



Smith ScholarWorks

Government: Faculty Publications

Government

2-1-2022

A New Measure of Digital Economic Activity and its Impact on Local Opportunity

Karen Mossberger

Arizona State University

Scott LaCombe

Smith College, slacombe@smith.edu

Caroline J. Tolbert

University of Iowa

Follow this and additional works at: https://scholarworks.smith.edu/gov_facpubs

Part of the Political Science Commons

Recommended Citation

Mossberger, Karen; LaCombe, Scott; and Tolbert, Caroline J., "A New Measure of Digital Economic Activity and its Impact on Local Opportunity" (2022). Government: Faculty Publications, Smith College, Northampton, MA.

https://scholarworks.smith.edu/gov_facpubs/45

This Article has been accepted for inclusion in Government: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact scholarworks@smith.edu



A new measure of digital economic activity and its impact on local opportunity[☆]

Karen Mossberger ^{a,*}, Scott LaCombe ^b, Caroline J. Tolbert ^c

^a Arizona State University, USA

^b Smith College, USA

^c University of Iowa, USA



ARTICLE INFO

Keywords:

Broadband
Domain name hosts
Local economic development
Microbusinesses
Measuring broadband use
Community impacts

ABSTRACT

Online businesses and platform work can create the impression that the digital economy is ephemeral and placeless. But the digital economy is experienced locally, and its effects are spatial. Measuring them requires better community-level data on economic activities online. While new government data measures broadband subscriptions down to neighborhoods, existing public data do not measure how broadband is used in local communities, and whether this digital activity affects economic outcomes. We analyze new monthly data on over 20 million domain name hosts/websites in the United States from November 2018 to November 2020 drawing on customer data. Surveys show that 3 out of 4 of these domains are commercial, including microbusinesses as well as websites for both online and brick-and-mortar establishments. How is the density of domain name hosts in a community (the number in a zip code or county divided by the population) related to local economic opportunity, controlling for other known factors? Using statistical matching and time series data, results show the density of domain name hosts positively predicts community economic prosperity, recovery from the 2008 recession, and change in median income. Interactions between the density of these hosts and broadband subscriptions also predict lower monthly unemployment rates over time, including after the March 2020 pandemic. Commercial data can improve our understanding of broadband's impacts, including its potential for inclusive growth in diverse communities.

1. Introduction

Life moved online at breakneck pace during spring of 2020, with schools and businesses closing their doors and households sheltering in place in the wake of the Covid-19 epidemic. Yet, persistent disparities in internet access and use meant that some individuals, businesses, and communities were unprepared for survival in this new environment. The crisis clearly demonstrated the extent to which some places lack widespread broadband connectivity, despite an increasingly digital economy prior to the pandemic.

The digital economy at the local scale is often measured by the share of employment in information technology, or by venture

[☆] De-identified data and funding for analysis were provided by GoDaddy. The academic research team independently determined all research questions, methods, and interpretation of the findings.

* Corresponding author. Arizona State University, School of Public Affairs Arizona State University, 411 N. Central Ave., Suite 450, Phoenix, AZ, 85004, USA.

E-mail address: karen.mossberger@asu.edu (K. Mossberger).

capital investments in IT (see [Jackson et al., 2019](#); [Moretti, 2012](#)). These measures focus on the IT industry rather than applications throughout the economy, including for small businesses, microbusinesses, or gig workers. Online businesses and platformed work can create the impression that the digital economy is ephemeral and placeless. But the digital economy is experienced locally, and its effects are spatial. There is a need to empirically determine how online economic activities are distributed across communities, and their impacts on local economies. This study provides a new measure of digital economic activity at the local level (for zip codes and counties) – the density of domain name hosts per capita – to investigate the impacts of this activity more generally as well as during Covid-19.

In the aftermath of the pandemic, new public investments in broadband infrastructure and its use are expected to contribute to economic recovery as well as other policy goals ([Taglang, 2021](#)). There is indeed evidence that broadband adoption has had positive economic impacts for communities in the past ([Lehr et al., 2006](#); [Whitacre et al., 2014](#); [Gallardo et al., 2020](#)). But there has been a lack of sufficiently granular data to understand how that technology is employed by businesses and residents at the local level ([National Telecommunications and Information Administration/National Science Foundation, 2017](#), 3). Measuring digital microbusinesses and economic activities provides a more accurate view of how a general-purpose technology such as broadband is being used and what outcomes might be expected from these uses.

National surveys lack sample sizes sufficient to measure activities online for small geographies such as counties and zip codes. Individual and household-level data on uses of the internet – for work, commerce, job search, health information, education and more – are available at the national or state level through sources such as Pew Research Center and the US Bureau of the Census Current Population Survey supplements. Specific digital activities, for example online news consumption, are better predictors of political participation than general measures of internet access, according to one meta-analysis of early internet and politics research ([Boulian, 2009](#)). Similarly, understanding how broadband is used for economic activity may better explain differences in local economies. Yet there is a lack of national survey data for measuring online activity across US counties and zip codes.

Data on 20 million “ventures” in the United States can help to fill this gap. These data are active domain name hosts (websites/ssl/email) and their redirects, representing over half of all U.S. domain name hosts, available monthly and in real time.¹ Through collaboration with GoDaddy, the world’s largest registrar of domain names, we gained access to de-identified monthly customer data on these domain hosts from November 2018 to November 2020. Using customer-provided zip codes, the website data are geocoded and aggregated to the 30,000+ inhabited zip codes for which we have data. For the purposes of this study, the zip code data is then aggregated to the county level for analysis.²

The density of domain name hosts represents a new measure of participation in the digital economy for local communities. While domain names may serve businesses, nonprofits, or other pursuits, 3 out of 4 are commercial. Is the density of domain name hosts in a community (the number of websites per 100 people in a county), related to local prosperity and economic opportunity, controlling for other known factors?

To examine this question, we briefly discuss trends in growing spatial inequality across US communities to motivate our study of economic opportunity and discuss a key outcome variable in the statistical analysis. Next, we review previous literature on broadband impacts and the digital economy, and then discuss domain name websites as a measure of digital economic activity. We describe the data and demonstrate how it differs from other measures of digital or business activity in communities: broadband subscriptions, IT employment and small business density. The analysis examines the effects of the density of domain name hosts across counties for annual prosperity, change in prosperity, change in median income, and monthly unemployment before and after the pandemic shutdowns, using time series analysis and statistical matching to address concerns about causation. The statistical results indicate that this online activity is strongly related to prosperity and economic resilience for local communities. These outcomes affect residents as well as businesses, with spillover benefits for the economic well-being of communities.

2. Technology and growing inequality across places

Can technology use reduce growing economic inequality across places, and help communities prosper? Trends in earnings and income since the 1980s show that there is greater income inequality over the past four decades ([Rose, 2018](#); [Bartels, 2008](#)) that is not just due to rising fortunes at the top ([Krause & Sawhill, 2018](#)). Even before the coronavirus pandemic of 2020–21, upward mobility had declined in recent decades. Only a little over half of the age cohort born around 1980 earn more than their parents did, in comparison with nearly 90 % of children born in the 1940s ([Chetty et al., 2017](#)).

This dwindling opportunity is unequally distributed across the landscape and affects places as well as individuals. The Economic Innovation Group’s Distressed Communities Index is one effort to measure these differences across communities with data for zip codes and US counties (see [Appendix Fig. 1](#)). This measure combines seven different indicators of social and economic conditions to measure overall well-being in communities. The map shows a regional pattern with more distressed places in the South and Southwest, as well as some rural parts of the West. There are 50 million Americans living in distressed zip codes. Since the 2008 financial recession, many of the most distressed communities are rural, though they are found in urban and suburban areas as well; people of color are also overrepresented in these disadvantaged zip codes.

¹ A venture is the term used by GoDaddy to describe unique domain name hosts and their redirects to a single domain name (website/email/ssl). Metrics from Data Provider (Netherlands) using WebCrawler data indicate there are roughly 40 million ventures in the US.

² Other analysis has been conducted for zip codes and metropolitan areas—CBSA. Variables such as unemployment are available at the county level, and this is a meaningful unit of analysis for many broadband impact studies in general.

This growing economic inequality across communities preceded the last recession, reflecting structural changes in the economy (Giannone, 2017; Moretti, 2012, p. 105). Following World War II, wage gaps narrowed across communities, but have progressively widened since the 1980s, due to skill-biased technological change and increasing returns to a college education (Giannone, 2017). Knowledge-intensive industries, including technology firms, cluster in urban ecosystems of innovation (Moretti, 2012, chapter 4) concentrating economic activity where workers have higher levels of education and skill.

In recent decades, the economies of a select group of “superstar cities” have boomed on the coasts and in other technology hubs, while many cities and towns in the Midwest and South, along with rural areas, have lagged in economic growth and incomes (Hendrickson et al., 2018; Florida, 2017, p. 6). Economist Enrico Moretti has called this the “great divergence” (2012, 73) and noted that differences across metropolitan areas are now greater than those within them. Chetty and Hendren (2017) provide evidence that places have a causal influence on economic opportunity, and that community-level trends and outcomes are critical for public policy and the nation’s future. These place-based inequalities have long-term effects and are likely to worsen as the economy struggles from the Covid-19 epidemic. Can broadband use, including digital economic activity, help mitigate these trends?

3. Prior research on broadband's community impacts

Technology use has been shown to have positive benefits for individuals, firms and communities. Broadband (high speed internet) lowers costs for communication, transportation, and search for goods and services, influencing the geography of economic activity (Greenstein et al., 2018). Past community-level research has primarily studied the effects of broadband infrastructure investments, availability of service, or the number of providers, though some more recent studies include measures of speed, subscriptions, or devices. Broadband infrastructure generates local economic benefits including increased business activity (Lehr et al., 2006; Holt & Jamison, 2009; Kolko, 2012; Jayakar & Park, 2013; Atasoy, 2013; Mack, 2014).³ Research on broadband deployment has associated it with growth in the number of businesses (Lehr et al., 2006; Kim & Orazem, 2012). This may be because of increased firm entry or improved firm survival (Abrardi & Cambini, 2019), and ultrafast networks have been found to increase sole proprietorships (Hasbi, 2017).

How does broadband affect economic outcomes for residents? Past research has found that local availability of broadband service raises levels of employment (Kolko, 2012; Atasoy, 2013) and reduces unemployment rates (Jayakar & Park, 2013). Yet, Kolko’s (2012) national study of broadband deployment revealed that this growth occurred without generating higher household income or employment for local populations. The extent to which less-educated residents benefit has varied across studies (Akerman et al., 2015; Mack & Faggian, 2013).

Measures of broadband use, such as subscriptions, are associated with stronger effects on local economies than deployment (Whitacre et al., 2014; Gallardo et al., 2020; Mossberger et al., 2021). In their study of rural counties, Whitacre et al. (2014) found that availability had limited effects on economic outcomes, but broadband subscriptions were strongly related to income growth and decreased unemployment. Propensity score matching strengthened the argument that broadband subscriptions had a causal impact on these outcomes. Similarly, Gallardo et al. (2020) compared broadband subscriptions for all counties to measures of availability from the Federal Communications Commission (FCC), and found that indicators of use were better predictors of county productivity than deployment. With time series data over the past two decades, Mossberger et al. (2021) showed that lagged broadband subscriptions in counties predicted growth in median income over time, while measures of broadband deployment did not. Broadband subscriptions were related to a variety of economic benefits for counties and metros over nearly two decades (Mossberger et al. (2021)).

As discussed above, there has been a lack of systematic nationwide data from government sources to measure how this technology is used for different online activities at the community level. Census data on e-commerce covers only online sales,⁴ missing other applications such as marketing and communications for brick-and-mortar businesses, or the gig workers who are not well accounted for in government survey data (Abraham et al., 2018). Some past research has examined internet use across industries using commercial data. Forman et al. (2012) found that between 1995 and 2000, advanced internet use in firms (beyond email or web browsing) was associated with higher wages, but only in counties that were large, had highly skilled workers, and a concentration of the IT industry. While the market research data they used included different types of internet use across counties, it was available only for larger companies. With metrics that measure large and small businesses, the results may differ nearly two decades later.

Researchers have increasingly turned to private sector sources to fill gaps in publicly available data (e.g. Chetty et al., 2020; Gupta et al., 2020). Sources of commercial data describing online transactions include Adobe’s Digital Economy Index and Yelp’s Economic Average. The Adobe index, however, focuses on industry sectors rather than subnational geographies and Yelp’s index of businesses and reviews has a sample of only 50 metros. Neither source can adequately describe local participation in the digital economy across urban and rural communities (zip codes and counties). There is a need for nationwide data that can more comprehensively capture participation in the digital economy.

4. Measuring the digital economy through domain sites

The impact of technology in the economy includes use of digital tools to help businesses function, not just the creation of digital products (Giones & Brem, 2017; Antonizzi & Smuts, 2020). Measures of the digital economy have traditionally relied upon the

³ See also Falck, 2017, Bertschek et al., 2016 and Abrardi & Cambini, 2019 for international overviews.

⁴ 2018 E-commerce Multi-sector Data Tables (census.gov).

proportion of the local economy in the information sector (see Jackson et al., 2019; Moretti, 2012). However, this ignores the extent to which the entire economic system has been altered through digital transformation, including the application of technologies across sectors (Brynjolfsson & Saunders, 2010). For example, a small boutique with a website would not be considered part of the information sector, but that website allows for the business to be found using search engines, reducing barriers such as distance and time to deliver services and products (Auger, 2005).

Domain names are the underlying address book of the internet that facilitates such search, and are governed by the global Internet Corporation for Assigned Names and Numbers (ICANN).⁵ Domain name websites can be used to employ technology in a variety of businesses, large and small, lowering communication and transaction costs for brick-and-mortar businesses, gig workers, and online entrepreneurs. Improved communications can connect niche businesses with consumers, expand visibility beyond the immediate neighborhood, or link commercial activities in sparsely populated areas to broader markets (Greenstein et al., 2018). This provides an opportunity for local businesses to raise regional exports. Websites can also provide valuable customer feedback by hosting other digital applications, such as email, social media, and data analytics. Online sales may reduce costs for real estate, utilities, and insurance. This lowers transaction costs and barriers for firm entry, including for solo entrepreneurs or start-ups that can test an idea online first, with less capital investment than in the past. Websites also enable businesses to be more nimble. During the pandemic, for example, websites allowed businesses to advertise virtual appointments, curbside pickup or delivery, changed hours, and reopening. Overall, use of the internet for commerce should affect communities in much the same way that entrepreneurial activity in general can contribute to local development (Fortunato and Adler 2015; Feldman, 2014), creating positive externalities.

Data on domain name registrations was used to track the geographic diffusion of commercial use of the internet in early studies (Kolko, 1999; Moss & Townsend, 1997) through the address for the site owner. These studies relied on data from the 1990's before internet use was widespread; today domain name websites could be expected to cover a broader range of businesses and regions of the country. Registrations can also include many inactive domains. In contrast, the density of domain name sites used in this study distinguishes between levels of domain activity (traffic, in-links, out-links).

5. Research hypotheses

This study hypothesizes that communities with more domain name hosts will generate positive spillover benefits and have stronger economies, all else equal. Consistent with past research, we expect that broadband subscriptions also will have positive externalities for communities.

H1. The density of domain name hosts (or ventures) and highly active domains in a county have positive and significant effects for each of the dependent variables below:

An index of economic prosperity

Change in prosperity

Change in median household income

H2. The density of domain name hosts (or ventures) will have a negative and significant effect on monthly unemployment rates between January 2019 and November 2020 (the range of available venture data); higher venture density will lead to decreased county unemployment, both before and after the Covid-19 shutdowns.

Commercial data made available by the world's largest registrar of domain names enables us to examine these hypotheses in new ways.

6. Data: domain name hosts

6.1. What do these domain name sites represent?

The density of domain name websites in a community can be used to learn whether the collective online presence of local entrepreneurs creates spillover benefits for the local economy. GoDaddy refers to these domains and their redirects as "ventures," and in the aggregate they provide a footprint of digital activity at the community level that is more specialized than the general use enabled by broadband subscriptions.

With data on 20 million domain name hosts, this data accounts for more than half the population of domain name hosts in the US. Redirects linking to one main domain name are counted as a single venture. The data analyzed in this study are raw, de-identified data geo-coded by the zip code of the site's owner. GoDaddy provided the authors with monthly customer data which the authors aggregated to the zip code-level, but they have not determined the questions we have asked or how the data is analyzed. GoDaddy has also made these data publicly available on its Venture Forward website.⁶

Because counts of domain name websites are aggregated to the zip code and county level for this analysis, we focus on the *characteristics of geographies* in which these websites are located. We have some limited information on owners and their purposes

⁵ <https://www.icann.org/resources/pages/welcome-2012-02-25-en>.

⁶ <https://www.godaddy.com/ventureforward/>.

gathered through random-sample surveys of customers. Using machine learning and web scraping to categorize information from the home pages, data scientists at GoDaddy estimated that approximately 80 % of the domain names were commercial.⁷ Random sample online surveys of GoDaddy customers, with questions designed and analyzed by the authors, were also administered by the polling firm Advanis (<https://www.advanis.net/>) in August 2019 and July 2020, yielding over 2000 responses in each year.⁸ The surveys confirmed the prior machine analysis, as 3/4 of respondents said their ventures were commercial in both years.

The survey results indicated that businesses supported by most domain hosting websites are quite small. Approximately 8 % represented organizations with more than 10 employees, and 55 % were operated by solo entrepreneurs. Nearly half of site owners in 2020 (48 %) considered their ventures as their main source of employment, while another 29 % said that it was a second job or side employment.⁹ Respondents were largely similar across the two surveys, with a few exceptions. Some digital microbusinesses overlap with brick-and-mortar businesses, but in 2020 one-third (35 %) reported that they were online only, in contrast with 1 in 5 in 2019. The growth in online-only establishments is likely a result of the pandemic shutdowns.¹⁰ In the 2019 survey, 35 % of respondents reported that their audience was local, but the rest aimed for the wider markets or audiences that contribute to the community's export base, attracting additional dollars into the local economy. What the descriptive statistics show clearly is that these data measure digital economic activity. In some instances, websites are supplementing brick and mortar operations, while others are online businesses.

Unlike measures of the digital economy that rely on the size of the IT sector, domain name hosts cover all industry sectors. This study relies on the population of US GoDaddy domains with data gathered monthly. This provides a dynamic measure of economic activity that is not reliant on large scale surveys that are both resource and time-intensive, with significant temporal gaps and lags in availability.

6.2. Density of domain name hosts and maps

The predictor variable in this study is a count of all domain name websites per zip code or county divided by the adult population to create a density measure. These data are geocoded by zip code and then aggregated to county level. Fig. 1 shows the density of domain name websites per 100 for zip codes and Fig. 2 the same metrics for the 3007 counties, both measured in 2018.¹¹ Zip codes more accurately depict variation in ventures avoiding geographic distortion from large counties (especially in the West). Counties are relatively large units of analysis, while zip codes approximate urban neighborhoods or smaller communities.

In Fig. 2, counties with a higher density of ventures, in darker shades, are visible across much of the nation's interior as well as along the coasts. Metropolitan areas tend to be darker, but evidence of activity stretches out well beyond the immediate counties around Midwestern cities like Minneapolis, Chicago and Detroit, and borderland regions near El Paso. The heavy presence of domain name hosts in parts of the West, Texas, Mid-South, and Florida includes rural areas. Communities with a high density of sites are spread throughout the country, with concentrations evident in some rural counties as well as cities.

Displaying domain name sites by zip code shows a less solid wall of activity in the West (since rural counties are so large in this region), and more variation in other regions as well. Yet this does not change the pattern of high venture density across different types of communities in the US. Domain name sites serve diverse communities and are not confined to areas more traditionally associated with the digital economy, such as coastal regions or tech hubs.

Domain name hosts can be sorted into four distinct groups based on their activity over time—by website age, by demand (how busy is the website in terms of traffic and economic footprint with Amazon crawler data), by connection (how networked is the venture across the internet, both in-links and out-links), and by breadth (upgrades, downgrades or products added or subtracted).¹² Low levels of activity characterized approximately a third of all domain name hosts, with moderate activity for another third. Moderate-high levels of activity represented just over a quarter of all websites. Very high-activity ventures comprised approximately 10 percent of the total depending on the time period. In the analysis that follows, we group together moderate-high and high activity websites, which represent 1/3 of domain names, as "highly active." Places with more highly active domain name websites should have stronger economic outcomes, as the commercial activity on the sites is greater. The geographic distribution for the four clusters is comparable across counties and zip codes, with similar patterns based on overall domain name densities (see Appendix Fig. 2). Places that have many commercial websites with low to moderate activity also have many highly active websites at both the zip code and county level. The correlation is very strong (.85); geographic areas with extensive digital commercial activity have more domain name hosts at all levels.

⁷ GoDaddy data scientists used Word2Vec plus a deep model (Recurrent Neural Network) with specialized memory units (LSTM) that improve learning the context from text to classify domains over 265 industry categories.

⁸ The survey was conducted online, with respondents contacted with an email invitation. The survey samples had 2006 respondents recruited through random stratified sample in 2019 and 2330 in July 2020. Respondents in both years were restricted to venture owners or employees responsible for the site, and contractors were excluded in an initial screening question. Participants were incentivized to participate with an Amazon gift card.

⁹ Around 22 % in 2020 were not working when they started their ventures, as they were laid off, retired, disabled, students, stay-at-home parents, etc. Thus, ventures represent a variety of businesses including small start-ups and side "gigs," not always well accounted-for in government business data or contingent employment (Torpey & Hogan, 2016; Abraham et al., 2018).

¹⁰ Additionally, 61 % of venture owners said their website helped their businesses to persevere during Covid-19.

¹¹ Not all zip codes are inhabited, and while there are over 35,000 zip codes, domain name data only exists for approximately 30,000.

¹² For a subsample of just over 10 million domain name hosts, in addition to measuring counts of domain name hosts in zip codes, data from web crawlers provides additional measures of domain name activity.

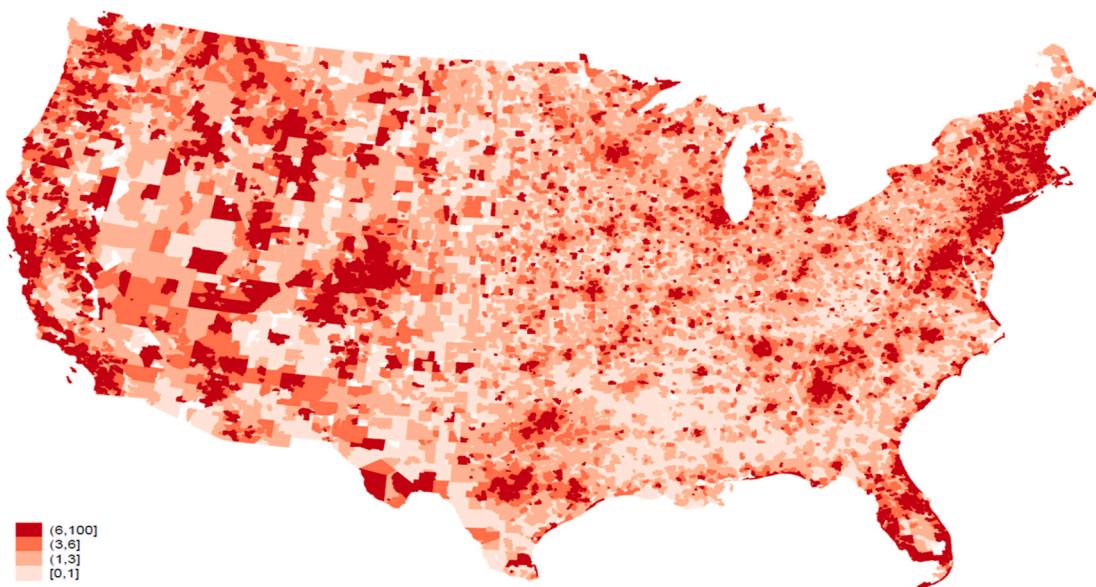


Fig. 1. Density of Domain Name Hosts and their Redirects per 100 people, by Zip Code.

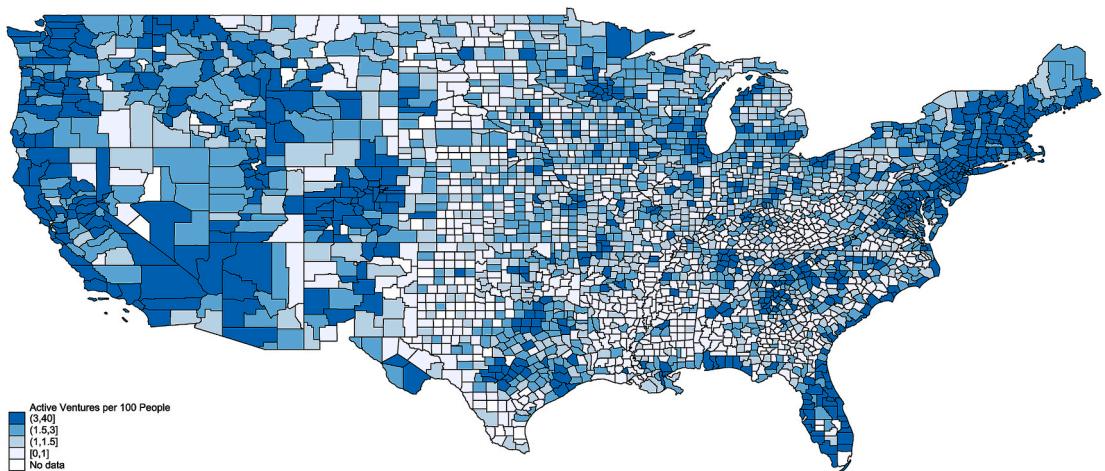


Fig. 2. Density of Domain Name Hosts and their Redirects per 100 people, by County.

7. Outcome measures and methods

7.1. Measuring economic prosperity and recession recovery

The measure of economic prosperity and recession recovery used in this study is a modification of the Distressed Communities Index using updated 2019 Census data for US zip codes and counties; The index is a “comparative measure of the economic vitality and well-being of U.S. communities” (EIG 2018, 2), and has been used in other scholarly research.^{13,14} We are interested in what contributes to prosperity rather than distress, so we have inverted the original index, categorizing communities from distressed (low

¹³ <https://eig.org/dci/methodology>. The EIG includes a measure of recovery from the recession that compares local performance on the seven metrics in two periods- 2007 to 2011 and 2012–2016.

¹⁴ The original Distressed Communities Index has been used as an independent variable in at least a dozen published studies predicting the influence of socio-economic factors in the community on health outcomes (e.g. Hawkins et al., 2019).

values) to prosperous (high values).¹⁵

The index combines data on seven component metrics¹⁶ including 1) percent of the adult population without a high school diploma (or equivalent), 2) housing vacancy rate, 3) percent of the population aged 25–64 not in the work force, 4) poverty rate, 5) median household income as a ratio (percentage) of the state median household income (to adjust for cost-of-living differences), 6) percent change in the number of jobs, and 7) percent change in the number of business establishments.¹⁷ The zip code data are drawn from the Census's 2019 American Community Survey 5-year estimates and from 2019 Zip Code Business Patterns. As outcome variables in the regression analysis (discussed more below), economic prosperity is measured in 2019 and economic recovery measures change in the index from 2007 to 2019.

The index is used to evaluate economic outcomes because it provides a more holistic understanding of a community's economic well-being, operationalizing the concept of spillover community benefits from business activity. Prosperity is often defined simply in terms of economic growth, though Feldman (2014) argues that the concept suggests something more about the overall quality of life and future trajectory of a community, as represented by the types of variables used here. As a measure of broader prosperity, the index combines commonly used indicators of economic growth such as change in the number of business establishments and jobs (Lehr et al., 2006; Kolko, 2012; Whitacre et al., 2014) with outcomes for residents such as poverty rates and labor force participation. These impacts are important to examine, as prior research on broadband availability revealed increases in jobs and businesses with no significant change in local unemployment or local wages (Kolko, 2012). While there is also value in understanding how this digital activity influences specific economic indicators, we argue that it is also important to understand their role in economic prosperity more broadly.

7.2. Modeling approach

Multivariate regression is used to control for other factors related to these outcomes, so the relationships we investigate go beyond correlation. Where appropriate we also use time series models and fixed effects for month. The models are weighted by county population, clustered by state and use robust standard errors. As in Whitacre et al. (2014), we also use statistical matching to ensure that counties with high and low density of domain name websites are comparable. We use 2019 Census data to measure community economic prosperity one year later than the domain name host data, measured in 2018.

Incorporating *change* in our statistical models helps control for endogeneity or confounding factors and helps ensure directionality as we lag our measure of the density of domain name websites by one year. Most of the explanation for the annual growth in median income, for example, can be attributed to the zip code or county median income in the previous years, so by examining change in median income, we control for the effects of income or prosperity in the earlier period. Measuring economic prosperity from two time periods (2007 and 2019) is especially important for addressing whether venture density can influence local economies, such as recovery from recession. We also evaluate monthly change over time for unemployment rates from the Bureau of Labor Statistics (BLS). Measuring change, whether change in prosperity, median income or unemployment rates is incorporated into the models to ensure that the predictor variable was measured prior to the outcome.

8. Results

8.1. Descriptive statistics

How do domain name sites differ from other measures of digital and economic activity? Data on broadband subscriptions are now available from the 2018 Census for all US counties, zip codes and tracts. The correlation between broadband subscriptions and domain name websites per capita for US counties is only moderate, at .53. This means these two variables are distinct. While broadband enables the creation and use of websites, internet use for commercial activity should have a more direct effect on economic outcomes.

Does this merely reflect other aspects of the digital economy such as the concentration of the tech industry? There is only a 0.37 correlation with the percentage of the population employed in IT jobs and the density of ventures for counties.¹⁸ This is a moderate correlation, suggesting these are quite different measures. At the zip code level this correlation is only 0.24 for ventures per capita.

The Census also publishes data on small businesses. The measure of small business density from the US Census used in this study is the number of businesses with 100 or fewer employees per 100 people. Despite moderate overlap, the domain name host data capture something other than the small businesses counted by the Census,¹⁹ whether small businesses are defined as having 100 employees or

¹⁵ Our Economic Prosperity Index (EPI) measure is 100 minus the distress score calculated by the Economic Innovation Group, ranging from 0 to 100.

¹⁶ To derive the distress score, "Each component is weighted equally in the index, which is calculated by ranking communities on each of the seven metrics, taking the average of those ranks, and then normalizing the average to be equivalent to a percentile," according to EIG (2018, 3).

¹⁷ Inverting the index does not change its distribution since each of the 7 component variables are rankings, not raw numbers. Since the study's theoretical focus is on predicting prosperity rather than distress, we reversed the direction of the high/low values. When we modeled the original distressed communities index instead of the new prosperity index, we get identical results for coefficients (except the intercept), but with the signs reversed.

¹⁸ As measured by the US Bureau of the Census North American Industrial Classification System (NAICS, 2017) data aggregated at the county level.

¹⁹ Economic Census, Bureau of the Census, 2018.

less (correlation with ventures is 0.43) or even 10 employees or less (0.53 correlation) at the county level. A major advantage of these data is not only their granularity and spatial coverage, but their availability over time to track where US ventures are increasing or decreasing beginning in 2018. What can this new digital footprint of local economic activity tell us?

8.2. Multivariate regression, statistical matching and covariates

Multivariate regression is used to examine the relationships between the number of domain name hosts per capita and the community prosperity index, change in prosperity (recovery from recession), change in median income, and monthly change in unemployment rates across counties. A particular challenge to using this data is the inherent endogeneity to economic development and technology activity. To overcome this hurdle, we use coarsened exact matching (Blackwell et al., 2009) to identify counties that share similar characteristics but differ in the density of domain name hosts. The treatment is defined as belonging to a community with above average domain name host density and the control belonging to a community below the mean on this variable. We then match counties by the following characteristics: broadband subscription rates, small businesses per capita, the racial demographics of a county, and the county population. Broadband subscription data is from the 2019 American Community Survey 5-year estimates and is defined as all types of broadband (cable, satellite, DSL, fiberoptic, and mobile).²⁰ Small business density includes establishments with 100 or fewer employees (2019 County Business Patterns). For both sets of treatments, roughly 2000 counties were matched along these criteria. This allows us to approximate an experimental design and produces weights that can be used in regression.

After matching counties, we identify covariates that are likely to predict economic prosperity. Given the geographic clustering of human capital in recent decades, economic prosperity scores are likely related to educational attainment. The models control for education (percent high school graduates and percent college), testing whether domain name sites still have a statistically significant effect. Including age cohort controls for the share of the population of working age also permits us to examine the effects of millennial presence. Metropolitan areas with more dynamic local economies have attracted millennials in recent years (Frey, 2018), and this may also be related to economic growth. Other control variables used in the analysis have been shown to influence local economies in prior research. The percentage of the population employed in various industries represents the structure of the local economy and opportunities for residents. The relative industry mix can affect local growth and decline for businesses and jobs as well as wages (Kemeny & Storper, 2015). Data is included on employment by 16 major industry sectors, using the high-level 2-digit classifications in the North American Industry Classification System (NAICS) codes from the Census.²¹

8.3. Outcome variable 1: what predicts local economic prosperity?

The results in the first column of Table 1 show that the density of domain name hosts per capita is a significant and positive predictor of county prosperity, controlling for other factors and using statistical matching to control for possible selection bias. The same results hold for the second model that shows the change in prosperity scores from 2007 to 2019. Higher density of domain name hosts is associated not only with counties being more prosperous, but also with higher levels of economic recovery following the 2008 recession. The effects are even stronger, however, for highly active domain name hosts (in columns 3 and 4). For each additional highly active domain per 100 people, a county's prosperity score increases by 1.2 points, all else equal, compared to .58 for all domain name hosts. As broadband subscriptions increase in a county, economic prosperity increases as well. These additive results suggest that a community's digital activity—including broadband adoption and use—is critical in local economic development today. Appendix Table 1 replicates these models using the full 3000 counties without using statistical matching. Similar results are reported.²²

Fig. 3 reports the substantive relationship between the density of domain name websites and economic prosperity, with all other variables held constant at mean values using probability simulations and confidence intervals.²³ This provides a graphical representation of how levels of prosperity change, so we are evaluating the slope of the line rather than the predicted levels of prosperity, which are a function of holding all other covariates at their means. The graph on the right, for highly active websites, slopes more sharply upward compared to the graph for all active ventures on the left. Fifty percent of counties have prosperity scores that range between 32 (25th percentile) and 56 (75th percentile), so this is a large substantive impact from adding just one domain name host per

²⁰ To generate our measure of broadband subscription rates we subtracted the percentage of the population who relied solely on cell phones from those with high-speed internet access.

²¹ The following variables were drawn from the 2019 5-Year American Community Survey (ACS, 2019): broadband subscriptions (percent of the population with a high-speed internet subscription, all types but excluding dial-up internet), percent Black, percent Asian-American, percent Native American, percent Hispanic, percent high school graduates, percent college graduates, percent employed, and age group cohorts. Percent of the population employed in industry code data is available in the 2019 ACS supplemental data, table K202403.

²² As a robustness check, we estimated the models from Table 1 in parallel analyses, one with broadband subscriptions only and one with domain name hosts only. In separate analyses both broadband and domain name hosts are positively and significantly related with prosperity and changes in prosperity. Additionally, the coefficients were larger (although only slightly larger) in the separate specifications compared to the model that includes both measures together. This again suggests that these measures both have a role in understanding the relationship between digital activity and economic prosperity.

²³ The predicted values were generated using the *margins* command in STATA 14. The reason the range for prosperity is much narrower than presented in the map is that the margins command holds all other factors constant (and defaults to their mean value). This allows us to isolate the relationship between ventures and prosperity directly. The reader should focus on the slope, not the intercept of the line as the values could move up or down depending on what the other covariates are set to.

Table 1

Predicting economic prosperity (2019) and change in prosperity (2007–2019) US counties (results from coarsened exact matching).

	Prosperity Index	Change Prosperity Index- 2007-2019	Prosperity Index (Highly Active Ventures)	Change Prosperity Index 2007–2019 (Highly Active Ventures)
Broadband Subscriptions	0.4728** (0.0366)	0.1308** (0.0440)	0.4405** (0.0385)	0.1369** (0.0466)
Venture Density (domain name hosts per capita)	0.5851** (0.1426)	0.7182** (0.1686)		
Highly Active Venture Density (domain name hosts per capita)			1.2401** (0.5192)	1.4061** (0.6118)
Small Business Density	-2.5881** (0.4092)	-5.9374** (0.5006)	-1.5396** (0.4130)	-4.4933** (0.5322)
Percent Black	-0.1574** (0.0286)	0.0476 (0.0353)	-0.2414** (0.0356)	0.0452 (0.0419)
Percent Native American	-0.0254 (0.1194)	-0.0045 (0.1417)	0.1005 (0.1206)	-0.0482 (0.1425)
Percent Asian	-0.7295** (0.3200)	-1.5484** (0.3797)	0.1243 (0.3453)	-0.4655 (0.4068)
Percent Hispanic	-0.1812*** (0.0301)	-0.0472 (0.0359)	-0.1640** (0.0329)	-0.1007** (0.0394)
Percent Agriculture	-0.0927** (0.0456)	-0.2474** (0.0552)	-0.2432** (0.0438)	-0.2873** (0.0580)
Percent Construction	0.5643*** (0.1102)	0.2178* (0.1314)	0.7350** (0.1137)	0.2256* (0.1369)
Percent Wholesale	0.3392* (0.1864)	-0.4406* (0.2276)	0.5452** (0.1911)	-0.2741 (0.2330)
Percent Retail	-0.1054 (0.1022)	0.1418 (0.1242)	-0.1950* (0.1064)	0.0816 (0.1275)
Percent Transport	-0.5495** (0.0972)	-0.8169** (0.1208)	-0.3004** (0.1017)	-0.6114** (0.1264)
Percent Info	-1.3949** (0.2865)	-0.7953** (0.3676)	-0.1215 (0.2588)	0.1066 (0.3276)
Percent Finance	0.1818 (0.1268)	-0.2400 (0.1512)	0.2573* (0.1503)	-0.4431** (0.1796)
Percent Professional	0.3266** (0.1331)	0.3335** (0.1636)	0.3125** (0.1375)	0.5380** (0.1687)
Percent Education	-0.7203** (0.0635)	-0.5533** (0.0765)	-0.7151** (0.0638)	-0.6011** (0.0771)
Percent Other	-0.4055** (0.1857)	-0.2923 (0.2235)	0.0986 (0.1958)	0.2541 (0.2339)
Percent Public	-0.3418** (0.0888)	-0.5606** (0.1076)	-0.2629** (0.0914)	-0.5305** (0.1099)
Percent High School	0.1128* (0.0584)	-0.2897** (0.0704)	0.1759** (0.0614)	-0.2838** (0.0740)
Percent College	0.8155** (0.0618)	0.1500** (0.0736)	0.7636** (0.0662)	0.0242 (0.0791)
Percent Millennial	-0.6671** (0.0934)	-0.4960** (0.1145)	-0.5921** (0.0994)	-0.4441** (0.1229)
Percent Gen X	0.5429** (0.1148)	0.0171 (0.1391)	0.5968** (0.1185)	0.1975 (0.1438)
Percent Baby Boomer	-0.5141** (0.1627)	-0.7667** (0.2002)	-0.5556** (0.1607)	-0.5812** (0.2040)
Percent Employed	1.0234** (0.0443)	0.0977* (0.0530)	1.0131** (0.0469)	0.0482 (0.0562)
Constant	-29.0082** (7.0003)	47.0695** (8.4044)	-37.9326** (7.1847)	36.9165** (8.6739)
F-Stat	187.355	33.113	143.628	29.997
R-Squared	0.700	0.297	0.647	0.283
AIC	14079.880	14353.58	13894.50	14022.31
Observations	1947	1899	1904	1843

Standard errors in parentheses, clustered by state and weighted by population.

*p < .1.

**p < .05.

100. Given the concentration of counties with a score near 50, even a one or two unit increase in prosperity is meaningful and represents a substantively large change in economic outcomes.

As expected, broadband subscriptions are also statistically related to county prosperity, independent of ventures as shown in Table 1; a 1 percent increase in broadband subscriptions leads to a 0.4 increase on the economic prosperity index, all else equal. Communities with higher rates of broadband subscriptions are more prosperous; but using this connectivity to create commercial and other activity online has an even larger effect on prosperity. Communities that are more reliant on agriculture, construction, retail, and

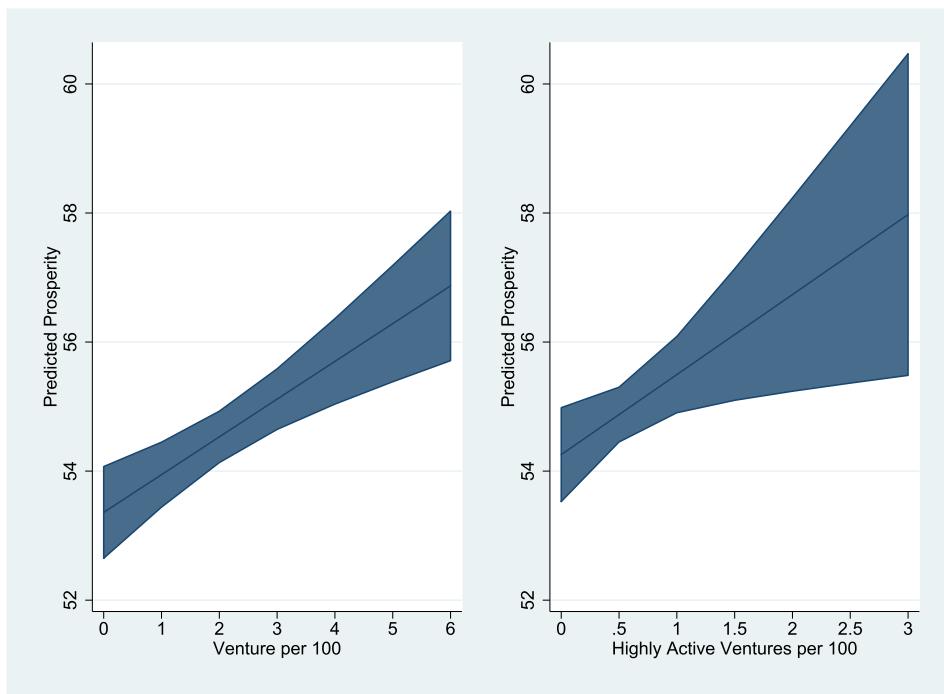


Fig. 3. Predicted 2019 economic prosperity varying density of domain name hosts (left) and highly active domain name hosts (right).²⁶

education for employment have lower prosperity. Given the modest correlation between small business density and commercial domains, the covariate for small businesses in the models represents those without a digital interface (i.e. after the covariance between the two variables has been removed from both variables in the multivariate regression analysis). It is less surprising then, that small business density does not produce positive economic effects after accounting for the positive influence of digital microbusinesses.

8.4. Outcome variable 2: recovery from the 2008 recession

As discussed above, another way of addressing whether this digital commercial activity is related to economic outcomes is to measure change over time. We measure change in economic prosperity from 2007 to 2019, with an additional test of causal relationship. Table 1 (columns 2 and 4) indicates that counties that experienced the greatest recovery from the recession had more ventures and highly active ventures in 2019, compared with the counties that recovered the least. Results from the multivariate regression model weighted through the statistical matching and shown in Table 1 demonstrate that both venture density and broadband subscriptions are positive and significant predictors of changes in prosperity over this period. Fig. 4 uses probability simulations to graph the substantive relationship between ventures and the change in prosperity. Again, both ventures and highly active venture density predict greater increases in prosperity from the 2008 recession, and highly active ventures again have a much stronger relationship.

8.5. Outcome variable 3: three-year growth in median household income

The prosperity index includes multiple outcomes, and a central one of concern to policymakers and their communities is whether economic activity increases household income. We therefore examine separately the relationship between the density of domain name hosts and changes in median household income across counties, holding other factors constant. We ask what factors account for changes in median household income from 2016 to 2019. By measuring change, we control for the fact that places with high incomes in 2016 are likely to also have high incomes in 2019; this more favorable starting position may be responsible for the median income in 2019 rather than domain name sites. We are therefore interested in what role ventures might have in growth or decline. We again use the coarsened exact matching method outlined above to model the relationship between ventures and highly active ventures with prosperity.

On average, across the counties, median household income increased by \$5,520 between 2016 and 2019. Two-thirds of counties experienced a change between \$2,000 and \$8,000, with some outlier counties that increased median household income by more than \$10,000 and others that experienced income decline.

The multivariate regression models in Table 2 indicate that domain name sites have a positive coefficient but are statistically insignificant after controlling for broadband subscriptions, small business density, demographics, education, industry mix, etc. But highly active ventures (clusters 3 and 4) have a statistically significant effect with a 90 % confidence interval. The models in the

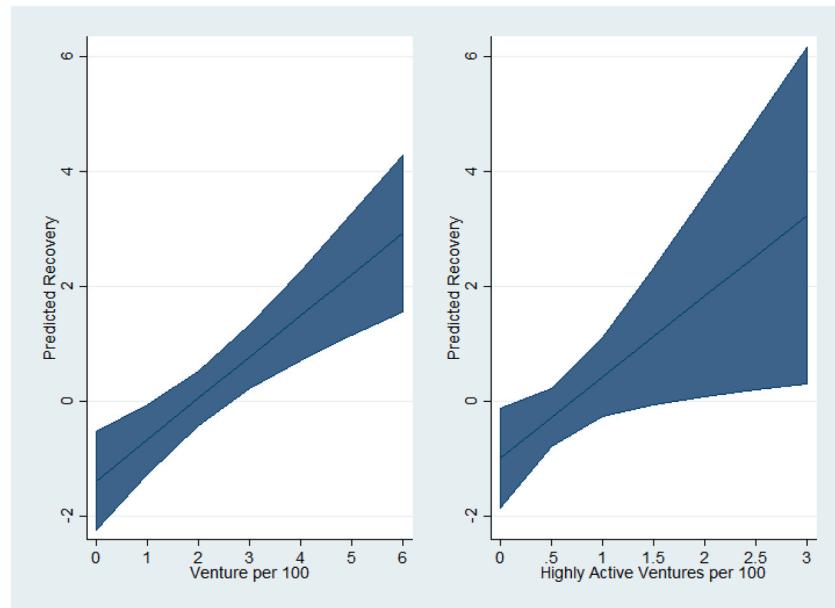


Fig. 4. Counties with more domain name hosts recovered more as measured by change in the prosperity score from 2007 to 2019, US counties.²⁶

Appendix using the full sample report a 95 % confidence interval. Controlling for the same variables mentioned above, median household income rises by an additional \$370 for each highly active site per 100 people in the county. Since the average county increase is \$5,500, this is a substantively large increase of 7 percentage points. Adding three highly active ventures per 100 people increases county median household income by over \$1,000 over the three-year period, for a 20 percentage-point increase.

8.6. Outcome variable 4: monthly unemployment rates January 2019–November2020

Finally, we leverage the temporal richness of our data by modeling monthly unemployment rates from January 2019 to November 2020 using outcome data from the Bureau of Labor Statistics (BLS). This period incorporates both the pre-Covid trends in a relatively strong economy and the economic shock following shutdowns in March and April. In contrast with the earlier cross-sectional models, the unemployment rate is measured on a monthly basis, meaning we have many repeated observations of the same county over a 2-year period, providing almost 60,000 county-month observations.²⁴ Combined with the more comprehensive prosperity measure used earlier, this allows us to demonstrate the value of domain name hosts in both the broader economic context, and across time on a specific component of economic well-being.

The mean county unemployment rate peaked at over 12 % in April 2020, compared to a mean rate of 4 % in February of 2020. Given the other positive economic effects of domain name sites, how do they affect local unemployment? See Fig. 5 below for a graph of monthly unemployment rates across counties. When interpreting the coefficients, it is important to remember that the dependent variable has a much narrower range (99 % of all data between 0 and 17) than the prosperity measure, so smaller coefficients do not indicate a smaller effect size.

We estimate two regressions with time fixed effects by month to model the influence of venture density on county-level unemployment, where domain name hosts lagged (in the previous month) are used to predict unemployment rates in the following month. In Table 3 the model in the first column includes the full monthly time series from January 2019 to November 2020. The data in the second column includes just unemployment rates from March–November 2020. We split the results to understand whether domain name hosts have a differential effect in the Covid era, where the digital economy has been predicted to play a more central role under social distancing conditions. We use the same covariates as in the earlier models but now can model month-to-month shifts in both the

²⁴ The prosperity index can only be updated annually due to Census bureau data availability, and data is not available until December of the following year. Data for a 2020 index will not be available until December 2021.

Table 2

Predicting 3-year Change in Median Household Income (CEM results).

	(1)	(2)
	Change Median Income 2016–2019	Change Median Income 2016–2019
Broadband Subscriptions	5.4443 (18.6329)	18.8168 (22.7415)
Venture Density (domain name hosts per capita)	23.8628 (43.0274)	
Highly Active Venture Density (domain name hosts per capita)		370.2699* (203.6186)
Small Business Density	-741.6828** (258.4008)	-450.7671* (232.2533)
Percent Black	-11.9962 (11.4605)	-11.6656 (10.9730)
Percent Native American	75.0322 (48.2934)	50.0802 (47.1777)
Percent Asian	-252.7283* (147.6568)	-242.9095 (169.0443)
Percent Hispanic	-16.5320 (18.0181)	-15.9263 (16.5578)
Percent Agriculture	-88.3548** (30.1923)	-120.6812** (30.1593)
Percent Construction	15.6369 (53.6990)	87.7603 (60.9622)
Percent Wholesale	113.8891 (116.2385)	175.7562 (123.4600)
Percent Retail	-93.2812* (49.2040)	-97.5248* (52.1165)
Percent Transport	-30.0364 (98.7115)	14.9208 (108.9969)
Percent Info	63.7199 (167.6605)	-28.6719 (138.0413)
Percent Finance	-99.7102 (71.5813)	-35.1840 (74.3892)
Percent Professional	52.7529 (74.6881)	51.2904 (68.5910)
Percent Education	-78.9158** (32.3552)	-36.5062 (37.9201)
Percent Other	-152.0718 (103.9459)	-145.9857 (108.9152)
Percent Public	16.6109 (50.7579)	50.9990 (46.3794)
Percent High School	-32.7161 (25.6423)	-11.1596 (29.7058)
Percent College	122.3036** (41.2000)	72.5204** (27.6629)
Percent Millennial	-68.2983 (61.7115)	-147.1542** (58.6834)
Percent Gen X	-30.9996 (83.6473)	-22.1223 (79.2682)
Percent Baby Boomer	64.6973 (90.5867)	-39.4650 (93.8254)
Percent Employed	66.8783** (28.3244)	83.1553** (28.8052)
Constant	6608.9326** (3213.5854)	4805.6758 (3916.8101)
F- Stat	10.36	9.523
R-Squared	0.180	0.173
AIC	36576.27	35830.20
Observations	1948	1905

Standard errors in parentheses, clustered by state and weighted by population.

* $p < .1$.** $p < .05$.

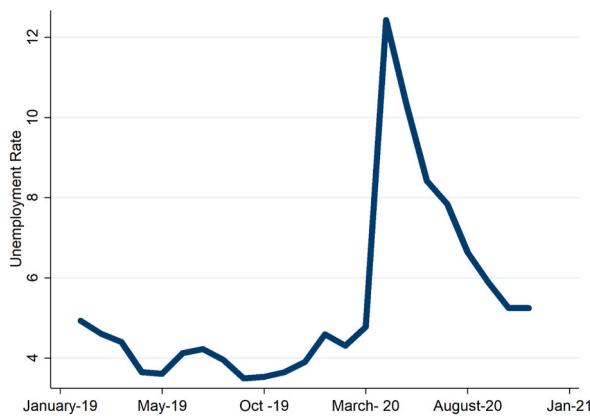


Fig. 5. Mean county unemployment rate by month.

number of ventures and the unemployment rate. We also interact domain name sites with broadband subscription rates to explore how they affect unemployment rates together, given the statewide shutdowns related to Covid-19 that forced commercial activity online.²⁵

The time series results predicting monthly unemployment show domain name site density does not reduce unemployment rates over the year and half period. However, this null result masks an important, statistically significant, interaction between the density of domain name hosts (ventures) and broadband subscriptions (the percent of the population with high-speed internet) shown in both column 1 (full time-period) and column 2 (pandemic period). When broadband subscription rates are very low, more domain name

Table 3
Modeling monthly unemployment rates January 2019–November 2020

	Unemployment Rate Jan 19–Nov 2020	Unemployment Rate March–Nov 2020 (Pandemic)
Venture Density- Lagged (domain name hosts per capita)	0.1893* (0.0688)	0.5270* (0.1393)
Broadband Subscriptions	0.0330* (0.0029)	0.0838* (0.0065)
Lagged Venture Density X Broadband Subscriptions	-0.0018** (0.0008)	-0.0054** (0.0016)
Small Business Density	-0.0583* (0.0085)	-0.0657 (0.0841)
Percent Black	0.0323* (0.0019)	0.0567* (0.0040)
Percent Asian	0.0392* (0.0087)	0.1108* (0.0171)
Percent Hispanic	0.0500* (0.0031)	0.0910* (0.0055)
Percent Native American	0.0455* (0.0030)	0.0609* (0.0061)
Percent Agriculture	0.0109 (0.0078)	-0.0975* (0.0156)
Percent Construction	-0.2108* (0.0152)	-0.3316* (0.0293)
Percent Wholesale	-0.0409 (0.0424)	-0.2385* (0.0788)
Percent Retail	-0.0700* (0.0124)	-0.1144* (0.0236)
Percent Transportation	0.0398*	0.0365

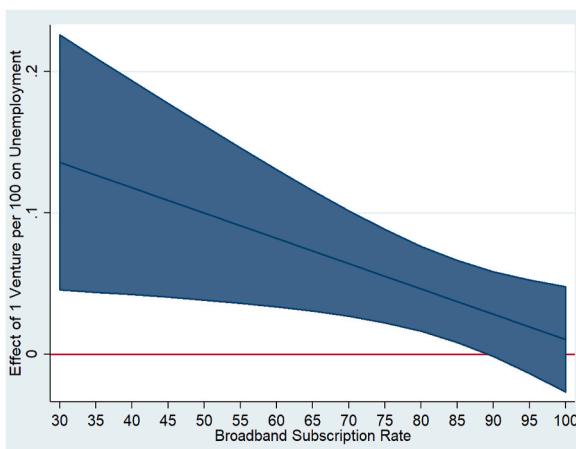
(continued on next page)

²⁵ We use slightly different specifications in these models than the earlier models for a variety of reasons. First, the data are fundamentally different with this data representing repeated observations of the same data (time series cross-sectional data, nearly 70,000 observations), while the early analysis was a 1 period cross section (only 1800 observations). Second, the Covid crisis represented an exogenous shock to communities that led many areas with the highest employment rates to experience the largest increases in unemployment because communities were forced to shut down virtually overnight, regardless of prosperity level.

²⁶ Note that the different ranges along the X axis are due to a higher density of ventures than highly active ventures. Box axes reflect the plausible range that each measure is observed.

Table 3 (continued)

	Unemployment Rate Jan 19-Nov 2020	Unemployment Rate March–Nov 2020 (Pandemic)
Percent Information	(0.0159) 0.1474 (0.1003)	(0.0307) 0.2725 (0.1501)
Percent Finance	−0.0502* (0.0145)	−0.0816* (0.0285)
Percent Professional	−0.0005 (0.0177)	−0.0316 (0.0290)
Percent Educational	0.0709* (0.0081)	0.0730* (0.0159)
Percent Public	−0.0585* (0.0078)	−0.1713* (0.0173)
Percent High School	0.0140* (0.0037)	0.0367* (0.0139)
Percent College	−0.0862* (0.0052)	−0.1337* (0.0135)
Percent Millennial	0.0537* (0.0104)	0.1282* (0.0185)
Percent Gen X	0.0675* (0.0131)	0.0766* (0.0271)
Percent Baby Boomer	0.2163* (0.0154)	0.5989* (0.0431)
Constant	−4.1253* (0.9811)	−10.5287* (1.7490)
F-Stat	387.36	154.73
R-Squared	0.7530	0.6685
AIC	290306.59	129761.41
Observations	68,527	27,234

**Fig. 6.** Marginal effect of ventures on unemployment.

sites are associated with an increase in the unemployment rate. However, as broadband subscription rates increase, domain name hosts start to have a positive effect in reducing the unemployment rate. So, a community must have **both** moderate to high broadband subscription rates and more online commercial businesses to see the beneficial effects. Fig. 6 plots the marginal effects of domain name hosts at different levels of broadband subscription. When broadband subscription rates are below 60 %, domain name hosts increase unemployment, whereas when broadband subscription rates are very high, they reduce unemployment.

When looking at the effect size, it is important to remember that a single percentage change in the unemployment rate represents hundreds of thousands of jobs. As can be seen in Table 3, the effect size in Table 3 column 2 is roughly three times the size (0.0054) compared to the pooled model in column 1 (0.0018). Given that there were 160 million Americans either employed or looking for a job, a reduction in the unemployment rate by 0.0054 instead of 0.0018 represents nearly 5800 more jobs saved for every additional domain name host per 100 people during Covid-19 than in the larger time period, holding all else constant in a given month. In total, adding one additional venture per 100 people saved roughly 8600 jobs per month. Given that the nationwide average venture density is 3 during the 9 months of observation during the pandemic, this translates to 230,000 jobs preserved by ventures, a substantively

important effect size.

These results indicate that both broadband subscriptions and digital commercial websites are critical components for integrating a community into the 21st century economy. A digitally connected community may represent a more robust local market for online transactions and communications, especially during the pandemic. Furthermore, a community with high levels of broadband adoption will not see economic benefits without its application for commercial purposes, such as ventures. These results highlight that policy makers need to do more than just increase broadband access, but also foster an environment where digitally enabled businesses grow.

9. Conclusion: a new measure of the digital economy for local communities

A challenge for local communities across the country is how to create opportunity for businesses and residents. This is especially true in the face of the pandemic's economic disruption, and for places that had already languished in an era of widening inequality. Digital economic activity, as measured by the density of domain name hosts, adds to the body of research on broadband and its economic impacts, improving data on specific uses of technology beyond broadband connectivity or IT employment. By leveraging commercial data, we overcome the lack of community-level data on activities online.

These domain name hosts are significant predictors of community prosperity and other economic benefits, controlling for other influences on local economies. Statistical matching is used to mimic an experiment with observational data; the treatment we examine is counties with high venture density and the control is counties with low venture density. The results show that as the density of domain name hosts rises in a county, so does the county's prosperity, as determined by an index of components that include poverty, labor force participation, jobs and outcomes for residents, as well as business growth. Examining change on this index over more than a decade shows that places with digital economic activity recovered more fully from the last recession. Between 2016 and 2019, median household income also rose by an additional \$370 (or 7 percent) for each highly active domain name host per 100 people in a county, controlling for other factors. Monthly data available for January 2019 to November 2020 demonstrates that a combination of higher levels of broadband subscription with domain site density led to lower unemployment in counties. In counties where at least 60 % of the population were connected, more domain name websites decreased unemployment rates. This was the case before the Covid shock and afterward.

Domain name hosts provide an example where commercial data can offer new insights in comparison with public sector sources (King, 2016). This data draws on the largest registrar of domains and accounts for half the market, but it does not cover the entire population. As a robustness check, we compared GoDaddy data with a measure of domain density using a sample drawn using WebCrawler. The correlation between the WebCrawler sample of domains from any source and GoDaddy data was extremely high (above 0.9) providing high confidence that our data is representative of the distribution of domains across the country. Future research will explore the economic impacts of this digital activity by industries where data allows.

Measuring and evaluating economic uses of the internet on a local scale are increasingly important as federal Covid-19 funding and infrastructure proposals in Congress promise historic investments in broadband infrastructure and adoption. Our research suggests that promotion of digital skills for commercial website use may benefit local economies, in addition to more widespread broadband adoption in the community. As the economy recovers, an online presence may facilitate new startups online as well as the return of businesses on Main Street. Even solo entrepreneurs or microbusinesses can improve their visibility, communications with customers, and access to broader markets through their websites. Domain name hosts are the footprint of a more digitally enabled economy that may offer strategies for addressing place-based economic disparities and inclusive growth in a diversity of communities.

Appendix

Appendix Table 1
Predicting Economic Prosperity (2019) and Change in Prosperity (2007–2019) US Counties (Full sample of US Counties)

	(1) 2019 Prosperity	(2) 2011–2019 Recovery	(3) 2019 Prosperity	(4) 2011–2019 Recovery
Broadband Subscriptions	0.7239** (0.0460)	0.3113** (0.0432)	0.7077** (0.0425)	0.2941** (0.0436)
Venture Density	0.3306* (0.1865)	0.4638** (0.2149)		
Highly Active Venture Density			2.2499** (0.7743)	2.2247** (1.0988)
Small Business Density	-3.1809** (0.5717)	-4.0208** (0.6848)	-3.4450** (0.5261)	-3.8732** (0.7052)
Percent Black	-0.1694** (0.0227)	-0.0410 (0.0267)	-0.1694** (0.0217)	-0.0312 (0.0257)
Percent Native American	0.0390 (0.0517)	0.0153 (0.0511)	0.0377 (0.0509)	0.0208 (0.0506)
Percent Asian	-0.1506** (0.0613)	-0.2078** (0.0753)	-0.1560** (0.0620)	-0.2290** (0.0764)

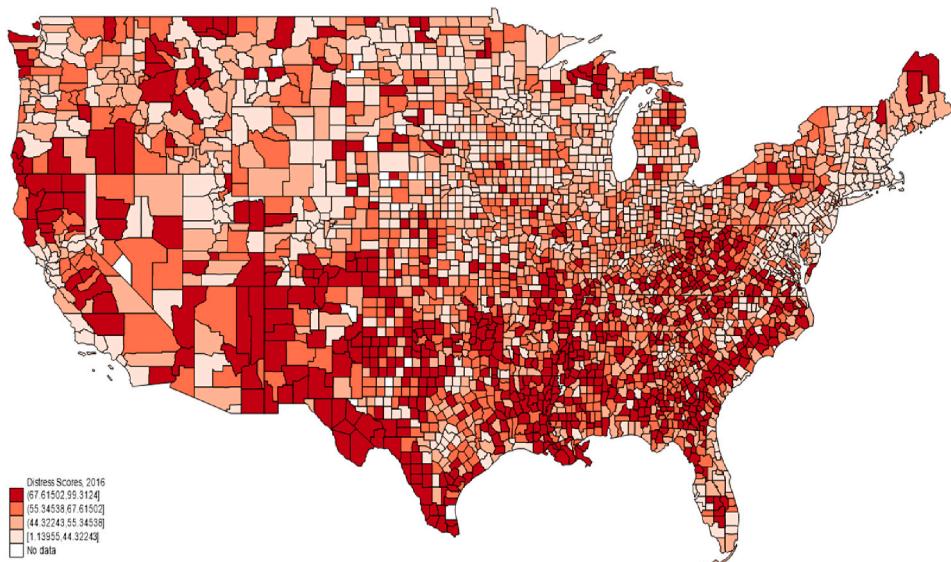
(continued on next page)

Appendix Table 1 (continued)

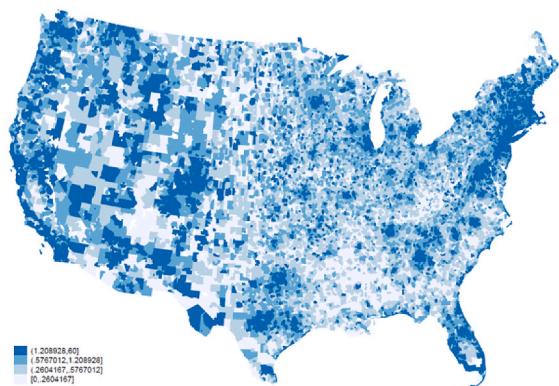
	(1) 2019 Prosperity	(2) 2011–2019 Recovery	(3) 2019 Prosperity	(4) 2011–2019 Recovery
Percent Hispanic	-0.1345** (0.0209)	-0.0675** (0.0249)	-0.1232** (0.0189)	-0.0512** (0.0225)
Percent Agriculture	-0.1599** (0.0755)	-0.3624** (0.1034)	-0.1389* (0.0755)	-0.3542** (0.1033)
Percent Construction	1.0354** (0.1820)	0.0870 (0.2442)	1.0020** (0.1813)	0.0306 (0.2492)
Percent Wholesale	0.5006 (0.3307)	0.6692 (0.4373)	0.5703* (0.3235)	0.6460 (0.4304)
Percent Retail	0.4184** (0.1566)	-0.0832 (0.1726)	0.3820** (0.1533)	-0.1285 (0.1695)
Percent Transportation	0.2683* (0.1577)	0.0402 (0.1713)	0.3166** (0.1458)	0.0832 (0.1695)
Percent Information	-0.1692 (0.4547)	1.1382* (0.6488)	-0.3233 (0.4259)	1.0036 (0.6230)
Percent Finance	0.4016** (0.1482)	-0.1055 (0.1928)	0.4084** (0.1421)	-0.0917 (0.1867)
Percent Professional	-0.4031** (0.1613)	0.0632 (0.1948)	-0.4171** (0.1539)	0.0898 (0.1847)
Percent Education	-0.5072** (0.0817)	-0.6004** (0.0882)	-0.4976** (0.0816)	-0.6040** (0.0937)
Percent Other	-0.2906 (0.3462)	-0.5303 (0.3572)	-0.2679 (0.3380)	-0.4532 (0.3534)
Percent Public	-0.0305 (0.1148)	-0.5781** (0.1543)	-0.0081 (0.1164)	-0.5767** (0.1593)
Percent High School	-0.1406** (0.0652)	-0.5798** (0.0806)	-0.0945 (0.0647)	-0.5387** (0.0788)
Percent College	0.5226** (0.0767)	-0.3742** (0.0904)	0.5404** (0.0744)	-0.3592** (0.0869)
Percent 20–34	-0.7372** (0.1290)	0.2237 (0.1562)	-0.7649** (0.1236)	0.2531* (0.1527)
Percent 35–54	0.0487 (0.1549)	0.2938* (0.1614)	0.0696 (0.1469)	0.3515** (0.1580)
Percent 55–64	-0.7120** (0.2492)	-0.2248 (0.2720)	-0.6934** (0.2334)	-0.1488 (0.2654)
Percent Employed	0.8585** (0.0623)	-0.2353** (0.0617)	0.8687** (0.0606)	-0.2353** (0.0639)
Constant	-21.0303** (9.5748)	43.8893** (10.3039)	-23.0153** (9.1720)	39.7850** (9.9613)
Observations	2958	2918	3101	3042

Appendix Table 2
Summary Statistics

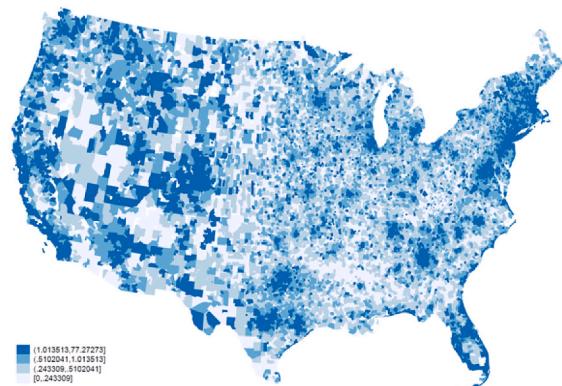
Variable	N	Mean	SD	Min	Max
Economic Prosperity Index	2958	42.12	21.40	0.47	96.47
Venture Density	3093	2.70	2.43	0.02	18.67
Highly Active Venture Density	3104	0.61	0.61	0.00	8.15
Broadband Subscriptions	2959	60.60	11.32	22.30	91.30
Percent Black	2959	8.85	14.29	0.00	86.18
Percent Native American	2959	1.54	6.30	0.00	81.31
Percent Asian	2959	1.20	2.25	0.00	34.15
Percent Hispanic	2959	8.84	13.45	0.00	98.96
Percent Agriculture	2959	6.62	6.97	0.00	59.30
Percent Construction	2959	7.17	2.24	0.70	22.50
Percent College	2959	20.72	9.01	4.90	73.70
Percent Wholesale	2959	2.41	1.16	0.00	20.50
Percent Retail	2959	11.48	2.29	1.40	30.90
Percent Transport	2959	5.42	1.94	0.00	22.50
Percent Info	2959	1.42	0.81	0.00	15.80
Percent Finance	2959	4.61	1.87	0.00	20.60
Percent Prof	2959	6.78	3.07	0.00	28.70
Percent Educ	2959	23.16	4.46	9.10	45.50
Percent Other	2959	4.69	1.18	0.00	11.50
Percent Public	2959	5.51	3.04	0.00	34.00
Percent Millennial	2959	18.02	3.81	6.10	43.90
Percent Gen X	2959	25.24	2.35	13.40	34.10
Percent Baby Boomer	2959	23.91	4.31	6.20	49.90



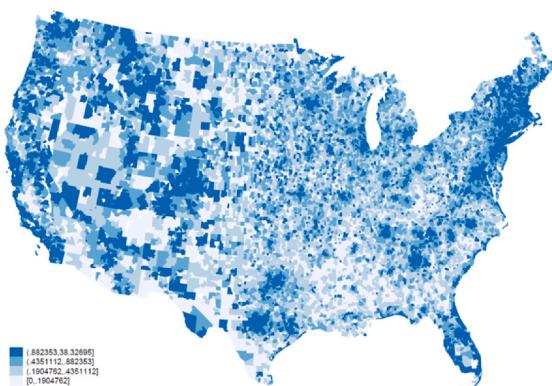
Appendix Fig. 1. Distressed Communities Index by County, 2016 ([Economic Innovation Group, 2018](#)). Source: Authors' map based on raw data from [Economic Innovation Group \(2018\)](#).



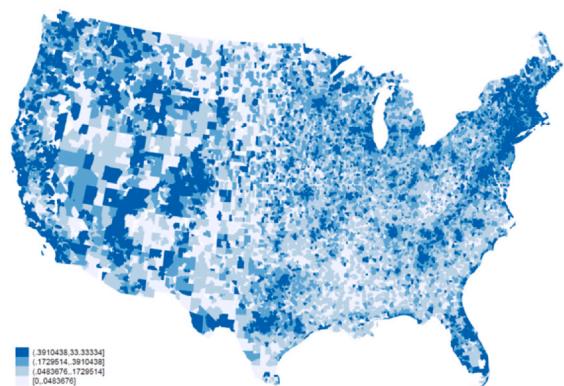
Cluster 1 – low activity



Cluster 2 – low/moderate activity

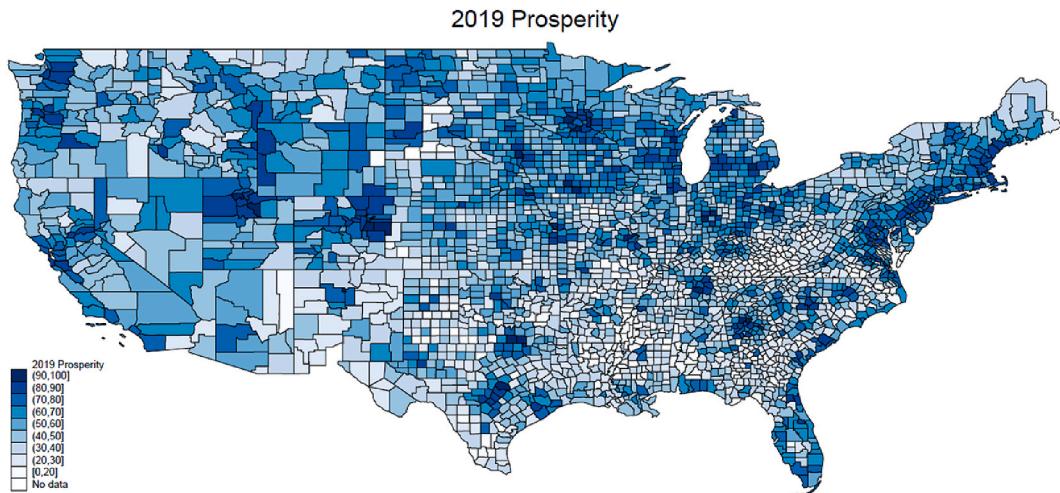


Cluster 3 – moderate/high activity



Cluster 4 – high activity

Appendix Fig. 2. Density of Domain Name Hosts by Activity Level and Zip Code, Nov 2018.



Appendix Fig. 3. 2019 Economic Prosperity Index for US Counties.

References

- Abraham, K. G., Haltiwanger, J. C., Sandusky, K., & Spletzer, J. R. (2018). *Measuring the gig economy: Current knowledge and open issues*. NBER Working Paper 24950. Retrieved from <https://www.nber.org/papers/w24950>.
- Abrardi, L., & Cambini, C. (2019). Ultra-fast broadband investment and adoption: A survey. *Telecommunications Policy*, 43(3), 183–198.
- Akerman, A., Gaarder, I., & Mogstad, M. (2015). The skill complementarity of broadband internet. *Quarterly Journal of Economics*, 130(4). <https://doi.org/10.1093/qje/qjv028>
- American Community Survey. (2019). Retrieved from <https://www.census.gov/programs-surveys/acs/>.
- Antonizzi, J., & Smuts, H. (2020). The characteristics of digital entrepreneurship and digital transformation: A systematic literature review. *Responsible Design, Implementation and Use of Information and Communication Technology*, 12066, 239–251.
- Atasoy, H. (2013). The effects of broadband internet expansion on labor market outcomes. *ILR Review*, 66(2), 315–345.
- Auger, P. (2005). The impact of interactivity and design sophistication on the performance of commercial websites for small businesses. *Journal of Small Business Management*, 43(2), 119–137.
- Bartels, L. B. (2008). *Unequal democracy: The political economy of the new gilded age*. Princeton, NJ: Princeton University Press.
- Bertschek, I., Brügelauer, W., Huschelrath, K., Kauf, B., & Niebel, T. (2016). *The economic impacts of telecommunications networks and broadband internet: A survey*. Discussion Paper (No. 16-056). ZEW Centre for European Economic Research. Retrieved from <http://ftp.zew.de/pub/zew-docs/dp/dp16056.pdf>.
- Blackwell, M., Iacus, S., King, G., & Porro, G. (2009). cem: Coarsened exact matching in Stata. *STATA Journal*, 9(4), 524–546.
- Boulianne, S. (2009). Does internet use affect engagement? A meta-analysis of research. *Political Communication*, 26(2), 193–211.
- Brynjolfsson, E., & Saunders, A. (2010). *Wired for innovation: How information technology is reshaping the economy*. Cambridge, MA: MIT Press.
- Chetty, R., Friedman, J. N., Hendren, N., & Stepner, M. (2020). *The economic impacts of COVID-19: Evidence from a new public database built from private sector data* (September). Retrieved from https://opportunityinsights.org/wp-content/uploads/2020/05/tracker_paper.pdf.
- Chetty, R., Grusky, D., Hell, M., Hendren, N., Manduca, R., & Narang, J. (2017). The fading American dream: Trends in absolute income mobility since 1940. *Science*, 356(6336), 398–406.
- Chetty, R., & Hendren, N. (2017). The impacts of neighborhoods on intergenerational mobility II: County-level estimates. In *Equality of opportunity project*. Stanford University. Retrieved from http://www.equality-of-opportunity.org/assets/documents/movers_paper2.pdf.
- Economic Innovation Group. (2018). *New study shows how recovery from the great recession transformed the geography of U.S. economic well-being*. Retrieved from <https://eig.org/news/new-study-shows-how-recovery-from-the-great-recession-transformed-the-geography-of-u-s-economic-well-being>.
- Falck, O. (2017). Does broadband infrastructure boost employment? *IZA World of Labor*, 2017, 41.
- Feldman, M. (2014). The character of innovative places: Entrepreneurial strategy, economic development, and prosperity. *Small Business Economics*, 43, 9–20. <https://doi.org/10.1007/s11187-014-9574-4>
- Florida, R. (2017). *The new urban crisis: How our cities are increasing inequality, deepening segregation, and failing the middle class—and what we can do about it*. New York, NY: Basic Books.
- Forman, C., Goldfarb, A., & Greenstein, S. (2012). The internet and local wages: A puzzle. *The American Economic Review*, 102(1), 556–575.
- Fortunato, M. W. P., & Alter, T. (2015). Community entrepreneurship development: An introduction. *Community Development*, 46(5), 444–455. <https://doi.org/10.1080/15575330.2015.1080742>. Retrieved from.
- Frey, W. H. (2018). February 7). Where do the most educated millennials live? Are they living in the Amazon HQ2 finalist places? *The avenue*. Retrieved from <https://www.brookings.edu/blog/the-avenue/2018/02/06/where-do-the-most-educated-millennials-live/>.
- Gallardo, R., Whitacre, B., Kumar, I., & Upendram, S. (2020). Broadband metrics and job productivity: A look at county-level data. *Annals of Regional Science*. <https://doi.org/10.1007/s00168-020-01015-0>. Retrieved from.
- Giannone, E. (2017). *Skilled-biased technical change and regional convergence*. University of Chicago. Retrieved from http://home.uchicago.edu/~elisagiannone/files/JMP_ElisG.pdf.
- Giones, F., & Brem, A. (2017). Digital technology entrepreneurship: A definition and research Agenda. *Technology Innovation Management Review*, 7(5). Available at SSRN: <https://ssrn.com/abstract=2984542>.
- Greenstein, S., Goldfarb, A., & Forman, C. (2018). How geography shapes—and is shaped by—the internet. In G. Clark, M. Feldman, M. Gertler, & D. Wojcik (Eds.), *The new Oxford Handbook of economic geography* (pp. 269–285). Oxford: Oxford University Press.
- Gupta, S., Nguyen, T. D., Rojas, F. L., Shayam, R., Lee, B., Bento, A., Simon, K., & Wing, C. (2020). *Tracking public and private responses to the COVID-19 epidemic: Evidence from state and local government actions*. NBER Working Paper No 27027 (April). Retrieved from <https://www.nber.org/papers/w27027>.
- Hasbi, M. (2017). Impact of very high-speed broadband on local economic growth: Empirical evidence. Retrieved from <https://www.econstor.eu/bitstream/10419/168484/1/Hasbi.pdf>.

- Hendrickson, C., Muro, M., & Galston, W. A. (2018). *Countering the geography of discontent: Strategies for left-behind places*. Brookings Institution. Retrieved from <https://www.brookings.edu/research/countering-the-geography-of-discontent-strategies-for-left-behind-places/>.
- Holt, L., & Jamison, M. (2009). Broadband and contributions to economic growth: Lessons from the US experience. *Telecommunications Policy*, 33(10–11), 575–581.
- Jackson, J., Lee, J., Lin, M., & Ratnatunga, M. (2019). *Best performing cities 2018: Where America's jobs are created and sustained*. Milken Institute. Retrieved from <https://www.milkeninstitute.org/reports/best-performing-cities-2018-where-americas-jobs-are-created>.
- Jayakar, K., & Park, E. A. (2013). Broadband availability and employment: An analysis of county-level data from the National Broadband Map. *Journal of Information Policy*, 3, 181–200.
- Kemeny, T., & Storper, M. (2015). Is specialization good for regional economic development? *Regional Studies*, 49(6), 1003–1018.
- Kim, Y., & Orazem, P. F. (2012). *Broadband internet and firm entry: Evidence from rural Iowa*. Working Paper #12026. Retrieved from https://lib.dr.iastate.edu/econ_las_workingpapers/79/.
- King, G. (2016). Preface: Big data is not about the data! In R. M. Alvarez (Ed.), *Computational social science: Discovery and prediction*. Cambridge: Cambridge University Press. Copy at <https://j.mp/2nSwPwD>.
- Kolko, J. (1999). The death of cities? The death of distance? Evidence from the geography of commercial internet usage. <http://citeseerx.ist.psu.edu/viewdoc/download?jsessionid=DCB82B557E04DAE9901A46E805C328C9?doi=10.1.1.35.3217&rep=rep1&type=pdf>.
- Kolko, J. (2012). Broadband and local growth. *Journal of Urban Economics*, 71(1), 100–113. <https://doi.org/10.1016/j.jue.2011.07.004>
- Krause, E., & Sawhill, I. (2018). *Seven reasons to worry about the American middle class*. Brookings Institution. Retrieved from <https://www.brookings.edu/blog/social-mobility-memos/2018/06/05/seven-reasons-to-worry-about-the-american-middle-class/>.
- Lehr, W. H., Osorio, C., Gillett, S. E., & Sirbu, M. A. (2006). *Measuring broadband's economic impact. National technical assistance, training, research and evaluation project (No. 99-07-13829)*. Prepared for the U.S. Department of Commerce, Economic Development Administration. Retrieved from http://cfp.mit.edu/publications/CFP_Papers/Measuring_bb_econ_impact-final.pdf.
- Mack, E. A. (2014). Broadband and knowledge intensive firm clusters: Essential link or auxiliary connection? *Papers in Regional Science*, 93(1), 3–29.
- Mack, E. A., & Faggian, A. (2013). Productivity and broadband: The human factor. *International Regional Science Review*, 36(3), 392–423. <https://doi.org/10.1177/0160017612471191>. Retrieved from.
- Moretti, E. (2012). *The new geography of jobs*. New York, NY: Mariner Books.
- Mossberger, K., Tolbert, C. J., & LaCombe, S. (2021). *Choosing the future: Technology and opportunity in communities*. NY: Oxford University Press.
- Moss, M. L., & Townsend, A. (1997). Tracking the net: Using domain names to measure the growth of the Internet in U.S. cities. *Journal of Urban Technology*, 4(3), 47–60. <https://doi.org/10.1080/10630739708724566>
- National Telecommunications and Information Administration/National Science Foundation. (2017). *The national broadband research agenda* (Vol. 2017). Key priorities for broadband research and data. Retrieved from <https://www.ntia.doc.gov/files/ntia/publications/nationalbroadbandresearchagenda-jan2017.pdf>.
- Rose, S. J. (2018). *How different studies measure income inequality in the US: Piketty and company are not the only game in town*. Urban Institute. Retrieved from https://www.urban.org/sites/default/files/publication/99455/how_different_studies_measure_income_inequality_0.pdf.
- Taglang, K. (2021). *Show us the money: Federal broadband support during the Covid-19 pandemic*. Benton Institute. April 23, 2021. Retrieved from <https://www.benton.org/blog/show-us-money-federal-broadband-support-during-covid-19-pandemic>.
- Torpey, E., & Hogan, A. (2016). *Working in a gig economy*. U.S. Bureau of Labor Statistics, United States Department of Labor. Retrieved from <https://www.bls.gov/careeroutlook/2016/article/what-is-the-gig-economy.htm>.
- United States Census Bureau. North American industrial classification system (NAICS). Retrieved from. <https://www.census.gov/eos/www/naics/>.
- United States Census Bureau –. (2018). *Economic Census*. Retrieved from <https://www.census.gov/programs-surveys/economic-census.html>.
- Whitacre, B., Gallardo, R., & Strover, S. (2014). Broadband's contribution to economic growth in rural areas: Moving towards a causal relationship. *Telecommunications Policy*, 38(11), 1011–1023.