., five, ten year, or twenty years), own the servers and program them, and let others manage the building. At the other extreme, firms with unique computing needs, such as Facebook, Apple, Microsoft, Amazon, Oracle, and Google, own everything. They operate the largest private data centers in North America and configure the building and servers to suit their applications.

Today a cloud service involves a data center that rents its services, with the additional feature that users can request any size and turn the service off and on at will. The major cloud providers also increasingly offer additional software services for a nominal charge or for none at all. The demand for cloud services has grown as the services improve in quality and declined in price.[40] The appeal of the cloud comes from its flexibility in capital commitment, and the option (for business users) to substitute variable costs for fixed costs on a balance sheet, which has strong appeal to cash-constrained entrepreneurial firms. Sophisticated firms increasingly utilize complex architecture to balance the loads from user demands, such as using CDNs for rapid response to requests for timely content, cloud facilities for secondary response, and remote servers for requests of the least popular content.

Just as with CDNs, the growth of data centers and the cloud illustrates an important economic question about the impact of investment in frontier digital infrastructure. How do the gains distribute between cloud providers who operate the servers, the content providers who use them, and the users who enjoy previously unobtainable content? Due to such externalities, the private incentives for improvement appear to be lower than the wider gains.

The data center and cloud market has received some attention. In the first paper on its productivity, Jin and McElheran (2018) examine use of cloud computing in US manufacturing, and find it predicts productivity growth among young firms and new units in established firms. Use of the cloud also predicts productivity, conditional on survival, in uncertain environments.

[40] Estimates of its market share depend on the definition of sales. See the discussion in Byrne et al. (2018).

The evidence is consistent with the highest gains accruing to firms who take advantage of the flexibility and lower costs of learning about IT needs in spite of uncertainty.

Tensions between the size of the investment and localization of demand shapes the location of data centers. Greenstein and Pan Fang (2019) posit a framework that focuses on the tension between the "distaste for distance," [41] which creates localization of demand, and different supply conditions across geography, causing variance in the costs of meeting local demand. That leads to facilities spreading out to match local demand. These compete with facilities that "aggregate" the demand from many locations. The costs of supply reflect variance in two related factors, namely, economies of scale and variance in operational costs. Both fixed and variable costs vary with cost of inputs, such as land, electricity, cooling, and technical labor. These lead to variance in costs across different locations, and firms respond to this tension with entry and capacity decisions. They forecast a "minimum threshold" of local users under which no entry occurs, and find evidence consistent with this model. That suggests data centers and cloud services have an urban bias, favoring bigger and denser cities.

An open economic question concerns the future localization of demand. If buyers perceive shorter distances – between users and the data centers for cloud services – as an important attribute of cloud services, then that will create further potential for tension around the localization of supply. The first evidence about demand for cloud services suggests users will place value on distance (Wang et al., 2019). While ubiquitous frontier infrastructure confers large societal benefits, such frontier infrastructure tends not to be available in low-density regions or in areas with a concentration of low-income households.

III.3. Open questions about components

As noted, uneven geographic supply of infrastructure creates regional variance in the impact of complementary digital infrastructure. The uneven supply arises for many reasons. Among them,

[41] It arises from a mix of three factors. The first two—user dislike for latency and user desire to avoid congestion—look alike in reducing distances between users and facilities. A third factor, "server hugging," arises from managerial preferences for nearby physical facilities, which facilitates monitoring.

the costs of supply may reflect economies of scale, such as in installing and operating CDNs, cell towers, and data centers. These structures endogenously exist near a greater number of densely located users. The demand for higher quality can also drive unequal dispersion, and firms prefer to initially buildout in more affluent and urban locations where a greater number of buyers are more willing to pay for the expensive frontier quality. Or, as noted with data centers, Marshallian agglomeration may further reinforce the biases towards urban locations.[42]

Two distinct views animate open questions about the geography of digital infrastructure supply. A somewhat optimistic outlook anticipates more diffusion to more users, more regions, and a greater set of applications. It interprets the state of digital infrastructure at a point in time as temporary and transient, in the midst of diffusion, and not settled. According to this outlook, productivity differences between users melt over time as once-expensive infrastructure, which incubated in a few cities, spreads to new users and new locations. The most important open question concerns the determinants of the speed of diffusion, which then determines how fast laggard regions and buyers catch up to frontier regions and buyers. In this view, policy focuses on speeding up diffusion by removing deterrence to adoption, or subsidizing emergence of more supply.

A more pessimistic outlook stresses urban-biased technical change. In this view, modern digital infrastructure will achieve higher productivity in dense locations due to economies of scale in equipment, increased productivity from colocation, and availability of skilled labor in urban areas. In this view, large differences may persist, and in the most pessimistic view, only a few locations enjoy the benefits of the frontier.

On the surface, the experience with CDNs supports the optimistic view, while the experience with data centers supports the less optimistic view. More detail and scrutiny is required, however, before either view has solid factual support. Research can resolve some of the tension between these two outlooks. For example, though many users prefer local supply of infrastructure when it is available, it may be possible to use remote data centers, cloud storage and/or satellites instead. Similar to the tradeoff between satellites and wireline broadband supply,

[42] See the longer discussion of Forman et al. (2018).

feasible subsidies and user preferences frame the question regarding the tradeoff between a local supply of infrastructure and remote data centers, cloud storage, and/or satellites. These questions would be, and could be, informed by estimates of the elasticity of demand.

Policy analysis needs more than just estimates of demand, however, and this observation comes face to face with the biggest challenge for this research, namely, translating tradeoffs into pecuniary terms. Governments try to employ infrastructure to have an impact on non-economic outcomes, such as informing a citizen's knowledge, furthering the education of children, or contributing to the public health of a local population. How much does society want to spend for those goals to encourage, say, supply of wireless broadband or improvement in wireline transmission speeds where it would otherwise not arise from market forces? How much should it pay to build out the internet in low-density places to organizations with public missions, such as libraries, schools, and hospitals? As with the demand for many public goods, it is naïve to presume an easy answer. With a frequently moving frontier, the public policy issues are especially vexing.

IV. Governance and Protocols

In addition to funding the building of prototypes for backbone and complementary hardware, government research and development (R&D) helped create the protocols that support the internet. First, it directly helped subsidize the research that created the protocols and the institutions that support them. Second, it made a number of investments that contributed to the use of, and network effects around, those protocols. Third, it funded the testbed for many early users in universities. That is, it funded much of the science that made use of the early internet inside many universities and research labs that acted as "lead users" for these protocols.

Out of this long incubation came a set of *protocols* and *protocol stacks*, intended to make digital equipment universally compatible.[43] Although complex, the protocol stack design for sending

[43] Protocols are the set of rules and regulations that determines how data makes it through the network. A networking protocol defines conventions for processes, which includes definitions for both the format of data packets, and as well as for recovery in the event of transmission errors. *Protocol stacks* are comprised of a family of related protocols assembled together, and they act as a reference model for designers, who largely aspire to make

data packets along the least congested route is an essential feature of today's digital infrastructure that relies on fast delivery of data for applications. Its importance leads to an economic question: what is the economic impact from improving protocols? That question is open, in part because every user and supplier has access to the same protocols, resulting in no meaningful variance in adoption across the globe with which to make estimates of impact. Moreover, such non-rivalry in use, combined with *de facto* lack of excludability in most of the institutions that support development of global protocols, gives software protocols a set of properties isomorphic to classic public goods. Long experience has taught that, despite its importance, estimating the value of public goods from a government-supplied source is quite challenging. Estimating demand is no easier just because a consortium of privately financed standard setting organizations designed it.

Protocol development today does not reside exclusively with governments. Several non-profit organizations design and upgrade the protocol stack used for global internet infrastructure. The governance of these protocols has enormous consequences, and accordingly, many stakeholders contribute to improvements. For example, the Internet Society oversees the Internet Engineering Task Force, which designs protocols behind TCP/IP (Transmission Control Protocol/Internet Protocol), BGP (Border Gateway Protocol), and other protocol stacks. The Internet Society and other organizations[44] subsequently charge little for their use.

The basic economics suggests that improvements in protocols could confer quite large gains to users. That motivates questions about the factors that determine the (in)efficacy of these development processes. An important example of such research is Simcoe (2012), which examined the speed with which the IETF generated new protocols for the internet before and after privatization. He traces variance in speed to its underlying determinants, such as the composition of the committees making new protocols. His research focuses attention on the role of disagreements between participants with varying interests. It stresses the importance of multi-

---

compatible equipment. For longer descriptions, see, e.g., Clark (2018), Kneips and Bauer (2016), or Greenstein (2015).

[44] The Institute of Electronic and Electrical Engineers (IEEE) maintains 802.11, the standard underlying Wi-Fi, as well as other technical standards. The Internet Corporation for Assigned Names and Numbers (ICANN) governs assignment of domain names and updates the routing tables used by every switch and router on the internet. A routing table contains information about the topology of a network. See e.g., Clark (2018).

stakeholder institutions that (do or do not) become sclerotic as they admit more participants. It frames the question about whether such process necessarily cannot avoid becoming slower as their designs touch a wider breadth of the economy, resulting in a greater breadth of voices developing conflicting stakes in the details of protocol improvement.

An open avenue for research explores the incentives for and gains from improvement in protocols, as designed by quasi-public organizations. How large are the shared benefits from improving protocols? More broadly, today a mix of publicly subsidized and privately funded research finances protocol development. How large are the contributions from members in relation to those benefits? Unlike the past, government users no longer acts as the major test bed for protocol development. Whose experience has the most salience for the direction of improvement?

One approach to these issues focuses on estimating the existence and size of the externalities from deployment of a key piece of infrastructure. For example, Nagle (2019) takes a novel approach to this topic by examining a set of externalities in protocol improvement that *were not* global. Counter-intuitively, he makes progress by focusing on quite the opposite, externalities from software in which the spillovers were particularly *localized* in scope. He focuses on the spillovers from a French government mandate to use Linux, a program adopted as part of general policies to encourage use of open source software. He finds it had consequences for the rate of new business formation in complementary digital areas. The analysis takes advantage of a natural placebo test in events, in which the Italian government did not enforce a similar decree within its own borders. Nagle's estimates suggest that the externalities can be substantial if governments enforce their policies. It frames a big open question: what conditions lead local supply of talent to respond, and what limits that response?

While the gains from improvement in protocols could spread across many, the private costs and commercial risks often are concentrated. That motivates another open research topic, about the size and incidence of the externalities from improvement in digital infrastructure and the gaps between private incentives to innovate and society returns. This brings us back to an important topic, how government R&D policy focuses funding on areas where these gaps are largest.

One approach to this topic examines the recent past, in which, with the benefit of hindsight, the economics were comparatively simple. Namely, the costs of R&D were defrayed against the benefits affiliated with meeting the mission of a federal agency (i.e., at DARPA and NSF). While these costs were concentrated, the external benefits to society were widely shared. That sets up a question: what were the economic gains from the public investments in the historical R&D that supported protocol development? As one approach, Greenstein and Nagle (2014) employ a method for estimating the value of unmeasured web servers in the United States in 2011. They show that these inputs make a positive contribution to economic growth in the United States. They further show that the returns from web servers alone generated enough economic gains to equal the U.S. government's R&D internet investment. That is an important conclusion, because they do not make a full account for all gains from the invention of the internet (and which is still an open question).

IV.1. Research at the boundaries

Although most observers agree that internet access services are part of digital infrastructure, many other specialty software are less unanimously categorized as such, even though these are necessary for delivering any service that relies on the internet. The governance of this ubiquitous software requires attention, though such activity falls far outside the scope of (traditionally) regulated markets or government-supplied services.

Concrete examples can illustrate the open questions. Contrast two starkly different models. Some privately supplied software has achieved ubiquitous use, such as the Microsoft Operating System, and Oracle Server, and Android/iPhone smart phone operating systems. Private firms supply this software, upgrade it, service requests, and exclude those who fail to pay an appropriate price. Another model also yields ubiquitous software. The World Wide Web Consortium is one such example. Managed by a non-for-profit consortium, this software experiences regular upgrades, and achieves ubiquity through non-exclusion, making each upgrade available without restriction. The continuing success of the Web illustrates a model that leads to widespread use and high impact. What economic factors lead to a good match between these governance models and market settings? How much difference does the governance model make to outcomes?

Webserver software raises similar questions, and offers different insights. Different users today largely employ three different servers: Apache, IIS, and Ngnix. The first one descended from earliest experiments with web servers at the University of Illinois, organized as an open source project. Microsoft offers IIS, generally as part of a range of the enterprise software it offers and certifies. The third, Ngnix, comes in a freemium form today, with a fast but limited version available for nothing. An enterprise version is available with payment for additional services. Apache and IIS had a large impact on the market in the first two decades of the Web, but Ngnix has recently enabled large gains in high-volume servers, which are essential for streaming. The tradeoffs between each of these organizational forms defy easy characterization. It is unclear how to characterize the impact of innovations to each of them.

A range of questions about the governance and ownership of software arise across the globe, and research studies today are just beginning to make progress into analyzing the determinants of investment in complementary infrastructure. For example, Athey and Stern (2014) ask why some countries use more pirated operating system software. Their framework contrasts two broad determinants: (1) variation in willingness to pay for software, which shapes economic incentives to pirate software, and (2) institutional enforcement of property rights, which shaped incentives for private actors to invest in software. Athey and Stern measure the former with variables, such as per capita income, while they measure the latter with variables, such as the country-specific history of respecting property rights. If the former is important, then sellers of proprietary software could potentially change their pricing strategies in settings where open source appeals to those who want to avoid expenses. If the latter is important, then pricing is unlikely to address the challenge, and sellers of proprietary software would have to work in different directions.

Their framework provides a pathway forward. Two differences between operating systems and other internet infrastructure potentially shape the economics of other digital infrastructure. In most settings, infrastructure must be available for continuous operations and compatible with other parts of the internet. Continuous operation and compatibility requires the range of complementary operations mentioned previously. Considerable data exists to measure variance across the globe in these complementary activities and supporting institutions (OECD, 2014, WEF, 2016).

An altogether different and important insight comes from research focused on now-casting in developing countries, i.e., using present economic activity to forecast events over a short time horizon, particularly, where GDP measurement apparatus is absent or primitive. Near ubiquitous digital infrastructure can offer a way forward in measurement. For example, Indaco (2018) uses Twitter activity (as measured thru GPS-labeled photos) to determine if geo-located IP addresses give as much information as light from satellite photos. The study correlates Twitter use with other measures of economic activity, such as the light from satellite photos, because the same types of advanced investments support both—namely, continuous electrical supply, skilled labor, and a range of complementary investments. Ackermann et al (2014) provide a similar exercise when they examine the distribution of IP addresses. Once again, this provides evidence of economic activity.

IV.2. Open questions about governance

As the above examples hint, the boundary between public and private is in flux across a wide set of activities. Some software is private, some is open source, and some employs a mixed model. Some software comes from consortia, others from standard setting organizations, and still others from private suppliers. Government policy plays a variety of roles – for example, in subsidizing research and invention, in workforce training in higher education, in providing some services, and in defining legal boundaries for different types of organizations.

Mapping software offers an example of the new frontier. Once thought to be solely a government function, digital mapping has passed to either proprietary or open source projects, which draw input from crowds, and these compete with one another. These platforms vary in their governance and source of input, as well as in response to new opportunities.[45] The next

[45] See e.g., Nagaraj (2017), and Nagaraj and Piezunka (2018).

generation of mapping for autonomous vehicles has entirely moved to private sources. Different firms use different models of how to use input from crowds. At the same time, all mapping depends critically on government-funded satellites that provide GPS (Global Positioning System) coordinates, so public support is never far away. Are the incentives to develop digital maps too low or too high, and do they result in too few or too many development projects? What are the incentives to share results, once they are developed?

The definition of infrastructure remains fluid and difficult to pin down in widely used software tools as well. For example, consider software repositories, such as Github, which has become so common, it is essential infrastructure for many software projects. Github aids sharing of code and reduced frictions in large-scale projects. Making collaboration across distance easier, Github's creation had a well-known productivity impact; and Microsoft recently purchased the entire platform for $7.5 billion in stock. How Microsoft's purchase shapes Github's productive impact remains an open question.

Next, consider this provocative example: Is Wikipedia digital infrastructure? Its ubiquity suggests it ought to be treated as such. It receives more than 15 billion pages views per month.[46] At the time of this writing, over 5.9 million articles grace its web pages in English alone, with more than 500 new articles added each day. Volunteers built the entire corpus of text. More to the point, due to its non-for-profit status and aspiration toward a neutral point of view, miniscule storage and transmission costs, the scale economies appear virtually limitless. It has become a focal site on which many others depend, including many search engines and Q&A sites. Many software firms also use it to complement their documentation efforts on Github, providing longer explanations and links. The Wikipedia example epitomizes the many open questions of this topic: What is and what is not infrastructure? Where are the boundaries, and what questions must we ask about them? Many basic questions about the impact of digital infrastructure contain considerable room for additional investigations, particularly research aimed at measuring the impact of frontier digital infrastructure on economic activity.

[46] https://stats.wikimedia.org/v2/#/all-projects

To finish this conversation, note that Wikipedia remains unavailable in China, where the government firewall blocks the site at the border. That example illustrates the possibility of the "splintering" of the internet, that is, the erosion of compatibility of complementary equipment and software, resulting in distinct regions of the globe pursuing their own direction of technical developments, each internally consistent within national boundaries, yet, inconsistent and incompatible across borders. Splintering has begun to happen at the application layer of the internet as different governments censor content, and impose limits on the operations of applications consistent with local preferences for privacy, security, copyright, and other government policy. Some of these policies have begun to migrate into complementary digital infrastructure, where governments impose, for example, distinctly different privacy and security policies on packet-inspection in routers and back-door design within operating systems. As governments increasingly do not accept voluntary technical standards without additional modification, it frames questions about the rate, direction, and consequence of such actions for seamless interoperability.

V. Conclusion

Long before it spread across the globe, it was fashionable to call the internet an "information superhighway." The label arose out of a combination of observations and aspirations. The *observation* contained a grain of truth about the physical layout of the internet. Many backbone lines followed existing rights-of-way for roads, bridges, and highways. The *aspiration* channeled a proposed vision for the future, one where government subsidized the capital expenditure and left the assets unpriced, as in, say, a freeway. That sought to advance an ideal in which information remained free to use, i.e., unpriced and subsidized by government support.[47]

With the benefit of several decades of hindsight, we can see that both the observation and aspiration bear only partial resemblance to the present state of commercial digital infrastructure. There is considerable insight and shortfall inherent in that comparison.

[47] The aspiration became associated with a range of policy initiatives that subsidized the internet for research in the late 1980s, and, eventually, it became associated with the specific aspirations of presidential candidate, Al Gore. See, e.g., Greenstein (2015), Pages 65-70, for a discussion of these policies.

Begin with insight. Like much infrastructure, in the last few decades economic actors shared the use of the long-lasting capital of digital infrastructure, and many economic actors employed it as an intermediate input in the production of goods and services. As an intermediate input, it acted much like a new road connecting two areas with previously poor connections, lowering frictions between potential transactions in different locations (Goldfarb and Tucker, 2018). Lowering of frictions created two distinct types of new opportunities, either fostering cooperative agreement between suppliers of complementary inputs, or encouraging competition from supplier who serve new customers in new areas. Hence, new applications emerged that otherwise would not have developed or deployed in the absence of the low-cost and reliable network infrastructure.

The metaphor goes only so far in illuminating the economics, however. The building of digital infrastructure was not a one-time event, and the continued improvement accounted for a salient feature of recent experience. The contrast with roads and highways could not be sharper: Roads typically do not undergo improvements in all their key attributes every few years. In digital infrastructure, however, innovations did change frequently, and those changes accumulated, producing a whole architecture for the system that no central planner or brilliant designer could have specified in advance. New digital infrastructure supported frequent reassessments, and heightened disagreements among distinct views about how to make valuable use of opportunities enabled by improved infrastructure. That enabled "innovation from the edges" (Greenstein, 2015), and raised the importance of bringing applications to market to settle question about whether value existed, and, if so, where the highest value lay.

More pointedly, digital infrastructure *does not* resemble roads and highways in its pricing or governance. Building and operating roads and highways are largely government functions, and, relatedly, most highways and surface streets remain unpriced and non-excluded, with the possible exception of several hundred toll roads and bridges. In contrast, private funding lay behind investment in the vast majority of digital infrastructure. Regulators compel providers to offer specific services in specific territories in some local areas, but nothing on the scale associated with roads, and not with the same level of tax subsidy. Moreover, availability of digital infrastructure alone does not guarantee use, because failure to pay the minimal price leads to denying service to users, and despite the societal importance of fostering widespread use and adoption of frontier services, providers often have unfettered discretion over price and other aspects of service.

More to the point (of this essay), modern suppliers face minimal mandates to become ubiquitous, reliable, and inexpensive beyond what market forces incentivize them to build. As of this writing, variance in the supply and use of innovative digital infrastructure arises within every developed country, as well as between developed and developing countries. If the past is prolog, then much of the modern economic experience depends on whether that variance will persist, grow, or decline, and the persistence of that variance sets the table for a large number of open research questions.

References

Ackermann, Klaus, and SD. Angus (2014). "A resource efficient big data analysis method for the social sciences: the case of global IP activity." *Procedia Computer Science* 29: 2360-2369.

Athey, S, and S Stern. 2015. The Nature and Incidence of Software Piracy: Evidence from Windows. A Goldfarb, S Greenstein, and C Tucker, Eds., *Economic Analysis of the Digital Economy,* U of Chicago Press.

Bjorkegren, Daniel (2019), "The Adoption of Network Goods: Evidence from the Spread of Mobile Phones in Rwanda" Review of Economic Studies, 2019

Boik, Andre (2017), The economics of universal service: An analysis of entry subsidies for high speed broadband, *Information Economics and Policy,* 40, 13-20

Boik, Andre, Shane Greenstein, and Jeff Prince (2019), "The Persistence of Broadband User Behavior: Implications for Universal Service and Competition Policy." *Telecommunications Policy.*

Brynjolfsson, Erik, and Joo Hee Oh (2012), "The Attention Economy: Measuring the Value of Free Digital Services on the Internet," 33rd International Conference on Information Systems.

Brynjolfsson Erik Brynjolfsson, Avinash Collis, and Felix Eggers, 2019, "Using Massive Online choice Experiments to measure changes in Well Being." PNAS. April 9, 2019 116 (15) 7250-7255

Burnham, Brad, Shane Greenstein, Neil Hunt, Kevin McElearney, Marc Morial, Dennis Roberson, Charles Slocum, (2013). *Issues in Data Caps and Usage Based Pricing,* A report of the Economic Impact Working Group, Open Internet Advisory Committee. May 5. http://transition.fcc.gov/cgb/oiac/Economic-Impacts.pdf

Byrne, David, Carol Corrado, Dan Sichel (2018). The Rise of Cloud Computing: Minding Your P's, Q's and K's NBER 25188

Byrne, David, and Carol Corrado (2019), Accounting for innovation in Consumer Digital Services: IT still matters.

Cardona, M., T. Kretschmer, and T. Strobel (2013), "ICT and Productivity: conclusions from the empirical literature." *Information Economics and Policy.* 25(3), pp. 109-125.

Cariolle, Joel (2019). "Telecommunications Submarine-Cable Deployment and the Digital Divide in Sub-Saharan Africa," CESifo Working Paper 7415. Available at SSRN, https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3338769##.

Chen, Yongmin and Savage, Scott, "The Effects of Competition on the Price for Cable Modem Internet Access," *Review of Economics and Statistics.*

Clark, David (2018), *Designing an Internet.* MIT Press.

Clark, David, Steven Bauer, William Lehr, kc Claffy, Amogh Dhamdhere, Bradley Huffaker, Matthew Luckie (2014). "Measurement and Analysis of Internet Interconnection and Congestion," *Journal of Information Policy.*

Connolly, Michelle, and JE Prieger, (2013). A basic analysis of entry and exit in the US broadband market, 2005–2008, *Review of Network Economics* 12 (3), 229-270

Destafano, Tim, Richard Kneller, and Jonathan Timmins (2018). "Broadband Infrastructure, ICT use and firm performance: Evidence for UK Firms," *Journal of Economic Behavior & Organization*, 155(C), pp 110-139.

Destafano, Tim, Richard Kneller, and Jonathan Timmins (2019). "Cloud Computing and Firm Growth." *Working Paper. Presented at NBER Winter Digitization Meetings.*

Downes, Thomas, and Shane Greenstein (2002), "Universal Access and Local Commercial Internet Markets," *Research Policy*, 31, pp 1035-1052.

Ewens, Michael, Ramana Nanda and Matthew Rhodes-Kropf (2019). "Cost of Experimentation and the Evolution of Venture Capital" Harvard Business School Working Paper, No. 15-070, Journal of Financial Economics, forthcoming.

Fernandes, Ana, Aaditya Mattoo, Huy Nguyen, Marc Shiffbauer (2019), "The Internet and Chinese Experts in the pre-Ali Baba Era." *Journal of Development Economics.* 138. Pp. 57-76.

Forman, Chris (2005), "The Corporate Digital Divide: Determinants of Internet Adoption," *Management Science* 51(4): 641-654.

Forman, Chris, Avi Goldfarb, Shane Greenstein (2005) "How did location affect adoption of the commercial Internet? Global village vs. urban leadership," *Journal of Urban Economics.* 58(3), pp. 389-420.

Forman, Chris, Avi Goldfarb, Shane Greenstein (2012), "The Internet and Local Wages: A Puzzle," *American Economic Review.* February. 102(1), pp. 556-575.

Forman, Chris, Avi Goldfarb, Shane Greenstein (2018). "How Geography Shapes – and is Shaped by – the Internet." in (ed.) Gordon Clark, MaryAnn Feldman, Meric Gertler, and Dariusz Wojcik, *The New Oxford Handbook of Economic Geography*, Oxford University Press. Pp. 269-285.

Goldfarb, Avi, and Jeff, Prince (2008). "Internet Adoption and Usage Patterns Are Different: Implications for the Digital Divide." *Information Economics and Policy*, 20, 1, 2-15.

Goldfarb, Avi, and Catherine Tucker (2019). "Digital Economics," *Journal of Economic Literature*, 57 (1).

Goolsbee, Austan, and Peter Klenow (2006), "Valuing Consumer Products by the Time Spent Using Them: An Application to the Internet." *American Economic Review*, 96(2), pp. 108-113.

Greenstein, Shane (2015). *How the Internet Became Commercial: Innovation, Privatization, and the Birth of a new Network*, Princeton University Press.

Greenstein, Shane (2020), "The Basic Economics of Internet Infrastructure," *Journal of Economic Perspectives.*

Greenstein, Shane, Martin Peitz and Tomasso Valleti (2016). "Net Neutrality: A Fastlane to Understanding the tradeoffs." *Journal of Economic Perspectives.* 30(2). Pp 127-150.

Greenstein, S., & R. McDevitt (2011). "The Broadband Bonus: Estimating Broadband Internet's Economic Value," *Telecommunications Policy*, 35, pp 617-632.

Greenstein, S, and F Nagle (2014). "Digital Dark Matter and the Economic Contribution of Apache," *Research Policy*. v43. Pp 623-631.

Greenstein and Pan Fang (2019), "Where the Cloud Rests: the Economic Geography of Data Centers," Working paper, Harvard Business School.

Haller, Stephanie, and Sean Lyons (2015). "Broadband Adoption and Firm Productivity: Evidence from Irish Manufacturing Firms." *Telecommunications Policy.* Pp. 1-15.

Haller, Stephanie, and Sean Lyons (2019). "Effects of Broadband Availability on Total Factor Productivity in Service Sector Firms: Evidence from Ireland." *Telecommunications Policy*, pp. 11-22.

Hitt and Tambe (2007) "Broadband Adoption and Content Consumption. *Information Economics and Policy*, 19, 3-4, 362-378.

Hjort, Jonas, and Jonas Poulsen (2019). The Arrival of Fast Internet and Employment in Africa, *American Economic Review*, vol. 109, no. 3, 1032-1079, March 2019

Huston, Geoff (2017). The rise and the rise of Content Distribution Networks, Workshop on Information Economics, http://www.caida.org/workshops/wie/1712/slides/wie1712_ghuston.pdf, CAIDA, UC San Diego, Accessed April 16, 2019.

Indaco, Augustin (2019), "From Twitter to GDP: Estimating Economic Activity from Social Media," http://aindaco.com/.

Jin, and Kristina McElheran (2018), "Economies before Scale: Learning, Survival, and Performance of Young Plants in the Age of Cloud Computing" (with Wang Jin). Nominated for best paper, Strategic Management Society Special Conference, Oslo. SSRN working paper series: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3112901

Knieps, G and JM Bauer (2016). The industrial organization of the Internet, (editors) M Latzer and JM Bauer, *Handbook on the Economics of the Internet*, 23-54, Edward Elgar Publishing.

Kolko, Jed (2012), Broadband and Local Growth, *Journal of Urban Growth.* 71(1), pp. 100-113.

Malone, Jacob and Nevo, Aviv and Williams, Jonathan W. (2017), The Tragedy of the Last Mile: Economic Solutions to Congestion in Broadband Networks (June 6, 2017). NET Institute Working Paper No. 16-20. Available at SSRN: https://ssrn.com/abstract=2849869

McElheran, Kristina (2015). "Do Market Leaders Lead in Business Process Innovation? The Case(s) of E-Business Adoption," *Management Science*, 61(6), June 2015, pp. 1197-1216.

McManus, Brian, Aviv Nevo, Zachary Nolan, and Jonathan W. Williams (2018). Steering Incentives and Bundling Practices in the Telecommunications Industry (October 1). NET Institute Working Paper No. 18-12. https://ssrn.com/abstract=3267060

Nagaraj, Abhishek (2017), "Information seeding and knowledge production in online communities: evidence from open street map." http://abhishekn.com/files/nagaraj_seeding.pdf.

Nagaraj, Abhishek, and Henning Piezunka, 2018, "Deterring the New, Motivating the Established— The Divergent Effect of Platform Competition on Member Contributions in Digital Mapping Communities." http://abhishekn.com/files/nagaraj_piezunka_competition.pdf

Nagle, Frank (2019). "Government Technology Policy, Social Value, and National Competitiveness." Harvard Business School Working Paper, No. 19-103, March 2019.

Nevo, Aviv, John Turner and Jonathan Williams (2016). "Usage Based Pricing and Demand for Residential Broadband," *Econometrica*, 84(2), March, 411- 443

Norton, William B (2014). The Internet Peering Playbook: Connecting to the Core of the Internet. DrPeering Press.

Nuechterlein, Jonathan, and Phillip Weiser (2005). *Digital Crossroads: Telecommunications Law and Economics in the Internet Age.* MIT Press.

OECD (2014), *Measuring the Digital Economy, A New Perspective,* OECD Publishing, Paris. https://doi.org/10.1787/9789264221796-en

Poliquin, Chris (2018), The effect of the Internet on wages, Working paper, UCLA Anderson School, https://poliquin.xyz/files/poliquin_jmp.pdf.

Rosston, Gregory L., Scott J. Wallsten (2019). "Increasing Low-Income Broadband Adoption through Private Incentives," Technology Policy Institute. Washington DC.

Rosston, Gregory, Scott Savage, and Scott J. Wallsten (2010) "Household Demand for Broadband Internet Service in 2010," B. E. Journal of Economic Analysis & Policy: Advances, Vol 10, No 1.

Seamans, Robert (2012). Fighting City Hall: Entry Deterrence and Technology Deployment in the Cable TV Industry. *Management Science*. 58(3): 461-475

Simcoe, Timothy (2012). Standard Setting Committees: Consensus Governance for Shared Technology Platforms, *American Economic Review*. 102 (1), pp. 305-36.

Skiti, Tedi (2019), "Strategic Technology Adoption and Entry Deterrence in Broadband," *Industrial and Corporate Change.*

Wallsten, Scott, and Colleen Mallahan (2013) Residential Broadband Competition in the United States, in (eds.) Avi Goldfarb, Shane Greenstein, and Catherine Tucker, *The Economics of Digitization*, Edward Elgar; Cheltenham, UK.

Wallsten, Scott, and James Riso (2010), "Residential and Business Broadband Prices, Part 2: International Comparisons." Technology Policy Institute Working Paper. December.

https://techpolicyinstitute.org/policy_paper/residential-and-business-broadband-prices-part-2-international-comparisons/.

Wang, Shuang, Jacob LaRiviere, and Aadharsh Kannan, 2019, "Spatial Competition and Missing Data: An Application to Cloud Computing." *Working Paper. Presented at NBER Winter Digitization Meetings.*

Whitacre, Brian, Roberto Gallardo, and Sharon Strover (2014), "Broadband's Contribution to Economic Growth in Rural Areas: Moving Towards a Causal Relationship," *Telecommunications Policy*, 38(11), pp. 1011-1023.

World Economic Forum, 2016, The Global Information Technology Report, https://www.weforum.org/reports/the-global-information-technology-report-2016.

.