Broadband Internet and Business Activity

Richard Beem[†]

Job Market Paper

September 29, 2021

Abstract

Does the diffusion of broadband Internet enhance business activity and entrepreneurship? Does the narrowing of the Digital Divide spur rural economic growth? To answer these questions, I empirically examine one of the Federal Communication Commission's largest broadband deployment programs — Phase II of the Connect America Fund. Exploiting plausibly exogenous variation in the deployment of wired broadband connections during 2015-18, my county-level difference-in-differences results show persistent gains in the number of firms, establishments, and entrepreneurs, as well as higher employment levels and average annual wages among treated counties. Firm growth is driven by small, young, and rural firms. To assess the social welfare implications of this program, a cost-benefit analysis reveals that the benefits from CAF II outweigh the costs by a factor of 42.

JEL Classifications: D22, H25, R11

Keywords: Broadband, business activity, subsidies, FCC, Connect America Fund

[†]Beem: Economics PhD Candidate, Department of Economics and Boyd Center for Business and Economic Research (BCBER), The University of Tennessee, 916 Volunteer Boulevard, Knoxville, Tennessee, 37996. Email: rbeem1@vols.utk.edu. I would like to thank Don Bruce, Matthew Harris, Maria Padílla-Romo, David Agrawal, Marianne Wanamaker, Matthew Murray, LeAnn Luna, Jim Alm, and Eunsik Chang for helpful comments and feedback on previous drafts. This paper benefited greatly from discussions with seminar participants at the 2020 and 2021 UTK Economics Department's Brown Bag Workshops, the 2020 Southern Economic Association's (SEA) annual meeting, the 2021 Eastern Economic Association's (EEA) annual meeting, the 2021 Canadian Economic Association's (CEA) annual meeting, and the 2021 International Institute of Public Finance (IIPF) annual congress.

1 Introduction

Broadband Internet is an essential productivity-enhancing technology (Akerman et al., 2015). Since the signing of the Telecommunications Act of 1996, the broadband industry has invested more than $1.7 trillion in broadband infrastructure (U.S. Telecom Association, 2019).[1] While ensuring that households and businesses have access to broadband-capable networks remains a policy objective of the Federal Communications Commission (FCC), roughly 31 percent of rural America was left disconnected in 2015, while urban America reported near-universal connectivity (FCC, 2015). This Digital Divide continues to serve as a market friction, affecting the start-up, operation, and expansion of businesses located beyond the borders of dense urban areas (Ivus and Boland, 2015).[2]

Despite the growing importance of broadband in today's economy, the existing literature paints an incomplete picture of its causal relationship with measures of business activity and entrepreneurship. The main hurdle researchers face has been empirical. Historically, broadband-capable networks have been deployed in an endogenous manner, with broadband first emerging in areas of greatest ability to pay.[3] Correlation between the timing and extent of broadband diffusion and underlying trends in business activity have made it challenging to isolate exogenous variation in the availability and adoption of broadband-capable networks.[4] Although prior work offers associative evidence from the examination of small-scale broadband deployment programs (Kandilov and Renkow, 2010, 2020; Shideler and Badasyan, 2012) or those that were conduced during the infancy stage

[1]For comparison, the Federal Aid Highway Act of 1956 generated an investment of roughly $500 billion in creating the interstate highway system (Infrastructure Report Card, 2016).

[2]Especially for small businesses, the GAO (2014a) found that broadband service improves production efficiency and leads to more streamlined operations. Broadband also reduces search frictions (Brown and Goolsbee, 2002; Kroft and Pope, 2014), reduces unemployment spells (Kuhn and Mansour, 2014), and allows businesses to fill vacancies faster (Bhuller et al., 2019).

[3]Much like the diffusion of rural electrification in the 20th century, remote areas with mountainous terrain, low-income populations, and higher deployment costs have been slow to receive broadband (GAO, 2014b; TACIR, 2017).

[4]In measuring the impact of broadband on economic growth, Czernich et al. (2011) use the existing telephone and cable TV infrastructure in OECD countries as an instrument for broadband diffusion, while Akerman et al. (2015) exploit the gradual rollout of broadband access points in Norway to examine the skill complementarity of broadband. In other work, researchers have used instrumental variables (IV) methods to isolate exogenous variation in broadband availability (Agrawal, 2021; Andersen et al., 2012; Bertschek et al., 2013).

of broadband's commercial rollout (Kim and Orazem, 2016), a rigorous analysis of broadband availability, firm dynamics, and the Digital Divide remains absent from the literature.[5]

In this paper, I present the first comprehensive analysis of one of the FCC's largest broadband deployment programs — Phase II of the Connect America Fund (CAF II) — by identifying the causal effect of broadband availability on measures of business activity and entrepreneurship.[6] The CAF II program, which operated during the 2015-20 period, issued \$1.675 billion per year in broadband installation subsidies to telecom providers for the purpose of installing wired broadband connections where connections did not previously exist. In addition to being large in denomination, this program reached 91 percent of counties, facilitating the connection of nearly 3.4 million households and small businesses by the end of 2019.[7] Exploiting this large-scale broadband supply shock, I examine whether greater broadband availability leads to (persistent) growth in the number of firms and establishments, employment levels, the average annual wage, the number of entrepreneurs, and annual entrepreneurial revenues.

By exploiting the unique features of the CAF II program, I am able to isolate plausibly exogenous variation (across geography and time) in the gradual rollout of wired broadband connections. Unique to this deployment program, the locations considered eligible for broadband installation subsidies were not endogenously chosen by telecom providers, but selected objectively by an engineering cost model developed by the FCC's Wireline Competition Bureau and the network modeling company, CostQuest. This exogenous allocation of broadband installation subsidies is an empirical hallmark of the CAF II program and thus provides a unique quasi-experimental research setting. The design and implementation of

[5] In some prior work, broadband availability has been measured using data from the FCC's Form 477 database (Atasoy, 2013; Kolko, 2012; Whitacre et al., 2014), but these data often overstate true broadband availability especially in large census blocks where a provider might only offer broadband service to a single household (Kolko, 2010).

[6] Since the signing of the Telecommunications Act of 1996, the definition of broadband has undergone several meaningful changes. The most-recent update to the benchmark occurred in 2015, requiring download/upload speeds of 25/3 Mbps. However, the CAF II program, which emerged out of the FCC's 2011 plan to modernize universal voice and broadband service, required that telecom providers offer broadband service with download/upload speeds of at least 10/1 Mbps. Going forward, this paper defines broadband according to this benchmark.

[7] Author's calculation using data from the Connect America Fund Broadband Map.

this program, unlike previous broadband deployment efforts in the United States, essentially eliminated the possibility of endogenenous selection into already expanding communities.

My empirical investigation begins with a standard two-way fixed effect (TWFE) design that exploits the staggered rollout of broadband into eligible counties. I follow Goodman-Bacon et al. (2019) and perform a decomposition of my baseline results to determine the extent to which negative weights impact the TWFE estimate. Although employment and entrepreneurial revenues appear to be the only outcome variables affected by greater broadband availability, the possibility of time-varying treatment effects suggests that the TWFE parameters might be misleading and biased downward (Goodman-Bacon, 2021; de Chaisemartin and d'Haultfoeuille, 2020; Sun and Abraham, 2020). Therefore, I employ the estimator developed by de Chaisemartin and d'Haultfoeuille (2020), which produces estimates that are robust to staggered treatment designs, to augment the principal analysis.

The difference-in-differences (DID) results, which allow for treatment-effect heterogeneity across time, reveal that greater broadband availability causes statistically significant gains in business activity and entrepreneurship. Three years after the initial installation of broadband, treated counties report 2.7 percent additional firms, 2.8 percent additional establishments, and 3.3 percent higher employment levels. Broadband-induced gains also translate into higher average annual wages (up 1.7 percent) and more entrepreneurs (up 1.9 percent).

Delving into several sources of heterogeneity, I unpack the baseline result by urbanicity to determine whether greater broadband availability disproportionately impacts rural communities. This allows for an empirical investigation of the core-periphery model (Krugman, 1991; Fujita et al., 1999) and whether agglomeration effects are present in this setting. In addition, I explore how greater broadband availability impacts firms differentially by firm age to better understand broadband's linkage with the recent findings of declining business dynamism (Haltiwanger, 2012; Fort et al., 2013; Decker et al., 2014; Pugsley and Sahin, 2019; Davis and Haltiwanger, 2019; Karahan et al., 2019; Haltiwanger, 2021), and explore how firm size impacts firm growth. I find that rural firms, small firms with fewer than 500 employees, and young firms no more than five years old drive the baseline result, recording the strongest and most-persistent responses to greater broadband availability.

Several robustness checks support the causal interpretation of the baseline results. First, I determine that treatment status and timing are effectively random during the 2015-18 period. Next, I determine that *unobservable* determinants of business activity and entrepreneurship are also unable to predict treatment status, strengthening the claim of broadband-deployment exogeneity. In recognition of the fact that the build-out of broadband in a county could spur local business growth, I test for the presence of endogeneity due to reverse causality and conclude that this is not problematic in my empirical setting. Finally, I recognize that with treatment-effect heterogeneity across time, endogenous migration could result in materially different county populations post treatment. I find that the household income characteristics of treated counties remain unaffected in the years that follow the initial installation of broadband. In other words, migration into and out of treated counties is not materially changing the composition of households.

In attempting to understand the social welfare consequences of this program, I perform a cost-benefit analysis. When given the option, I err on the side of overestimating broadband deployment costs and underestimating broadband-induced benefits. I consider both the CAF II program's external costs (i.e., universal service fund (USF) contributions) and internal costs (i.e., out-of-pocket deployment costs incurred by telecom providers). On the benefit side, I quantify the expected revenues gained from new broadband subscriptions and the present discounted earnings of newly employed individuals. Taken together, the CAF II program's benefits measure $242 million, outweighing the program's costs by a factor of 42.

This paper makes several contributions to the literature studying the economic effects of broadband on economic growth (Roller and Waverman, 2001; Czernich et al., 2011; Whitacre et al., 2014), labor markets (Akerman et al., 2015; Hjort and Poulsen, 2019; Atasoy, 2013; Ivus and Boland, 2015; De Stefano et al., 2014; Briglauer et al., 2019), and firm activity (Kandilov and Renkow, 2010, 2020; Shideler and Badasyan, 2012; Kim and Orazem, 2016; Bertschek et al., 2013; Falck et al., 2016). First, I complement the existing literature by focusing my empirical analysis around the quasi-experimental research setting generated by the CAF II broadband deployment program. The unique features of this program provide a framework for estimating causal effects. Unlike previous broadband deployment efforts, the selection of broadband-eligible households and small businesses was accomplished objectively by an engineering cost model. Naturally, this empirical setting offers a structural break in the way broadband installation subsidies were allocated to telecom providers. Second, my finding that rural counties report stronger business activity growth compared to urban counties stands in contrast to the theoretical predictions of the core-periphery model and informs the existing literature offering competing evidence on broadband deployment and agglomeration effects (Zuo, 2019; Atasoy, 2013; Ivus and Boland, 2015). While the industrialized "core" does benefit, the "periphery" appears to be the largest beneficiary. Finally, my results address the current deployment of broadband. In comparison to other studies that examine the economic effects of broadband during its infancy stage, this paper focuses on the current generation of broadband diffusion.

2 Background: The CAF II Program

The CAF II broadband installation program is unique from past broadband deployment efforts in that it represents a structural break in *how* households and businesses are offered broadband service. In the past, telecom providers endogeneously expanded broadband-capable networks based on expected network profitability and local economic characteristics (GAO, 2014b; TACIR, 2017). Empirically, this selection has prohibited researchers from estimating the causal effect of broadband on economic outcomes. With the CAF II program, the eligibility of households and businesses was determined objectively by an engineering cost model developed by the FCC's Wireline Competition Bureau and network modeling company, CostQuest.

The model first identified the exact locations of households and small businesses to determine where and to what extent broadband markets existed. Next, the network topology was designed. The model assembled and designed an efficient wireline network using existing spatial realities (e.g., road systems and relevant terrain). Terrain characteristics such as depth to bedrock, depth to water, rock density, and soil type were all considered in the optimization process. Lines were then connected from the end user to the central office to form the most efficient wireline network (i.e., the network that minimized turns, following existing roadways). Then, the model computed the deployment and operational costs necessary to create and service such a network. Computed network costs included engineering expenses, materials, and construction labor. Finally, the requisite support amount was calculated for the households and small businesses lacking service.[8]

High-cost census blocks, with a computed monthly cost-per-location between $52.50 (benchmark) and $198.60 (threshold), that were not served by subsidized or unsubsidized wireline competitors and were not subject to the Rural Broadband Experiments (RBE) qualified for CAF II subsidies. Telecom providers were then offered model-based support in exchange for a commitment to serve all eligible locations in their service territories.[9] If support amounts were declined, eligible census blocks were to receive broadband connections based on a competitive bidding process.[10] Beginning in 2015, the annual disbursement of CAF II funds measured $1.675 billion, targeting more than 4 million households and businesses.

3 Empirical Methods and Data

To uncover the causal effect of broadband availability on business activity and entrepreneurship, I leverage plausibly exogenous variation in the rollout of wired broadband connections across 2,200 counties in the United States during the 2015-18 period. In particular, identification in this analysis requires exogeneity in both

[8]In recent work, Alm et al. (2020) examine the factors influencing the designation of Qualified Opportunity Zones (QOZ) and find that although the selection process depends highly on the state of the local economy (e.g., unemployment rate, welfare recipients, median income), political factors also play a significant role. In the context of this broadband deployment program, it appears that the selection of eligible locations did not conform to this pattern.

[9]"In meeting its obligation to serve a particular number of locations in a state, an incumbent that has accepted a state-level commitment may choose to serve some census blocks with costs above the highest cost threshold instead of eligible census blocks (with lower costs), provided that it meets the public interest obligations in those census blocks, and provided that the total number of unserved locations and locations covered is greater than or equal to the number of locations in the eligible census blocks." FCC DA 14-534.

[10]In August 2018, the CAF II auction allocated $1.488 billion in annual broadband installation subsidies to telecomm providers to be dispersed over a 10-year period across 45 states encompassing more than 700,000 locations. In total, 103 providers were offered support. See https://docs.fcc.gov/public/attachments/DOC-353840A1.pdf. For a more detailed description of the auction process see https://www.bbcmag.com/law-and-policy/the-connect-america-fund-reverse-auction.

the *eligibility* of counties and the *timing* of broadband deployment, following Deshpande and Li (2019) and Baum-Snow and Lutz (2011). I remove from my sample ineligible counties that are never treated from my panel.[11] Departing from previous studies that have utilized broadband deployment data from the FCC's Form 477 database (Atasoy, 2013; Kolko, 2012; Whitacre et al., 2014), I examine the recent CAF II program, which offers a unique quasi-experimental research setting.[12] Importantly, individual locations (i.e., households and small businesses) without broadband access were classified as eligible to receive broadband installation subsidies based on the objective determination of an engineering cost model. To identify broadband's causal effect on business activity, I exploit this structural break in broadband allocation. And because my county-year panel offers extensive geographic and temporal variation in the *availability* of broadband-capable networks, inference is relatively isolated from external validity concerns. In what follows, I assess whether this increase in local connectivity translates into gains in business activity and whether broadband-induced growth persists.

3.1 *Identification Strategy*

To assess whether broadband availability has a causal relationship with various measures of business activity, I first employ a TWFE staggered treatment design. In this setting, business activity outcomes in newly treated counties are compared

[11] There are several reasons why some counties were considered ineligible for CAF II subsidies. If a county did not contain a single location with an estimated connection cost *below* the extremely high-cost threshold (EHCT) and *above* the funding benchmark, it was considered ineligible. Counties were also ineligible if they were served by unsubsidized or subsidized wireline competitors or participated in the Rural Broadband Experiments program. Relative to those that contain eligible locations, ineligible counties are sparsely populated and concentrated in the Midwest and Mountain regions. Their population characteristics and potentially diverging business activity trends suggest that they would not be a well-suited control group.

[12] As noted by Kolko (2010) and others, Form 477 broadband data often overstate true broadband availability, even at the more granular geographic levels in the United States. Twice per year, telecom providers are required to submit the census blocks in which they "can or do offer broadband-capable service." Especially for larger census blocks, broadband may be available, but only for a small subset of the population. In March 2020, the Broadband Deployment Accuracy and Technological Availability (DATA) Act was signed into law. The DATA Act requires the FCC to alter the way broadband availability data are collected from telecom providers, to publish granular broadband availability maps subject to independent audits, and establish the Broadband Serviceable Location Fabric, a map that would identify where broadband infrastructure is needed.

to outcomes in already-treated counties and eventually-treated counties, following Goodman-Bacon (2021). I estimate a regression equation of the following form:

$$y_{ct} = \gamma_c + \theta_t + \beta \times 1\{broadband\}_{ct} + \eta X_{ct} + \epsilon_{ct} \tag{1}$$

where y_{ct} is the outcome of interest for county c in year t. Specifically, I explore the natural log of six main outcome variables. These include the number of firms, the number of business establishments, the level of employment, the average annual wage, the number of entrepreneurs, and annual entrepreneurial revenues. I include county fixed effects (γ_c) to control for time-invariant determinants of business activity and year fixed effects (θ_t) to control for aggregate trends. I also include a vector of county-time-varying controls (X_{ct}) to mitigate potential omitted variable bias. This vector accounts for the FCC's concurrent broadband deployment efforts and cell tower availability.[13] My broadband availability indicator is defined as $1\{broadband\}_{ct} = eligible \times treated$. It equals 1 in the year in which an eligible county begins receiving wired broadband connections from telecom providers using CAF II installation subsidies, remaining 1 in the years that follow. Finally, ϵ_{ct} is the error term. The coefficient of interest (β) measures the effect of increased broadband availability on various measures of business activity and entrepreneurship. The log-linear nature of this specification means that β is interpreted as a semi-elasticity in percent terms; turning on the treatment indicator yields a percentage change in the outcome of interest. The estimation sample spans 2011-18. Standard errors are clustered at the county level.

Identification requires that the timing of treatment must be as good as random after controlling for county and year fixed effects (Bertrand et al., 2004) and that treated and comparison counties exhibit parallel trends in the years before treatment. The latter ensures that credible counterfactuals can be estimated. If, however, treatment effects are thought to vary over time, the weighted-average TWFE parameter estimated in the above equation may be misleading (Goodman-Bacon,

[13]Concurrent broadband expansion efforts conducted by the FCC include ACAM, or the Alternative Connect America Cost Model, which provides rate-of-return carriers with broadband installation subsidies in exchange for meeting defined broadband build-out obligations; Broadband Loop Support (CAF-BLS) provides support for the build-out of broadband and voice service; and the Rural Broadband Experiments (RBE) provides funding for the build-out of broadband in rural areas.

2021; de Chaisemartin and d'Haultfoeuille, 2020; Sun and Abraham, 2020).[14] As noted by Goodman-Bacon (2021), this does not imply a failure of the TWFE design, but does suggest that caution should be taken when attempting to summarize the average treatment effect (ATE) with a single TWFE parameter.

In my empirical setting, variation in treatment timing means that some counties receive broadband connections *earlier* than others. Table 1 summarizes the timing of treatment for the CAF II program. In 2015, telecom providers install broadband connections in 905 counties, roughly 32 percent of all eligible counties. This "treated share" increases over time, surpassing 76 percent by the end of 2018. In addition, broadband-induced gains in business activity are expected to grow over time. The path between broadband becoming *available* and broadband being *adopted* by households and small businesses may span several months or years. In addition, once broadband is fully adopted and utilized, it is not hard to imagine how the benefits might compound over time. For example, broadband has been shown to reduce job-search frictions and the length of unemployment spells (Kuhn and Mansour, 2014) and allows firms to fill vacancies faster (Bhuller et al., 2019). Both generate reductions in inefficiencies that have the potential to increasingly impact business activity in a positive manner.

Recognizing the plausibility that broadband's impact on business activity increases over time, I first explore the extent to which negative weights impact the TWFE parameters using the Bacon decomposition formula (Goodman-Bacon et al., 2019). Next, I allow treatment effects to be heterogeneous across time, employing the estimator developed in de Chaisemartin and d'Haultfoeuille (2020). Importantly, this estimator is valid even in the presence of negative weights and allows the identifying assumption of parallel pre-trends to be tested visually.[15] I estimate

[14]With variation in treatment timing, the TWFE parameter is a weighted average of all cross-group treatment effects (Goodman-Bacon, 2021). Negative weights arise when treatment effects vary over time (Goodman-Bacon, 2021; de Chaisemartin and d'Haultfoeuille, 2020). In this setting, the "control" group is composed of already-treated groups, thus changes in their treatment effects over time get subtracted from the overall TWFE estimate. This puts downward pressure on the average treatment effect defined by a single TWFE parameter.

[15]This test differs from the standard event-study pre-trend test, which Sun and Abraham (2020) have shown to be invalid in the setting with treatment-effect heterogeneity.

the following regression:

$$y_{ct} = \gamma_c + \theta_t + \underbrace{\sum_{\tau=0}^{\tau=3} \beta_{-\tau} \times 1\{broadband\}_{ct-\tau}}_{\text{Post-Treatment Effects}} + \underbrace{\sum_{\tau=1}^{\tau=4} \beta_{+\tau} \times 1\{broadband\}_{ct+\tau}}_{\text{Anticipatory Effects}} + \eta X_{ct} + (\zeta_s \times t) + \epsilon_{ct} \tag{2}$$

where the main difference from equation (1) is the addition of post-treatment and anticipatory (placebo) effects. In addition to the contemporaneous effect, I estimate three post-treatment effects to uncover whether broadband-induced gains in business activity grow over time, and four anticipatory effects to test for parallel trends in pre-treatment periods. I also include state-specific linear time trends $(\zeta_s \times t)$ to account for secular trends in the business environment that are specific to states but unrelated to the rollout of broadband under the CAF II subsidy program (Baum-Snow and Lutz, 2011; Lindo et al., 2018; Ohrn, 2019; Fox et al., 2020).[16] Similar to the interpretation of equation (1), the coefficients of interest $(\beta_0, ..., \beta_{-3})$ are interpreted as semi-elasticities. They represent the *current* percentage change in county-level business activity resulting from the *past* initial installation of wired broadband connections using CAF II installation subsidies.

3.2 *Defining Urban and Rural Counties*

To explore whether the relationship between increased broadband availability and business activity exhibits heterogeneity across the urbanicity spectrum, I classify counties as either urban or rural. To accomplish this, I turn to the United States Department of Agriculture's (USDA) rural-urban continuum codes (RUCC), which classify all counties in the United States according to their population density and

[16] As a practical matter, the inclusion of state-specific linear time trends does improve the comparability between treated and comparison counties. Consistent with Meer and West (2016) and Wolfers (2006), the magnitudes of the estimated treatment effects do respond to the inclusion of time trends. However, their inclusion rarely changes the statistical significance of post-treatment effects. Table A.2 reports the difference in ATEs estimated with and without state-specific linear time trends for all six outcome variables. After four years of treatment, the state-specific trends increase the treatment effects for firms, establishments, employment, average annual wage, and entrepreneurs by an average of 0.9 percentage point. Interestingly, the state-specific trends reduce the treatment effect for entrepreneurial revenue by 0.9 percentage point.

adjacency to a metropolitan statistical area (MSA).[17]

In Table A.1, I define urban counties as those with RUCC classifications of either 1, 2, 3, 4, or 6. This designation includes all metropolitan counties and those located tangentially to an MSA with urban populations of at least 2,500 persons. Consequently, rural counties are those with RUCC classifications of either 5, 7, 8, or 9. This amounts to completely rural counties and the remaining set of non-tangential "periphery" counties. Figure A.1 displays this geographic variation. This designation allows for an empirical investigation of the core-periphery model developed by Krugman (1991) and Fujita et al. (1999).[18] The basic intuition is that in a simplified world, with an industrialized "core" and agricultural "periphery," the one with a larger manufacturing labor force is seen as more attractive. This is because it boasts higher nominal wages and lower prices due to increased variety of locally manufactured goods.

With both urban and rural counties containing eligible locations, does the business-activity response of urban counties exceed that of rural counties, suggesting an agglomeration effect?[19] Conversely, do rural counties, which experience larger gains in connectivity as a result of the CAF II program, report larger and more persistent growth? In Section 4.3, I provide suggestive evidence of the relative importance of the agglomeration effect and treatment dosage in exploring urbanicity heterogeneity.

3.3 *Data and Summary Statistics*

In a major departure from previous studies that utilized Form 477 data from the FCC to examine the effects of increased broadband availability (Atasoy, 2013; Kolko, 2012; Whitacre et al., 2014), I make use of address-level broadband instal-

[17]For documentation on the RUCC classification process, see https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/.

[18]I thank Thiess Buettner for this suggestion.

[19]The presence of agglomeration effects has been documented in the literature. Kandilov and Renkow (2010) evaluate the USDA's Broadband Loan Program from the early 2000s and find that positive economic effects were driven by communities located closest to urban areas. Similarly, Kim and Orazem (2016) find that broadband availability influences the location decisions of new rural firms, with the effect being largest in more populated rural areas and those adjacent to metropolitan areas.

lation data from the Connect America Fund Broadband Map (CAF Map). This rich dataset documents where, when, at what speed, and from which telecom provider households and businesses receive wired broadband connections under the CAF II program.[20] Maintained by the Universal Service Administrative Company (USAC), broadband installation data collected from telecom providers are independently verified using a random sample of reported locations each year. I pair these data with the FCC's CAF II Final Eligible Areas Map to determine the number of eligible locations contained within each county and the denomination of subsidies distributed annually to participating telecom providers.

Table 2 presents summary statistics for several broadband measures for both urban and rural counties. There are several key takeaways. First, urban counties contain eight times as many price-cap locations (i.e., households and businesses located within telecom service territories) than rural counties. This is not a surprise since the urban designation contains metropolitan counties, which are densely populated. Figure A.2 shows that more than 40 percent of rural counties contain fewer than 5,000 price-cap locations. Second, the share of total locations that are eligible to receive broadband installation subsidies in rural counties is more than twice the share in urban counties. On average, nearly 19 percent of locations in rural counties are eligible for broadband installation subsidies compared to 9.3 percent in urban counties; half of urban counties report an eligibility share less than five percent (see Figure A.2). Third, the average broadband installation subsidy in rural counties is about 36 percent higher than in urban counties. Indeed, broadband-capable networks have been difficult to extend into unserved and underserved areas due to prohibitive construction and maintenance costs (GAO, 2014b; TACIR, 2017). Especially in rural counties, mountainous terrain makes extending fiber or cable technologies particularly costly; both of which must be buried underground or attached to elevated poles. This translates into larger broadband installation subsidies.

[20] Although the CAF II program was designed to extend broadband-capable networks to both households *and* businesses, telecom providers are not required to submit to the Universal Service Administrative Company the designation of newly connected locations. This was confirmed through email communication with the chief data administrator. The CAF Map also reports the build-out of broadband service under several alternative deployment programs. These include ACAM, CAF-BLS, RBE, and the CAF II auction.

In 2015, the CAF II program launched and ended the year having installed broadband connections at nearly 150,000 locations nationwide. Figure 1 displays the cumulative CAF II rollout of broadband (black) against the backdrop of the FCC's alternative deployment programs (gray). By 2019, more than 3.4 million households and businesses had received broadband connections from telecom providers that were issued CAF II installation subsidies. The geographic concentration of this rollout is depicted in Figure 2. Large portions of Maine, Wisconsin, Minnesota, Ohio, Kentucky, and Louisiana appear in dark blue, signifying large treatment dosages. Southern California and much of Northern Arizona also show large increases in connectivity, while many of the sparsely populated counties in the Midwest and Mountain regions report ineligibility.

Turning next to the task of quantifying business activity at the county level, I incorporate data from a variety of sources. Firm, establishment, and employment data come from the Business Dynamics Statistics (BDS), which provide annual measures of firm start-ups and shutdowns, establishment openings and closings, and employment creation and destruction.[21] Average annual wage data are from the Quarterly Census of Employment and Wages (QCEW), entrepreneurship data come from the Nonemployer Statistics (NES), and business application data are sourced from the Business Formation Statistics (BFS).[22]

Table 3 presents summary statistics for seven measures of business activity, again delineated by urbanicity. Not surprisingly, urban counties contain more firms, establishments, and entrepreneurs, employ more workers, and offer higher average annual wages than do rural counties. In addition, urban counties receive more business applications per year. Aside from the fact that urban counties are simply larger than rural counties, both types report skewness in business activity

[21] I choose to work primarily with BDS data as opposed to County Business Patterns (CBP) data for two primary reasons. First, the BDS data offer firm statistics; the CBP data only offer statistics for establishments and employment. Second, the redesigned BDS data describe the *dynamics* of firm activity, which I explore in Section 6. In exploring the CBP data, I find that the results for establishments and employment are not materially different than those generated using BDS data. If anything, the CBP results are slightly augmented. Relatively speaking, the BDS results are more conservative.

[22] To maintain a balanced panel during the 2011-18 period, I remove six counties that have county-year cells censored for confidentiality. These include three counties in Texas (King, Loving, and Hudspeth), two counties in Iowa (Ringgold and Adams), and one county in Pennsylvania (Sullivan).

measures. For example, urban counties contain an average of 2,766 firms, more than three times the median value of 761. For rural counties, this skewness is not as extreme, but still present, nonetheless. In recognition of the fact that few exceptionally large counties pull mean values to the upper end of the distribution, all measures of business activity that serve as outcome variables in equations (1) and (2) are transformed by taking the natural log.[23] This will help mitigate concerns that treatment effects are driven predominantly by scale and not the treatment itself.

4 Results

I first present the baseline TWFE results from estimating equation (1) as well as the Bacon decomposition (Goodman-Bacon et al., 2019) for each outcome variable. Next, I present the DID results from estimating equation (2), drawing comparisons with the TWFE results. Then, I explore sources of heterogeneity: firm size and urbanicity to better understand how broadband availability impacts business activity.

4.1 *Traditional TWFE Results*

Table 4 presents the results from estimating equation (1) describing the effect of greater broadband availability on six measures of business activity. Each regression includes county and year fixed effects as well as controls for concurrent broadband deployment programs and cell tower availability. The sample includes 2,511 counties during the 2011-18 period and standard errors are clustered at the county level.

The coefficient on $1\{broadband\}$ appears highly significant for two outcomes: employment and entrepreneurial revenues. The 0.0097 parameter for employment indicates that employment in counties receiving wired broadband connections from subsidized telecom providers increases by roughly 1 percent in the years following the initial installation of broadband. For the average treated county

[23]Because some small counties report zero business applications in some years, I add 1 to each outcome variable before taking the natural log. This helps preserve meaningful zeros in the business application data.

with payrolls measuring 45,254, this amounts to an additional 439 jobs. Entrepreneurial revenues increase a slightly stronger 1.3 percent, or an additional $5,680 for the average treated county. However, the number of firms and establishments, the average annual wage, and the number of entrepreneurs appear unaffected by increased broadband availability. It is noteworthy that these effects are smaller in magnitude compared to the coefficients for employment and entrepreneurial revenues.

To investigate the extent to which negative weights are driving down the *average* effect of increased broadband availability, I perform the Bacon decomposition (Goodman-Bacon et al., 2019) for each outcome variable. These results are presented in Table 5. For each measure of business activity (row), this table reports three treatment effects and their weights (columns) used to construct the ATE. The last column reports the overall ATE. The "timing groups" column captures comparisons between counties treated earlier versus later (and *vice versa*) in the sample, while the "never vs. timing" column compares treated counties with those that never receive treatment. When treatment effects vary over time, more weight attached to the "timing groups" estimate suggests that the ATE may be biased downward (Goodman-Bacon, 2021; de Chaisemartin and d'Haultfoeuille, 2020).

Indeed, Table 5 reveals that for each outcome variable, 62 percent of the ATE is derived from the *timing* of treatment between treated counties; the remaining 38 percent comes from the comparison between treated and never-treated counties. Figure 3 displays the decomposition of each ATE graphically. For each outcome, it is clear that the timing of treatment is driving the ATE. Especially if increased broadband availability is thought to *increasingly* affect business activity and entrepreneurship over time, an alternative estimation method must be employed to better understand the evolution of each outcome variable.

4.2 *Results with Treatment-Effect Heterogeneity Across Time*

To augment the above analysis, I estimate equation (2) using the estimator developed by de Chaisemartin and d'Haultfoeuille (2020). Importantly, this estimator allows treatment effects to be heterogeneous across time and remedies the complications that arise from negative weights in empirical settings with staggered treatment. Figure 4 presents results for all six measures of business activity and entrepreneurship graphically, with four placebo periods and four post-treatment periods. The sample spans 2011-2018 with the standard set of controls, linear state-specific trends, and standard errors clustered at the county level using 1,000 bootstrap replications.

Figure 4 shows that the effect of increased broadband availability is *not* constant over time. Instead, broadband-induced growth in business activity appears to compound in the years following the initial installation of wired broadband connections. Unlike the baseline TWFE result presented in Table 4, firms do appear to be increasingly responsive to connectivity gains at the county level. Indeed, the number of firms increases a statistically significant 0.4 percent in the first year, up nearly 2.7 percent after four years. Similar and slightly more pronounced growth paths exist for establishments and employment.[24] The latter result is consistent with Hjort and Poulsen (2019), who also report large positive employment effects.[25] Taken together, these results suggest that the average treated county reports an additional 57 firms, 70 establishments, and 1,484 jobs after four years as a result of the CAF II broadband deployment program.

Figure 4 also shows increasing gains in the average annual wage and the number of entrepreneurs. That the *average* annual wage increases monotonically during post-treatment years is consistent with the finding in Akerman et al. (2015) that skilled workers report larger wage gains relative to the wage declines reported by unskilled workers. Especially after a few years, treated counties also report significant gains in the entrepreneurial community. The average treated county reports nominal wage gains of $675 per year and 179 additional entrepreneurs after four years of increased broadband availability. However, entrepreneurial revenue growth does not persist beyond three years.

[24] Across all six outcomes, only two of 24 placebo effects appear statistically significant. In Table A.2, I compare the DID results with and without state-specific trends.

[25] Hjort and Poulsen (2019) find that the arrival of submarine Internet cables off the coast of Africa increases the probability of an individual being employed by at least 3.1 percent, up to 13.2 percent in some areas.

4.3 *Heterogeneous Results: Urbanicity*

To test for the presence of agglomeration effects, I augment my principal analysis by separately estimating effects for firms in rural and urban counties. While there is some evidence suggesting that rural areas record broadband-induced growth in excess of urban areas (Atasoy, 2013; Ivus and Boland, 2015), more-recent research suggests the opposite (Zuo, 2019). If broadband-induced growth experiences diminishing returns as initial levels of connectivity increase, then rural counties should record stronger growth. Recall that the share of eligible locations in rural counties (18.7 percent) is more than twice that in urban counties (9.3 percent). Figure 5 presents the results from estimating equation (2). Below each set of heterogeneous treatment effects, I include the difference between the results for either rural or urban firms and the baseline result for total firms.

While both rural and urban counties report significant and lasting firm growth, rural counties record stronger growth. Indeed, three years after initial broadband installations are made, firm growth in rural counties measures 3.3 percent, an effect nearly one-and-two-thirds the size of the effect in urban counties. The bar chart below also shows that rural-firm growth exceeds the overall ATE by 0.6 percentage point in $t + 3$; the opposite is true for urban-firm growth. Recall that the rural designation includes completely rural and periphery counties (i.e., those that do not share a border with an MSA). This result of stronger growth in rural "periphery" counties provides counter evidence for the canonical core-periphery model of economic geography developed by Krugman (1991) and Fujita et al. (1999) and empirical findings in Kandilov and Renkow (2010) and Kim and Orazem (2016). Put another way, this result suggests that treatment dosage (in rural counties) dominates any agglomeration effects at play.[26]

[26] In unreported results, I interact the treatment indicator with the share of locations eligible for CAF II broadband installation subsidies $1\{broadband\} = eligible \times treated \times eligible\ share$ and re-estimate equation (2). The results show significant and positive effects in the post-treatment years, suggesting that broadband availability matters more for counties starting out from a place of greater disconnectedness. This is especially true for rural counties, which report initial eligibility shares twice the size of urban counties.

4.4 *Heterogeneous Results: Firm Size*

Next, I explore whether firm size matters in determining the business activity response to increased broadband availability. In keeping with the Small Business Administration's (SBA) definition, I define small firms as those with fewer than 500 employees.[27] Small businesses are often referred to as the "backbone" of the American economy. In fact, more than 70 percent of start-up employment is generated by firms with fewer than 50 employees (Haltiwanger, 2012). Small firms, which are often more credit constrained (Fort et al., 2013), do not have the ability to scale operations like large firms. Gains in local connectivity could disproportionately impact small firms that are then able to expand into new and distant markets.

Figure 5 presents the results from estimating equation (2) for small and large firms. There are a few important results that require some discussion. First, small-firm growth increases monotonically and robustly since the initial installation of broadband. After four years, the number of small firms remains 2.8 percent above the pre-treatment level. And while large firms report parallel (and significant) gains through the third year, it appears that large-firm growth tapers off and loses significance in the fourth year. Consequently, large-firm growth remains 1.3 percentage points below the ATE for total firms by the fourth year. It is also important to note the distinction between small firms and young firms (Fort et al., 2013), the latter of which I investigate in Section 6.

5 Robustness Checks

To further support claims of causality and probe the identifying assumptions used in estimating equation (2), this section reports the results from a battery of robustness checks. First, I test whether treatment appears to be random by predicting both treatment *status* and treatment *timing* based on observable economic characteristics, following Deshpande and Li (2019).[28] To accomplish this, I exploit

[27] To maintain a balanced panel, I drop 51 counties (primarily in Montana, Nebraska, and Texas) that have incomplete or censored data during the 2011-18 period.

[28] Specifically, I estimate two equations for each treatment year (i.e., 2015, 2016, 2017, and 2018). The first is defined as $1\{broadband\}_c = \alpha + \zeta\Phi_c + \epsilon_c$ where $1\{broadband\}_c$ is the treatment indicator and Φ_c is a vector of current and lagged observable economic characteristics for county c. These

cross-sectional variation from 2,200 treated counties (treatment timing) and 3,102 total counties (treatment status) and present the results in Tables 6 and 7.

In Table 6, each column presents the results from regressing the treatment status on various observable economic characteristics. While some characteristics are significant predictors of treatment status in some years, no characteristic consistently (and significantly) predicts treatment status in all years.[29] It is also notable that the signs on the coefficients for each observable economic characteristic are not consistently positive or negative. For example, in 2017, counties with *higher* jobless rates, *fewer* cell towers, and *higher* average annual wages were more likely to receive broadband connections. In 2018, the reverse was true. Along with the lack of consistent significance across treatment years, this lack of consistency in the *sign* of coefficients suggests that the CAF II broadband rollout was essentially random.

Because my empirical setting includes counties that are not treated (but remain eligible) by the end of 2018, I also explore whether the timing of treatment is a function of observable economic characteristics. Table 7 presents the results from this exercise. Once again, no characteristic consistently predicts the year that counties are initially treated. These results suggest that treatment *timing* is effectively random, corroborating the institutional details of the CAF II program.

Next, I ask whether *unobservable* determinants of business activity are correlated with treatment status. Having shown that observable economic characteristics are not consistently correlated with treatment status or timing, the exogeneity of treatment could be called into question if unmeasured county-time-varying determinants of business activity are in fact correlated with treatment. To explore the potential severity of this concern, I first save the residuals from regressing each outcome variable on the treatment indicator and standard set of controls. Then, I plot the correlation between the residuals, which represent unobserved determinants of business activity, and the treatment indicator in an overlaid binscatter

include the jobless rate, cell tower availability, other broadband deployment programs, and the average annual wage. The second is defined as $treated\,year_c = \alpha + \zeta\Phi_c + \epsilon_c$ where $treated\,year_c$ is the year that telecom providers began installing broadband in county c. I include state fixed effects and cluster standard errors at the county level.

[29] In 2016 and 2018, the current and lagged average annual wage appear to be marginally significant.

displayed in Figure 6. Each of the linear fitted lines show that no statistically significant relationship exists; all fitted lines have slopes that are essentially zero.

It could also be the case that lagged business activity predicts treatment status. This source of endogeneity due to reverse causality could arise if telecom providers choose to extend broadband-capable networks into counties that experience strong business growth.[30] More firms, higher employment levels, and higher annual wages would certainly make a new territory more attractive to a telecom provider. To explore the extent to which potential reverse causality is a concern, I regress treatment status on lagged measures of business activity and present these results in Table 8. While the lagged number of entrepreneurs appears marginally significant, no measure of business activity significantly predicts treatment status.

Finally, because treatment effects vary over time, I explore whether treated counties benefit from economically different in-migrants. If treated counties also benefit from net-in-migration, are new residents sufficiently different than both out-migrants and non-movers? If in-migrants report higher household incomes they may also have better business acumen or have access to a greater array of financing channels that could translate into gains in business activity. To empirically test for this possibility, I gather migration data from the Internal Revenue Service's (IRS) SOI Tax Stats. I regress treatment status on three outcomes: (1) the difference between the number of in-migrant and out-migrant households, (2) the difference between in-migrant and out-migrant household adjusted gross income (AGI), and (3) the difference between in-migrant and out-migrant AGI per household. The results are presented in Table 9. The first column shows that treated counties are the recipient of 51 additional households, suggesting some degree of migration from untreated counties, similar to the finding in Kolko (2012). However, columns two and three show that treatment status has no significant effect on the difference in income characteristics of in-migrants and out-migrants. While broadband deployment results in small gains in county population, the difference in their financial positions remains unchanged.

[30] I thank Bryan McCannon for this suggestion.

6 Mechanisms

The results in this paper tell a story of greater connectivity *causing* significant and persistent growth in business activity and entrepreneurship. Broadband-induced growth is strongest in rural counties situated beyond the borders of metropolitan areas and small firms with fewer than 500 employees. But, through what mechanism(s) does greater connectivity fuel gains in business activity and entrepreneurship? In this section, I outline two potential mechanisms, including enhanced labor market matching and the entry of new firms. The latter of which I explore empirically to determine the margin (extensive or intensive) along which broadband-induced growth can be attributed.

One potential mechanism involves better labor market matching and job search. Since the early 2000s, it has been well documented that the Internet successfully reduced search frictions in a variety of markets (Brown and Goolsbee, 2002; Kroft and Pope, 2014). With regard to job search, unemployed workers who use the Internet to search for work are re-employed 25 percent faster than those who do not use the web (Kuhn and Mansour, 2014). From the firm's perspective, the expansion of broadband reduces the duration of job vacancies and the share of establishments with unfilled positions (Bhuller et al., 2019). In the context of the CAF II program, the connectivity gains that resulted might have led to gains in business activity and entrepreneurship through one or more of these channels.

Thus far, the baseline results have not been attributed to existing firms that choose to expand operations or the entry of new firms. I incorporate BFS data, which track the number of individuals applying for Employer Identification Numbers (EIN) at the county-year level (Bayard et al., 2018). I also use BDS data to define young and mature firms according to the definition used in Davis and Haltiwanger (2019); young firms have been in operation for no more than five years and mature firms have been in operation for more than five years.[31]

To determine how increased broadband availability impacts young and mature firms as well as the number of business applicants, I estimate equation (2) and present the results graphically in Figure 7. It is clear that young firms appear to

[31] To maintain a balanced panel, this designation results in the removal of 85 counties that have censored new-entrant firm data for at least one of the years spanning the 2011-18 period.

be the chief beneficiary of greater broadband availability, consistent with the findings in Falck et al. (2016) and De Stefano et al. (2014). Indeed, after four years, the number of young firms measures more than six percent higher than in the year before the initial installation of broadband. For mature firms, this growth is much more subdued, measuring 1.2 percent. At least part of this young-firm growth is stemming from increased business application activity. After four years, the number of business applications in treated counties increases sharply, up six percent. Taken together, these results suggest that increased broadband availability disproportionately affects new-firm activity. This is a particularly vital finding against the backdrop of decreased business dynamism (Davis and Haltiwanger, 2019; Karahan et al., 2019; Decker et al., 2014).

7 Cost-Benefit Analysis

The results in this paper provide strong evidence that the diffusion of broadband Internet generates statistically and economically significant gains in business activity and entrepreneurship. Furthermore, broadband-induced growth appears to be persistent for a variety of firm outcomes. However, installing new wired broadband connections is expensive, especially in many rural counties where the terrain is more challenging and existing roadways do not always offer direct routes between households. Given the magnitude of CAF II expenditures and the need for additional broadband infrastructure to achieve universal connectivity, it is essential to understand the social welfare consequences of this program.[32] In this section, I illuminate the scale of benefits accrued by counties receiving new broadband connections against the backdrop of various broadband deployment costs.

Broadly speaking, broadband installation costs are borne by two parties: USF contributors (i.e., households and businesses subscribing to telephone service) and telecom providers.[33] Between 2015 and 2018, the average treated county received 841 new broadband connections. Each receiving an average per-location subsidy of $448, the real cost burden (in 2019 dollars) on USF contributors measures $360,732. Assuming that it costs an average of $7,085 to install broadband at a

[32] More than 14 million people still do not have access to broadband-capable networks FCC (2021).

[33] See Table A.3 for a breakout of residential and non-residential USF contributions.

single location (see Table A.4), the real cost burden on telecom providers measured $5.3 million.

Broadband-induced benefits can be characterized by additional subscriber revenue collected by telecom providers and discounted earnings from broadband-induced employment gains. Assuming a take-up rate of 50 percent among newly connected households and an average monthly subscrition rate of $75, telecom providers stand poised to collect $18.9 million from new broadband subscribers. Turning to the future earnings of newly employed individuals, employment in treated counties increases by 1,513 between 2015 and 2018. Assuming a job tenure of four years (BLS, 2020) and using QCEW average annual wages with a discount rate of two percent, the real value of discounted earnings from broadband-induced employment growth measures roughly $223 million. Taken together, the net benefit from the CAF II program likely measured $236 million, with the program's benefits outweighing its costs by a factor of 42.

8 Conclusion

Universal broadband availability, as a federal policy objective, has been catapulted to the foreground in the wake of the COVID-19 Pandemic. With large swaths of the United States still disconnected from broadband-capable networks, it is imperative that researchers and policymakers understand how increased connectivity impacts local economies. The CAF II program, which served as a massive (and exogenous) shock to the *supply* of broadband, issued $6.7 billion in broadband installation subsidies during the 2015-18 period. Overcoming empirical obstacles that remained from previous work, this paper uses address-level broadband deployment data to construct a county-year panel and estimates the causal impact of broadband availability on various measures of business activity and entrepreneurship.

Results from a DID design robust to the staggered deployment of broadband suggest that increased broadband availability increases the number of firms and establishments, employment levels, the average annual wage, and the number of entrepreneurs. These gains persist three years beyond the initial installation of broadband. Driving these baseline results are small firms with fewer than 500 employees, young firms less than six years old, and rural firms located beyond the borders of metropolitan areas. Although seemingly expensive, broadband deployment costs pale in comparison to the labor market benefits that accrue.

With much debate and partisanship surrounding the efficacy of redistributive policies, the results in this paper show that the social benefits of the CAF II program outweigh its upfront costs by a factor of 42. Conservative estimates from a cost-benefit analysis show that discounted earnings from broadband-induced employment gains are large in denomination, valued at more than $223 million. This positive consequence of the CAF II program underscores the usefulness of economic policies that pool together small contributions on a nationwide scale. Going forward, if subsidized, sufficient in scale, and fair in allocation, universal connectivity appears achievable with significant gains in business activity and entrepreneurship.

References

Agrawal, D. R. (2021). The Internet as a Tax Haven? *American Economic Journal: Economic Policy*.

Akerman, A., I. Gaarder, and M. Mogstad (2015). The Skill Complementarity of Broadband Internet. *Quarterly Journal of Economics 130*(4), 1781–1824.

Alm, J., T. Dronyk-Trosper, and S. Larkin (2020). In the Land of OZ: Designating Opportunity Zones. *Public Choice*, 1–21.

Andersen, T. B., J. Bentzen, C.-J. Dalgaard, and P. Selaya (2012). Lightning, IT Diffusion, and Economic Growth Across US States. *Review of Economics and Statistics 94*(4), 903–924.

Atasoy, H. (2013). The Effects of Broadband Internet Expansion on Labor Market Outcomes. *Industrial and Labor Relations Review 66*(2), 315–345.

Baum-Snow, N. and B. F. Lutz (2011). School Desegregation, School Choice, and Changes in Residential Location Patterns by Race. *American Economic Review 101*(7), 3019–46.

Bayard, K., E. Dinlersoz, T. Dunne, J. Haltiwanger, J. Miranda, and J. Stevens (2018). Early-Stage Business Formation: An Analysis of Applications for Employer Identification Numbers. Technical report, National Bureau of Economic Research Paper 24364.

Bertrand, M., E. Duflo, and S. Mullainathan (2004). How Much Should we Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics 119*(1), 249–275.

Bertschek, I., D. Cerquera, and G. J. Klein (2013). More Bits More Bucks? Measuring the Impact of Broadband Internet on Firm Performance. *Information Economics and Policy 25*(3), 190–203.

Bhuller, M., A. Kostol, and T. Vigtel (2019). How Broadband Internet Affects Labor Market Matching. *Available at SSRN 3507360*.

BLS (2020). Employee Tenure Summary. https://www.bls.gov/news.release/pdf/tenure.pdf. Accessed: 7/6/2021.

Briglauer, W., N. S. Dürr, O. Falck, and K. Hüschelrath (2019). Does State Aid for Broadband Deployment in Rural Areas Close the Digital and Economic Divide? *Information Economics and Policy 46*, 68–85.

Brown, J. R. and A. Goolsbee (2002). Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry. *Journal of Political Economy 110*(3), 481–507.

Chow, M. C., T. C. Fort, C. Goetz, N. Goldschlag, J. Lawrence, E. R. Perlman, M. Stinson, and T. K. White (2021). Redesigning the Longitudinal Business Database. Technical report, National Bureau of Economic Research Paper 28839.

Czernich, N., O. Falck, T. Kretschmer, and L. Woessmann (2011). Broadband Infrastructure and Economic Growth. *Economic Journal 121*(552), 505–532.

Davis, S. J. and J. C. Haltiwanger (2019). Dynamism Diminished: The Role of Housing Markets and Credit Conditions. Technical report, National Bureau of Economic Research Paper 25466.

de Chaisemartin, C. and X. d'Haultfoeuille (2020). Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review 110*(9), 2964–96.

De Stefano, T., R. Kneller, and J. Timmis (2014). The (Fuzzy) Digital Divide: Universal Broadband Access and Firm Performance. *University of Nottingham Discussion Paper 14/06 14*(06).

Decker, R., J. Haltiwanger, R. Jarmin, and J. Miranda (2014). The Role of Entrepreneurship in US Job Creation and Economic Dynamism. *Journal of Economic Perspectives 28*(3), 3–24.

Deshpande, M. and Y. Li (2019). Who is Screened Out? Application Costs and the Targeting of Disability Programs. *American Economic Journal: Economic Policy 11*(4), 213–48.

Eckert, F., T. C. Fort, P. K. Schott, and N. J. Yang (2021). Imputing Missing Values in the US Census Bureau's County Business Patterns. Technical report, National Bureau of Economic Research Paper 26632.

Falck, O., A. Mazat, and B. Stockinger (2016). Broadband Infrastructure and Entrepreneurship. *IFO Institute*.

FCC (2015). Eighth Broadband Progress Report. https://www.fcc.gov/reports-research/reports/broadband-progress-reports/2015-broadband-progress-report. Accessed: 10/12/2020.

FCC (2021). Fourteenth Broadband Deployment Report. https://docs.fcc.gov/public/attachments/FCC-21-18A1.pdf. Accessed: 7/5/2021.

Fort, T. C., J. Haltiwanger, R. S. Jarmin, and J. Miranda (2013). How Firms Respond to Business Cycles: The Role of Firm Age and Firm Size. *IMF Economic Review 61*(3), 520–559.

Fox, W., E. Hargaden, and L. Luna (2020). Statutory Incidence and Sales Tax Compliance: Evidence from Wayfair. *Working Paper*.

Fujita, M., P. R. Krugman, and A. Venables (1999). *The Spatial Economy: Cities, Regions, and International Trade*. MIT press.

GAO (2014a). Report to Congressional Requesters: Federal Broadband Deployment Programs and Small Business. Technical report, United States Government Accountability Office Report GAO-14-203.

GAO (2014b). Report to Congressional Requesters: Projects and Policies Related to Deploying Broadband in Unserved and Underserved Areas. Technical report, United States Government Accountability Office Paper GAO-14-409.

Goodman-Bacon, A. (2021). Difference-in-Differences with Variation in Treatment Timing. *Journal of Econometrics*.

Goodman-Bacon, A., T. Goldring, and A. Nichols (2019). bacondecomp: Stata Module for Decomposing Difference-in-Differences Estimation with Variation in Treatment Timing.

Haltiwanger, J. (2012). Job Creation and Firm Dynamics in the United States. *Innovation Policy and the Economy 12*(1), 17–38.

Haltiwanger, J. C. (2021). Entrepreneurship During the COVID-19 Pandemic: Evidence from the Business Formation Statistics. Technical report, National Bureau of Economic Research.

Hjort, J. and J. Poulsen (2019). The Arrival of Fast Internet and Employment in Africa. *American Economic Review 109*(3), 1032–79.

Infrastructure Report Card (2016). Happy 60th Birthday, Interstate Highway System! `https://www.infrastructurereportcard.org/`. Accessed: 10/12/2020.

Ivus, O. and M. Boland (2015). The Employment and Wage Impact of Broadband Deployment in Canada. *Canadian Journal of Economics 48*(5), 1803–1830.

Kandilov, I. T. and M. Renkow (2010). Infrastructure Investment and Rural Economic Development: An Evaluation of USDA's Broadband Loan Program. *Growth and Change 41*(2), 165–191.

Kandilov, I. T. and M. Renkow (2020). The Impacts of the USDA Broadband Loan And Grant Programs: Moving Toward Estimating a Rate of Return. *Economic Inquiry 58*(3), 1129–1145.

Karahan, F., B. Pugsley, and A. Şahin (2019). Demographic Origins of the Startup Deficit. Technical report, National Bureau of Economic Research Paper 25874.

Kim, Y. and P. F. Orazem (2016). Broadband Internet and New Firm Location Decisions in Rural Areas. *American Journal of Agricultural Economics 99*(1), 1–18.

Kolko, J. (2010). A New Measure of US Residential Broadband Availability. *Telecommunications Policy 34*(3), 132–143.

Kolko, J. (2012). Broadband and Local Growth. *Journal of Urban Economics 71*(1), 100–113.

Kroft, K. and D. G. Pope (2014). Does Online Search Crowd Out Traditional Search and Improve Matching Efficiency? Evidence from Craigslist. *Journal of Labor Economics 32*(2), 259–303.

Krugman, P. (1991). Increasing Returns and Economic Geography. *Journal of Political Economy 99*(3), 483–499.

Kuhn, P. and H. Mansour (2014). Is Internet Job Search Still Ineffective? *Economic Journal 124*(581), 1213–1233.

Lindo, J. M., J. Schaller, and B. Hansen (2018). Caution! Men Not at Work: Gender-Specific Labor Market Conditions and Child Maltreatment. *Journal of Public Economics 163*, 77–98.

Meer, J. and J. West (2016). Effects of the Minimum Wage on Employment Dynamics. *Journal of Human Resources 51*(2), 500–522.

Ohrn, E. (2019). The Effect of Tax Incentives on US Manufacturing: Evidence From State Accelerated Depreciation Policies. *Journal of Public Economics 180*, 104084.

Pugsley, B. W. and A. Sahin (2019). Grown-Up Business Cycles. *Review of Financial Studies 32*(3), 1102–1147.

Roller, L.-H. and L. Waverman (2001). Telecommunications Infrastructure and Economic Development: A Simultaneous Approach. *American Economic Review 91*(4), 909–923.

Shideler, D. and N. Badasyan (2012). Broadband Impact on Small Business Growth in Kentucky. *Journal of Small Business and Enterprise Development 19*(4), 589–606.

Sun, L. and S. Abraham (2020). Estimating Dynamic Treatment Effects in Event Studies with Heterogeneous Treatment Effects. *Journal of Econometrics*.

TACIR (2017). Broadband Internet Deployment, Availability, and Adoption in Tennessee. Technical report, Tennessee Advisory Commission on Intergovernmental Relations.

U.S. Census Bureau (2020). What is New in the Redesigned Business Dynamics Statistics? https://www2.census.gov/programs-surveys/bds/updates/bds2018-release-note-20201112-enty-exit-rate-fix.pdf. Accessed: 5/26/2021.

U.S. Telecom Association (2019). United States Broadband Investment Continued in 2018. https://www.ustelecom.org/wp-content/uploads/2019/07/USTelecom-Research-Brief-Capex-2018-7-31-19.pdf. Accessed: 10/12/2020.

Whitacre, B., R. Gallardo, and S. Strover (2014). Broadband's Contribution to Economic Growth in Rural Areas: Moving Towards a Causal Relationship. *Telecommunications Policy 38*(11), 1011–1023.

Wolfers, J. (2006). Did Unilateral Divorce Laws Raise Divorce Rates? A Reconciliation and New Results. *American Economic Review 96*(5), 1802–1820.

Zuo, G. W. (2019). Wired and Hired: Employment Effects of Subsidized Broadband Internet for Low-Income Americans. *American Economic Journal: Economic Policy*.

Table 1: Treatment Timing

	2015	2016	2017	2018	2019
Number of Counties Treated	905	1,462	2,083	2,200	2,511
Share of Counties Treated (%)	31.8	51.4	72.9	76.9	87.3
Number of Counties Not Yet Treated	1,938	1,381	771	657	360
Share of Counties Not Yet Treated (%)	68.2	48.6	27.1	23.1	12.7

Note: In total, 2,841 counties contain eligible locations. Counties are classified as "treated" when at least one new broadband installation is made using CAF II subsidies, remainig treated thereafter. A very small share (one percent) of treated counties do not contain any eligible locations. This is because participating telecom providers are allowed to use CAF II subsidies to extend broadband service to locations classified as "extremely high cost" to meet its state-wide deployment obligation. By 2019, 30 counties were classified as both inelligible and treated.

Table 2: Summary Statistics of Broadband Rollout Measures

Urban Counties ($N = 1,902$)	Mean	SD	25th	50th	75th
Number of Price-cap Locations	73,657	187,202	10,458	22,968	60,125
Share of Locations Eligible for CAF II	9.3	11.4	0.9	4.8	14.1
Annual Broadband Installation Subsidies ($)	611,822	724,847	123,779	411,277	839,613
Annual Support per Eligible Location ($)	400	148	303	371	462
Rural Counties ($N = 939$)	**Mean**	**SD**	**25th**	**50th**	**75th**
Number of Price-cap Locations	8,994	9,123	2,693	6,061	12,715
Share of Locations Eligible for CAF II	18.7	17.5	5.8	14.1	27.1
Annual Broadband Installation Subsidies ($)	536,235	554,260	109,529	370,563	819,740
Annual Support per Eligible Location ($)	545	233	381	500	658

Note: Data are from the Federal Communication Commission's (FCC) Final Eligible Areas Map. Eligible counties include at least one household or business eligible for CAF II broadband installation subsidies. Price-cap locations are households and businesses located within the service territories of large telecom "price-cap" carriers (e.g., AT&T, Verizon, CenturyLink, Windstream).

Table 3: Summary Statistics of Business Activity Measures

Urban Counties ($N = 1,966$)	Mean	SD	25th	50th	75th
Firms	2,766	7,585	352	761	2,154
Establishments	3,265	9,139	378	829	2,467
Employment	58,411	176,505	4,846	12,209	40,914
Average Annual Wage	39,579	9,226	33,709	37,606	43,058
Business Applications	1,379	4,732	110	246	792
Entrepreneurs	11,589	37,430	1,367	2,958	8,157
Entrepreneurial Revenue	544,284	1,894,973	51,786	118,410	348,364
Rural Counties ($N = 1,136$)	**Mean**	**SD**	**25th**	**50th**	**75th**
Firms	322	376	100	185	401
Establishments	349	416	105	198	430
Employment	4,415	6,031	914	2,041	5,223
Average Annual Wage	35,069	7,351	30,492	33,908	38,094
Business Applications	94	156	27	54	108
Entrepreneurs	1,083	1,181	370	709	1,347
Entrepreneurial Revenue	44,619	54,886	13,651	26,964	53,134

Note: Firm, establishment, and employment data are from the Business Dynamics Statistics (BDS), average annual wage data are from the Quarterly Census of Employment and Wages (QCEW), business application data are ffrom the Business Formation Statistics (BFS), and entrepreneurial data are from the Nonemployer Statistics (NES). Entrepreneurs are defined as sole proprietors and entrepreneurial revenue is defined as annual sole proprietor receipts. The sample period is 2011-18. Included in the summary statistics are both eligible and ineligible counties.

Table 4: Baseline TWFE Estimation Results

	ln(*Firms*)	ln(*Establishments*)	ln(*Employment*)	ln(*Wages*)	ln(*Entrepreneurs*)	ln(*Revenues*)
$1\{broadband\} = eligible \times treated$	0.0008	0.0022	0.0097***	0.0016	0.0024	0.0130***
	(0.0014)	(0.0014)	(0.0027)	(0.0014)	(0.0017)	(0.0028)
Observations	20,088	20,088	20,088	20,088	20,088	20,088
Fixed Effect (county)	Yes	Yes	Yes	Yes	Yes	Yes
Fixed Effect (year)	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table reports the results from estimating a standard TWFE model. The treatment indicator is defined as $1\{broadband\} = eligible \times treated$ and equals 1 when at least one new broadband connection is made using CAF II subsidies. Wages refer to average annual wages, entrepreneurs are defined as sole proprietors, and revenues are defined as annual sole proprietor receipts. Firm, establishment, and employment data are from the Business Dynamics Statistics (BDS), average annual wage data are from the Quarterly Census of Employment and Wages (QCEW), and entrepreneurship data are from the Nonemployer Statistics (NES). Each estimation controls for the FCC's concurrent broadband deployment programs and cell tower availability. The sample period is 2011-18. Standard errors are clustered at the county level and presented in parentheses.

Table 5: Bacon Decomposition for TWFE Estimation Results

	Timing Groups (weight)		Never vs. Timing (weight)		Within Residual (weight)		Overall
ln(*Firms*)	0.0003	(0.62)	0.0037	(0.38)	-0.2292	(0.00)	0.0008
ln(*Establishments*)	0.0015	(0.62)	0.0056	(0.38)	-0.2393	(0.00)	0.0022
ln(*Employment*)	0.0053	(0.62)	0.0198	(0.38)	-0.3248	(0.00)	0.0097
ln(*Wages*)	0.0011	(0.62)	0.0021	(0.38)	0.0317	(0.00)	0.0016
ln(*Entrepreneurs*)	0.0009	(0.62)	0.0087	(0.38)	-0.4088	(0.00)	0.0024
ln(*Revenues*)	0.0134	(0.62)	0.0149	(0.38)	-0.2800	(0.00)	0.0130

Note: This table shows the Goodman-Bacon Decomposition of the TWFE estimates produced from the baseline model. The treatment indicator is defined as $1\{broadband\} = eligible \times treated$ and equals 1 when at least one new broadband connection is made using CAF II subsidies. Each row represents a separate estimation, each including controls for the FCC's other broadband deployment programs and the number of cell towers at the county level. Wages refer to average annual wages, entrepreneurs are defined as sole proprietors, and revenues are defined as annual sole proprietor receipts. The "timing groups" column captures two-by-two diff-in-diff comparisons of counties treated earlier versus later in the sample. The "never vs. timing" column captures two-by-two diff-in-diff comparisons of treated counties with never-treated counties. For reference, the overall estimates are presented in the last column.

Table 6: Economic Factors Correlated with Treatment

	Broadband Installation Treatment Status							
	$1\{broadband\}_{2015}$		$1\{broadband\}_{2016}$		$1\{broadband\}_{2017}$		$1\{broadband\}_{2018}$	
$Jobless\,Rate$	−0.0004	(0.0127)	−0.0081	(0.0141)	0.0148	(0.0199)	−0.0647**	(0.0276)
$Jobless\,Rate_{t-1}$	0.0043	(0.0122)	−0.0149	(0.0148)	−0.0111	(0.0169)	0.0542**	(0.0249)
$Cell\,Towers$	−0.0047	(0.0049)	−0.0044	(0.0059)	−0.0097**	(0.0045)	0.0084*	(0.0049)
$Cell\,Towers_{t-1}$	0.0054	(0.0050)	0.0056	(0.0060)	0.0104**	(0.0046)	−0.0081	(0.0050)
$Other\,Broadband\,Deployments$	−0.0001	(0.0001)	−0.0001*	(0.0000)	0.0000	(0.0000)	0.0000	(0.0000)
$Other\,Broadband\,Deployments_{t-1}$	0.0001	(0.0001)	−0.0002***	(0.0001)	−0.0001*	(0.0000)	0.0000	(0.0000)
$Average\,Annual\,Wage$	0.0105**	(0.0051)	0.0105*	(0.0055)	0.0150**	(0.0065)	−0.0092*	(0.0051)
$Average\,Annual\,Wage_{t-1}$	−0.0122**	(0.0050)	−0.0092*	(0.0054)	−0.0145**	(0.0066)	0.0101*	(0.0056)

Note: This table examines the correlation between observable economic characteristics and the treatment status of counties. The dependent variable is $1\{broadband\} = eligible \times treated$ and equals 1 when at least one new broadband connection is made using CAF II subsidies. Each estimation includes state fixed effects and 3,102 county-year observations. Standard errors are clustered at the county level.

Table 7: Economic Factors Correlated with the Timing of Treatment

	Timing of Initial Broadband Installation Treatment							
	$Year = 2015$		$Year = 2016$		$Year = 2017$		$Year = 2018$	
$Jobless\,Rate$	0.0560	(0.0379)	0.0548	(0.0376)	0.0081	(0.0403)	−0.1526**	(0.0642)
$Jobless\,Rate_{t-1}$	−0.0312	(0.0359)	−0.0287	(0.0374)	0.0217	(0.0333)	0.1648***	(0.0556)
$Cell\,Towers$	0.0087	(0.0114)	0.0078	(0.0103)	0.0014	(0.0115)	−0.0169	(0.0113)
$Cell\,Towers_{t-1}$	−0.0102	(0.0118)	−0.0093	(0.0106)	−0.0027	(0.0117)	0.0159	(0.0116)
$Other\,Broadband\,Deployments$	0.0004***	(0.0001)	0.0000	(0.0000)	0.0001	(0.0001)	0.0001	(0.0001)
$Other\,Broadband\,Deployments_{t-1}$	0.0001	(0.0004)	0.0004***	(0.0001)	−0.0001	(0.0000)	0.0001	(0.0001)
$Average\,Annual\,Wage$	−0.0234	(0.0156)	−0.0181	(0.0160)	0.0152	(0.0135)	0.0093	(0.0139)
$Average\,Annual\,Wage_{t-1}$	0.0184	(0.0155)	0.0123	(0.0162)	−0.0216	(0.0141)	−0.0156	(0.0149)

Note: This table examines the correlation between observable economic characteristics and the *timing* of treatment. The dependent variable is the year of initial treatment. Each estimation includes state fixed effects and 2,200 county-year observations. Standard errors are clustered at the county level.

Table 8: Checking for Reverse Causality

	Coefficient	Standard Error
$\ln(Firms)_{t-1}$	−0.7952	(0.9464)
$\ln(Establishments)_{t-1}$	0.5369	(0.9550)
$\ln(Employment)_{t-1}$	0.2081	(0.1288)
$\ln(Average\,Annual\,Wage)_{t-1}$	0.0214	(0.2057)
$\ln(Entrepreneurs)_{t-1}$	0.4566*	(0.2328)
$\ln(Entrepreneurial\,Revenue)_{t-1}$	0.0075	(0.1115)

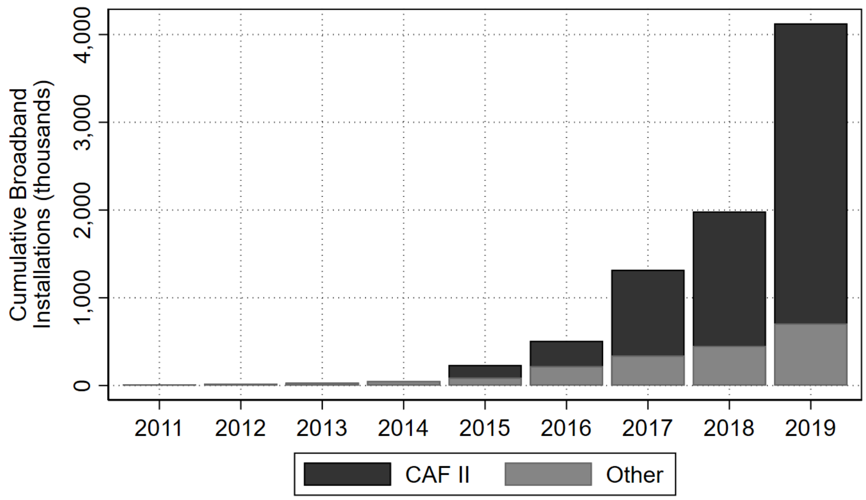
Note: This table checks for the presence of reverse causality between treatment status and lagged business activity. Entrepreneurs are defined as sole proprietors. Entrepreneurial revenue is defined as sole proprietor revenue. The estimation includes county and year fixed effects and the standard set of controls used in the baseline model. Standard errors are clustered at the county level.

Table 9: Does Treatment Impact Migration Flows?

	Measures of Net Household Migration		
	Number of Households	Adjusted Gross Income (AGI)	AGI per Household
$1\{broadband\} = eligible \times treated$	51.7**	1881.1	43.1
	(22.0)	(2407.4)	(36.9)

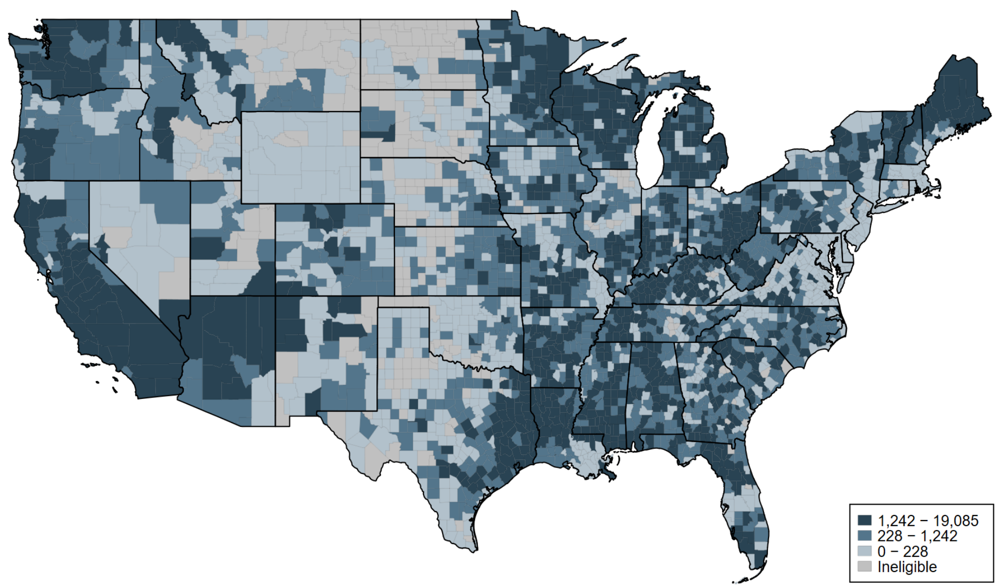
Note: This table investigates whether treatment status determines net household migration and/or net household income. The first dependent variable is defined as the difference between the number of in-migrant and out-migrant households. The second dependent variable is defined as the difference between in-migrant and out-migrant household adjusted gross income (AGI). The third dependent variable is defined as the difference between in-migrant and out-migrant adjusted gross income per household. Data are from the IRS SOI Tax Stats migration data. Included are county and year fixed effects. Standard errors are clustered at the county level.

Figure 1: Cumulative Broadband Rollout (2011-19)



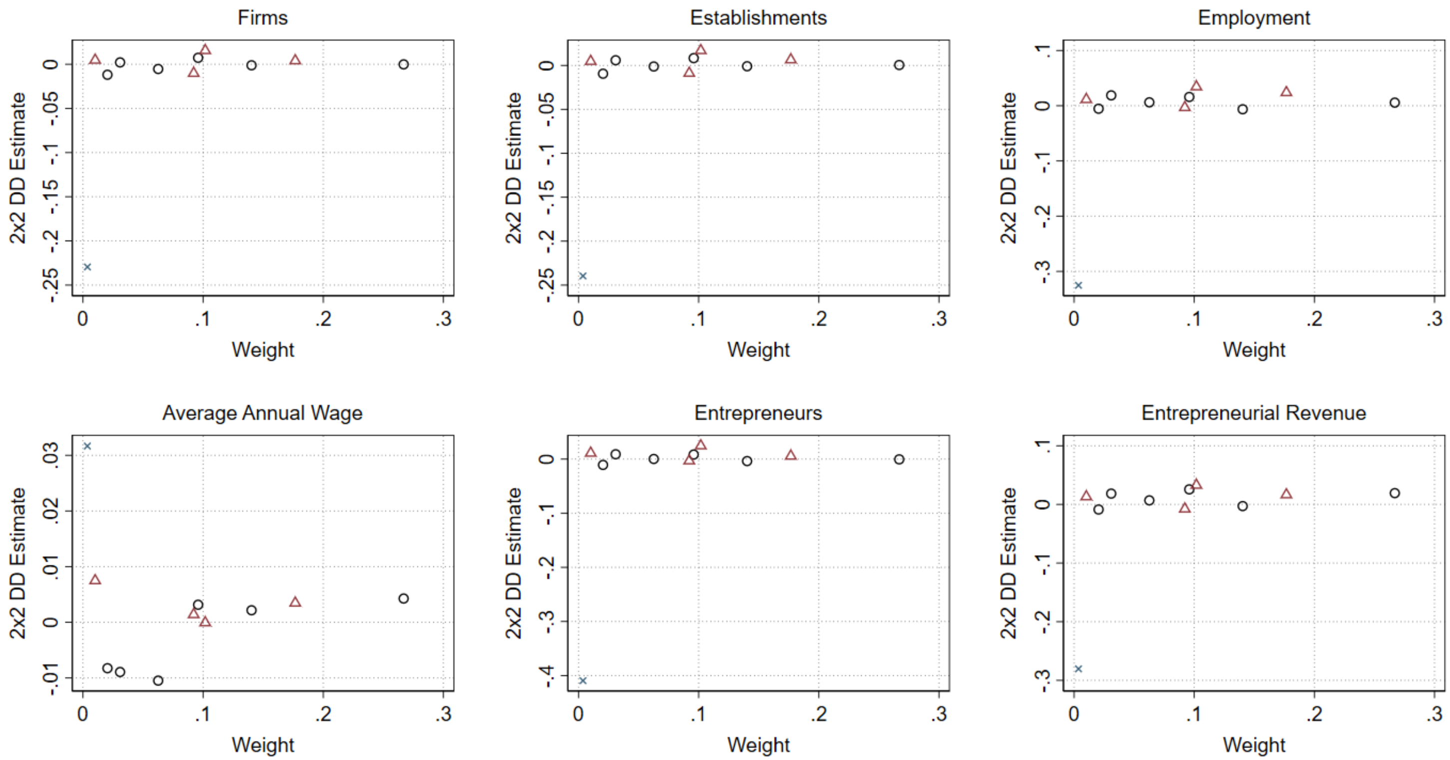
Note: This figure stacks the number of CAF II broadband installations on top of broadband installations from the FCC's other major deployment programs, including ACAM, CAF-BLS, and RBE.

Figure 2: CAF II Cumulative Broadband Rollout (2015-19)



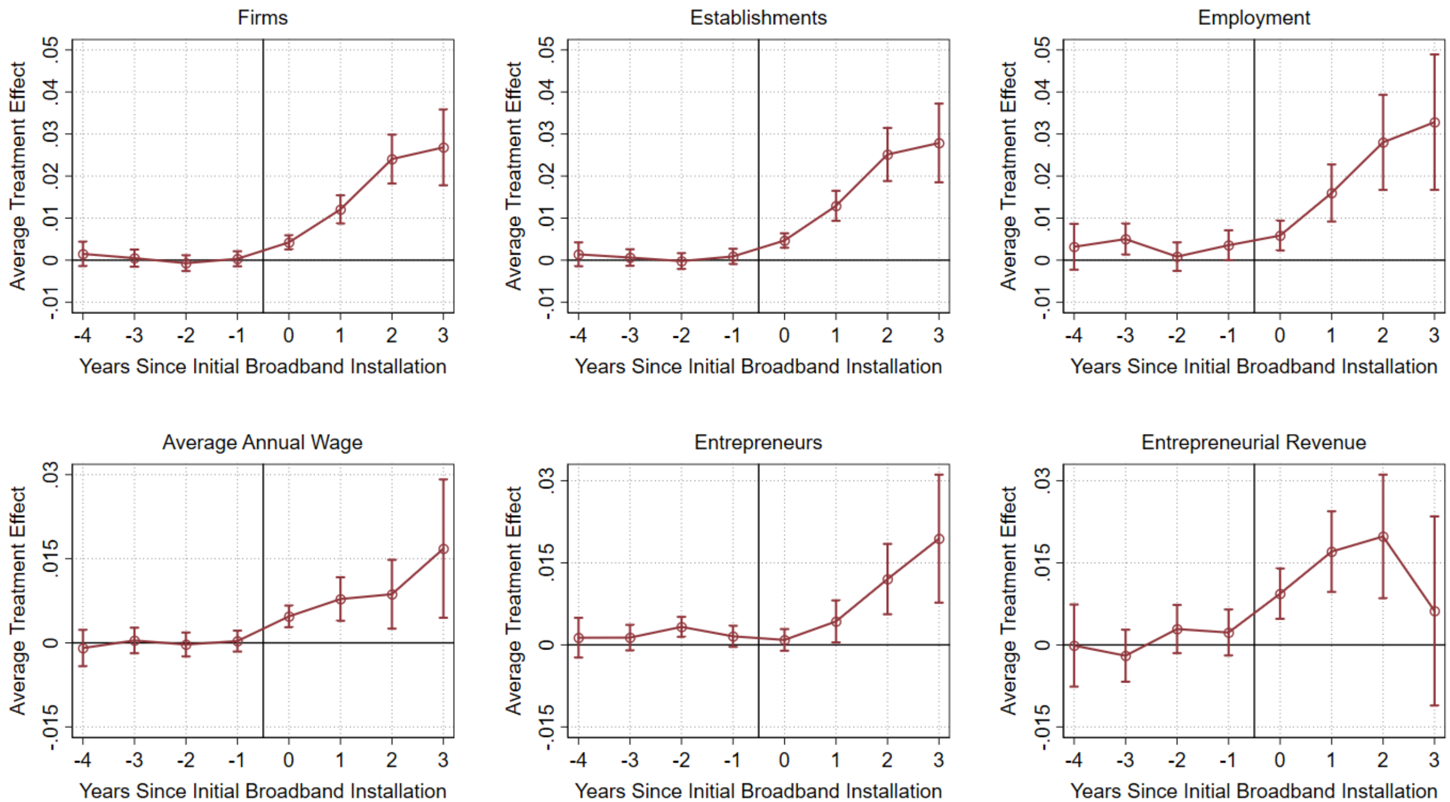
Note: Eligible counties contain high-cost locations with average monthly connection costs between \$52.50 and \$198.60. Eligible locations must have model-based connection costs between the benchmark of \$52.50 and the EHCT of \$198.60, must not be provided broadband service by an unsubsidized provider or subsidized wireline provider, and must not have been eligible for the Rural Broadband Experiments.

Figure 3: Bacon Decomposition for TWFE Estimation Results



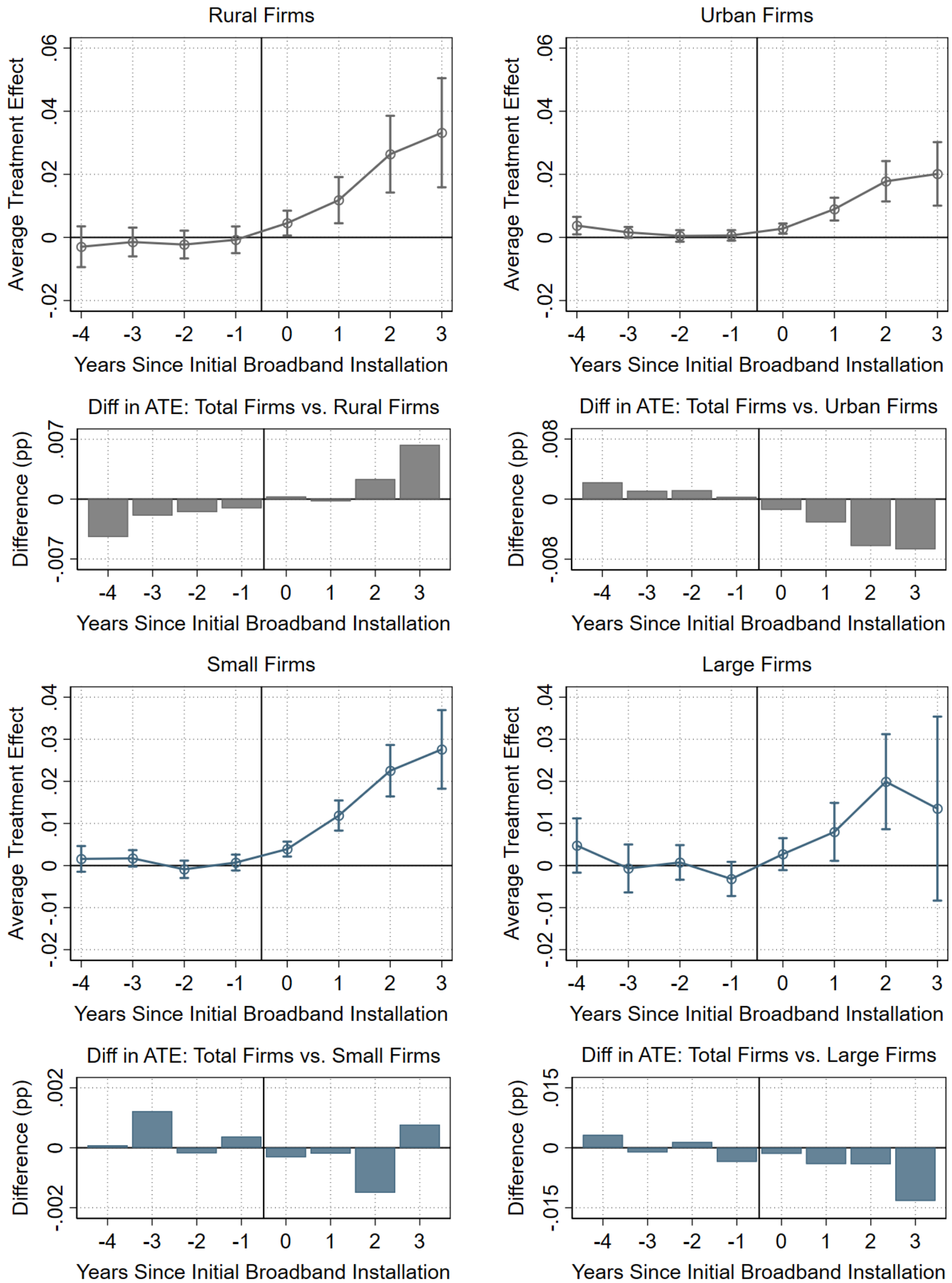
Note: Following Goodman-Bacon (2021), this figure plots each 2x2 difference-in-differences (DID) estimate from equation (1) for various outcome measures. Each DID estimate is plotted (vertical axis) against its weight (horizontal axis) used to generate the overall estimate. The open circles are DID terms comparing early treated counties with later treated counties; see the "timing groups" column in the Bacon decomposition table. The open triangles are DID terms comparing ever-treated counties with never-treated counties; see "never vs. timing" column in the Bacon decomposition table. Wages refer to average annual wages, entrepreneurs are defined as sole proprietors, and revenues are defined as annual sole proprietor receipts.

Figure 4: Baseline DID Estimates: Broadband's Impact on Firm Activity and Entrepreneurship



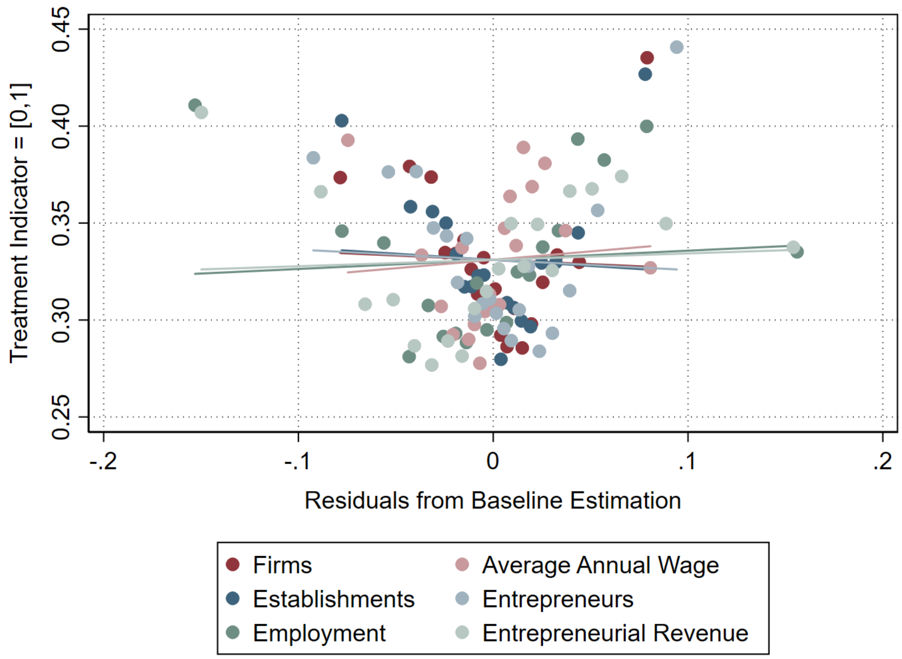
Note: This figure displays the DID estimates from estimating equation (2). Data are from the Business Dynamics Statistics (BDS), Quarterly Census of Employment and Wages (QCEW), and Nonemployer Statistics (NES). Entrepreneurs are defined as sole proprietors and entrepreneurial revenue is equal to sole proprietor revenue. The estimation sample is 2011-2018. Each treatment effect is bound by its 95-percent confidence interval. Each estimating equation includes county and year fixed effects, controls for concurrent broadband deployment programs, cell tower availability, and state-specific linear time trends. Standard errors are clustered at the county level and constructed using 1,000 bootstrap replications.

Figure 5: Broadband's Impact by Urbanicity and Firm Size



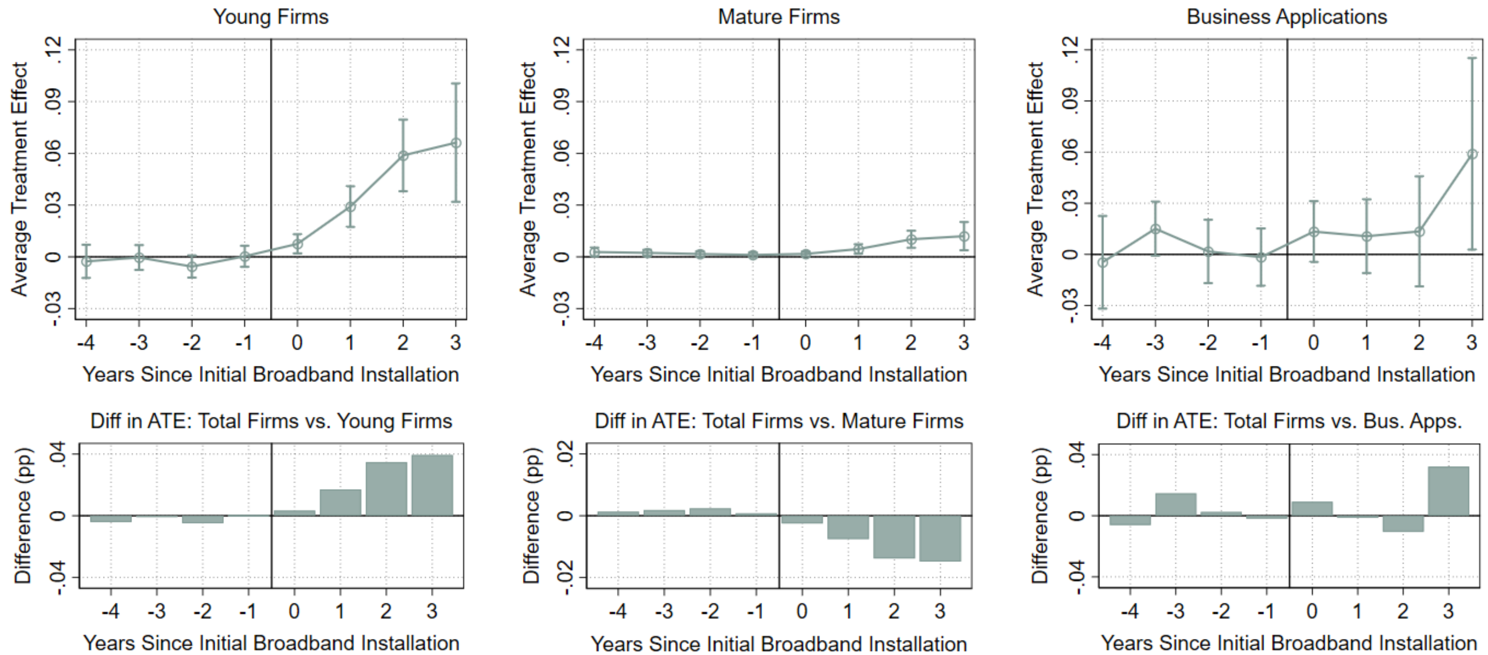
Note: This figure displays the DID estimates from estimating equation (2) for rural firms, urban firms, small firms (with fewer than 500 employees), and large firms (with at least 500 employees). Data are from the Business Dynamics Statistics (BDS). The estimation sample is 2011-2018. Each treatment effect is bound by its 95-percent confidence interval. Each estimating equation includes county and year fixed effects, controls for concurrent broadband deployment programs, cell tower availability, and state-specific linear time trends. Standard errors are clustered at the county level and constructed using 1,000 bootstrap replications.

Figure 6: Unobserved Business Activity Determinants and Treatment



Note: This figure investigates whether treatment status is a function of unobserved determinants of business activity and entrepreneurship. Each of the six main outcome variables were regressed on the treatment indicator, $1\{broadband\} = eligible \times treated$, and the standard set of controls with county and year fixed effects. The treatment indicator was then regressed on the residuals. The overlaid binscatter plots include linear fitted lines showing practically no correlation between treatment status and unobserved determinants of business activity.

Figure 7: Broadband's Impact on Young Firms, Mature Firms, and Business Applications



Note: This figure displays the DID estimates from estimating equation (2) for young firms (i.e., firms less than 6 years old), mature firms (firms that have been in existence for at least six years), and business applications. Data are from the Business Dynamics Statistics (BDS) and Business Formation Statistics (BFS). The estimation sample is 2011-2018. Each treatment effect is bound by its 95-percent confidence interval. Each estimating equation includes county and year fixed effects, controls for concurrent broadband deployment programs, cell tower availability, and state-specific linear time trends. Standard errors are clustered at the county level and constructed using 1,000 bootstrap replications.

Appendix for
"Broadband Internet and Business Activity"

Richard Beem
rbeem1@vols.utk.edu
916 Volunteer Boulevard, Knoxville, Tennessee, 37996

Appendix A Additional Tables and Figures

This appendix includes supporting tables that describe the urbanicity designation of counties, that compare the dynamic DID estimates with and without state-specific linear time trends, that outline the monthly USF contributions of households and businesses, and that provide estimates of broadband installation costs, as well as supporting figures, which present the geographic distribution of urban and rural counties and compare several broadband rollout measures by urbanicity.

A.1 *Additional Tables*

Table A.1: Defining Counties by USDA Urbanicity Codes

Urban Counties	
RUCC = 1	In metro areas with population ≥ 1,000,000
RUCC = 2	In metro areas with population ≥ 250,000 and < 1,000,000
RUCC = 3	In metro areas with population < 250,000
RUCC = 4	Urban population ≥ 20,000, adjacent to metro area
RUCC = 6	Urban population ≥ 2,500 and < 20,000, adjacent to metro area
Rural Counties	
RUCC = 5	Urban population ≥ 20,000, not adjacent to metro area
RUCC = 7	Urban population ≥ 2,500 and < 20,000, not adjacent to metro area
RUCC = 8	Completely rural or < 2,500 urban population, adjacent to metro area
RUCC = 9	Completely rural or < 2,500 urban population, not adjacent to metro area

Note: Rural-Urban Continuum Codes (RUCC) are published by the United States Department of Agriculture (USDA) every 10 years. The definitions used in this table are from the most-recent (2013) vintage. Source: https://www.ers.usda.gov/data-products/rural-urban-continuum-codes/.

Table A.2: Cpmparing DID Results With and Without State-Specific Time Trends

	Placebo Effects				Treatment Effects			
	$t-4$	$t-3$	$t-2$	$t-1$	$t=0$	$t+1$	$t+2$	$t+3$
$\ln(Firms)$	0.001	0.001	0.002	0.002	0.002	0.005	0.010	0.009
$\ln(Establishments)$	0.002	0.002	0.002	0.002	0.002	0.004	0.009	0.008
$\ln(Employment)$	0.004	0.004	0.002	0.001	0.001	0.002	0.004	0.006
$\ln(Average\,Annual\,Wage)$	0.002	0.003	0.002	0.001	0.001	0.002	0.005	0.010
$\ln(Entrepreneurs)$	0.000	0.001	0.001	0.001	0.001	0.002	0.004	0.011
$\ln(Entrepreneurial\,Revenue)$	−0.002	0.000	0.002	0.001	0.000	0.000	0.002	−0.009

Note: This table displays the differences between point estimates generated from the model with controls and state-specific linear time trends and the model with controls only. The differences are in percentage-point format.

Table A.3: Per-Household Contributions to the Federal Universal Service Fund

	Total = Residential + Business			Residential	
	High-Cost	Other	Total	Est (Low)	Est (High)
Mean (2015-17)	\$2.94	\$2.56	\$5.50	\$2.47	\$3.02

Note: This table apportions total monthly contributions to the Federal Universal Service Fund to businesses and households. Data are from the FCC's Universal Service Monitoring Reports. High-Cost contributions fund CAF II broadband installation subsidies. Residential estimates approximate monthly household contributions.

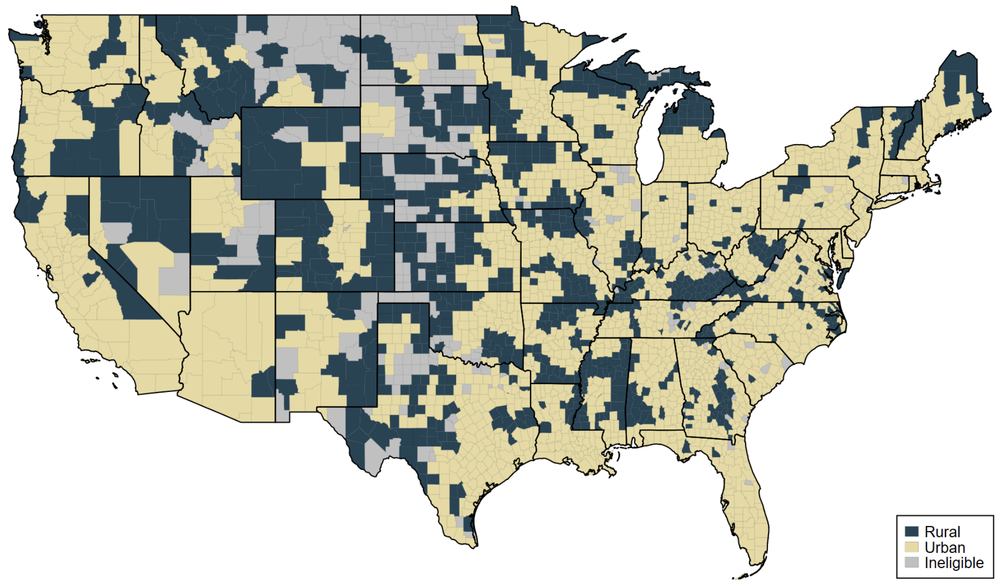
Table A.4: Estimating Broadband Deployment Costs for Disconnected Locations

Per Eligible Location	Type	Low Estimate	Mean Estimate	High Estimate
TNECD Report	fixed wireless	\$1,100	—	\$3,040
Southern Tier Wireless	fixed wireless	—	\$1,200	—
TNECD Report	FTTH	\$2,500	—	\$3,840
Industry Representatives	FTTH	\$5,000	—	—
Per New Subscriber				
TACIR Report	FTTH	\$2,391	—	\$10,870
TACIR Report	FTTH	\$2,750	—	\$12,500
CenturyLink	FTTH	\$4,000	—	\$10,000

Note: The Federal Communication Commission's model-based funding formula is based on the cost of building fiber-to-the-home (FTTH), although providers are allowed to expand service using different technologies as long as they offer service with download/upload speeds of 10/1 Mbps. The low estimate, from industry representatives, of \$5,000 to build out FTTH service is an estimate for rural areas. The mean estimate of \$1,200 for building out fixed wireless service, according to Southern Tier Wireless, is based on rural areas of New York. The FTTH cost estimates published in the TNECD Report include the "design, engineering, permitting, and fiber construction, including labor, materials, equipment, shelters, and all components of the outside plant infrastructure." CenturyLink's cost estimate is based on the expansion of fiber in Minnesota. TACIR = Tennessee Advisory Commission on Intergovernmental Relations. TNECD = Tennessee Department of Economic and Community Development.

A.2 *Additional Figures*

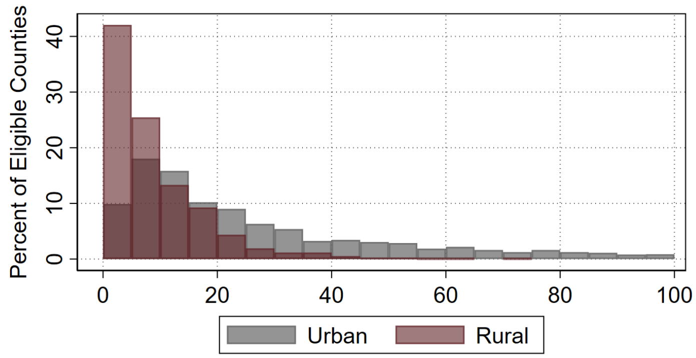
Figure A.1: Defining Rural, Urban, and Ineligible Counties



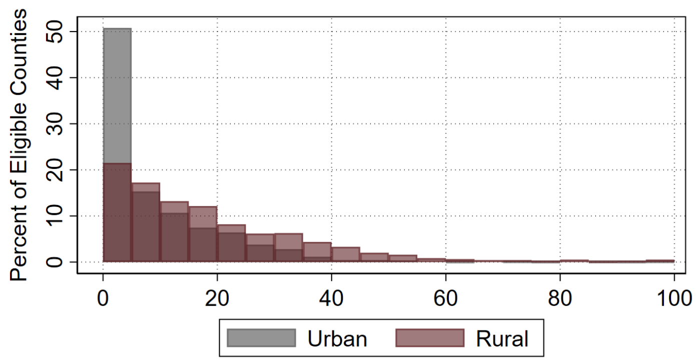
Note: This map displays the geographic dispersion of rural and urban counties. Rural counties are those with RUCC classifications of 5, 7, 8, or 9. This includes non-metropolitan counties *not* adjacent to a metropolitan county and counties classified as completely rural. Urban counties have RUCC classifications of 1, 2, 3, 4, or 6. This includes metropolitan counties and non-metropolitan counties that are adjacent to a metropolitan county. RUCC classifications are from the USDA's most-recent vintage (2013). The urban-rural delineation is generated by the author and not the USDA. Counties that are ineligible for CAF II broadband installation subsidies are shaded gray.

Figure A.2: Distribution of Broadband Rollout Measures by Urbanicity

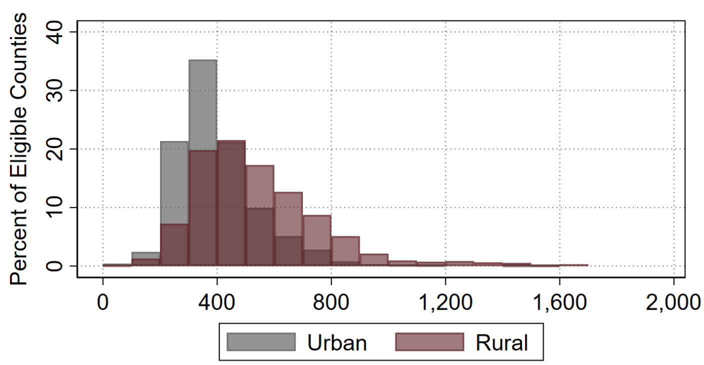
(a) Thousands of Locations within Telecom Service Territories



(b) Share of Locations Eligible for CAF II Subsidies



(c) Annual Broadband Installation Subsidy per Location



Note: This figure displays the distribution of three broadband rollout measures by urbanicity. In all panels above, "locations" refers to households and businesses. In Panel (a), counties with more than 100,000 locations are dropped to enhance readability, however, roughly 90 percent of counties remain.

Appendix B Description of Data Sources

B.1 *Connect America Fund Broadband Map*

In this paper, broadband deployment data come from the FCC's Connect America Fund Broadband Map (CAF Map), which is available at the following link: https://data.usac.org/publicreports/caf-map/. The interactive map displays the geographic boundaries that contain locations (i.e., households and businesses) eligible for various broadband deployment subsidies and identifies recipient locations at the address level. The CAF Map includes data for seven of the FCC's broadband deployment funds. These include the Alternative Connect America Model (ACAM) and ACAM II, the Alaska (AK) Plan, Phase II of the Connect America Fund (CAF II), the CAF II Auction, Connect America Fund Broadband Loop Support (CAF-BLS), and the Rural Broadband Experiment (RBE).[34]

The data underlying the CAF Map are submitted annually to the Universal Service Administrative Company (USAC) by participating telecom providers via the High Cost Universal Broadband (HUBB) portal.[35] To ensure data accuracy, the USAC verifies a random sample of reported locations each year. However, the USAC does not verify *all* data underlying the CAF Map.

For the years 2000 through 2019, the CAF Map offers the following data for locations receiving broadband installations using subsidies from one of the funds outlined above: the fund type from which the carrier receives broadband installation subsidies, the study area code assigned to the specific carrier, the name of the carrier, the longitude and latitude of the newly connected location, the deployment address, city, state, and ZIP code, the deployment date and year, the census block code, the location obligation of the carrier, the number of units associated with the deployed location (for example, a multi-family housing unit represents a single location, but numerous broadband connections), the target build-out for the carrier, the download/upload speed of the newly connected location, and the total amount of state support received by the carrier.

[34]ACAM provides broadband installation subsidies to rate-of-return carriers (i.e., small telephone companies). ACAM II provides subsidies to rate-of-return carriers that transitioned voluntarily from CAF-BLS funding. The AK Plan provides subsidies to rate-of-return carriers that operate in rural Alaska. CAF II provides subsidies derived from an engineering cost model to price-cap carriers (i.e., large telephone companies). The CAF II Auction provides subsidies to price-cap carriers and other Internet Service Providers (ISPs) that successfully bid to install broadband in service areas where the incumbent carrier declined model-based support from CAF II. CAF-BLS provides subsidies to rate-of-return carriers based on their costs and financial data. The RBE program subsidies to telecom carriers and other ISPs that successfully bid to install broadband in price-cap areas that are unserved.

[35]The USAC is an independent not-for-profit designated by the FCC and tasked with managing the contribution of revenue and distribution of funding from the Universal Service Fund (USF).

B.2 *CAF II Final Eligible Areas Map*

The CAF II Final Eligible Areas Map provides county-level data for the number of locations eligible to receive CAF II broadband installation subsidies, the total support amount received by carriers operating in each county, and the total number of locations in each county, which serves as an approximation of each county's total number of households and businesses. The Final Eligible Areas Map is available at the following link: `https://www.fcc.gov/reports-research/maps/connect-america-phase-ii-` In addition to describing *eligible* locations, the map also shades the areas that are *ineligible* for CAF II funds. Individual locations can be deemed ineligible for a variety of reasons. These primarily include the following: the location's model-based connection cost exceeds the extremely high-cost threshold (EHCT) of $198.60, the model-based connection cost falls below the benchmark of $52.50, or the local carrier has chosen not to participate in the CAF II program. Formally, eligible locations must have model-based connection costs between the benchmark of $52.50 and the EHCT of $198.60, must not be provided broadband service by an unsubsidized provider or subsidized wireline provider, and must not have been eligible for the RBE program.

B.3 *County Business Patterns*

Unlike survey data, the County Business Patterns (CBP) data are extracted from the Business Register (BR), which is the Census Bureau's most complete and up-to-date tabulation of business establishments in the United States.[36] The CBP data used in this paper come in two forms: raw data files constructed at the county-industry-year level and the county-industry-year panel with harmonized NAICS industry codes made available by Eckert et al. (2021). The raw CBP data files are made available at `https://www.census.gov/programs-surveys/cbp/data/datasets.html` and offer data describing the number of business establishments, employment during the week of March 12 of the reference year, annual payroll, and first-quarter payroll. In particular, the CBP data break out the number of business establishments by establishment size. These include the following employment size classes: 1-4, 5-9, 10-19, 20-49, 50-99, 100-249, 250-499, 500-999, and 1,000 plus. Although industry data are made available, the large degree of censoring does not permit a useful analysis. Therefore, I construct a county-year panel using the aggregate numbers for each county, maintaining the establishment size class detail.

Researchers face two main roadblocks when attempting to construct a county-year panel of CBP data over a relatively long time horizon. First, employment

[36]The BR is continuously updated, pulling data from economic censuses, quarterly and annual federal income and payroll tax records, and other administrative records programs.

for small counties and/or county-industry cells remains suppressed for confidentiality in the public-use files. Second, industry classifications change over time. This is true for both the Standard Industrial Classification (SIC) codes and the North American Industrial Classification System (NAICS) codes that replaced the former in 1998. Fortunately, Eckert et al. (2021) have made available a county-industry-year panel that offers both imputed employment for missing cells and a consistent NAICS (2012) industry classification. The data can be found at `http://fpeckert.me/cbp/`. To construct a balanced county-year panel, I replace suppressed employment levels from the raw files with imputed employment from the harmonized panel.[37]

Even after incorporating imputed employment, two counties remain incomplete. These include King County, Texas (FIPS=48269) and Loving County, Texas (FIPS=48301). In constructing a balanced panel, these counties are dropped from the analysis.[38]

B.4 *Nonemployer Statistics*

To complement the CBP data, which provide a detailed overview of business establishments *with* paid employees, the Nonemployer Statistics (NES) provide data on "establishments" *without* paid employees that are subject to federal income taxes with annual revenues of at least $1,000. The NES data can be found at the following link: `https://www.census.gov/programs-surveys/nonemployer-statistics/data/datasets.html`. These data provide a snapshot of entrepreneurial activity at the county level. Although most nonemployers are self-employed (i.e., sole proprietors), their earned income may or may not be their primary source of income.

During the 2010-18 period, three counties remained incomplete, with at least one county-year cell suppressed for confidentiality purposes. These include King County, Texas (FIPS=48269), Loving County, Texas (FIPS=48301), and Hudspeth County, Texas (FIPS=48229). Consequently, I drop the latter from the panel, which joins the former two that have already been removed from the data.[39]

[37] In total, 194 replacements were made to the raw data, accounting for 0.6 percent of the county-year cells spanning the 2010-19 period.

[38] King County, Texas was not eligible for CAF II broadband installation subsidies, but did receive a total of 27 broadband installations under the CAF-BLS program between 2017 and 2018. Loving County, Texas contained 14 locations eligible for CAF II broadband installation subsidies. All of which, plus an additional four locations, were connected in 2019. This is permissible because participating telecom providers can choose to connect locations that have connection costs above the EHCT so long as their total number of connections satisfies their CAF II obligation.

[39] Hudspeth County, Texas contained 155 CAF-II-eligible locations. In 2019, 198 locations were connected.

B.5 *Business Dynamics Statistics*

To understand how *dynamic* firm activity is at the county level, I turn to the Business Dynamics Statistics (BDS). The BDS data provide annual measures of firm startups and shutdowns, establishment openings and closings, as well as employment creation and destruction. In this paper, I incorporate county-level data from three BDS files: `bds2018_cty`, `bds2018_cty_fsize`, and `bds2018_cty_fage`, all of which are available at https://www.census.gov/data/datasets/time-series/econ/bds/bds-datasets.html. The BDS data are pulled from the confidential Longitudinal Business Database (LBD), which provides micro-data on business dynamics.

Recently, the BDS data were updated to account for the redesign of the LBD (Chow et al., 2021). The updated, or "redesinged", BDS data use a consistent NAICS industry classification for the 1978-2018 panel, offer improvements in source data, integration with the Statistics of U.S. Businesses (SUSB) program, better alignment with CBP data, and improvements to the linking methodology of the production processing (U.S. Census Bureau, 2020).

Using the `bds2018_cty` file, I extract county-year data documenting the number of firms, establishments, and employment for the years 2010-18. Similar to the CBP data, the BDS data do not offer a complete time series for two counties. These include King County, Texas (FIPS=48269) and Loving County, Texas (FIPS=48301). To maintain a balanced panel, I drop both counties.

To examine potential broadband-induced effects on firms by size, I incorporate data from the `bds2018_cty_fsize` file. For each county-year, these data report the number of firms, establishments, and employment (among other business dynamics) by three firm sizes: firms with fewer than 20 employees, firms with at least 500 employees, and firms that fall in-between. In keeping with the Small Business Administration's (SBA's) definition of "small" firms, I generate variables for `small` and `large` firms, establishments, and employment. However, some county-size-year cells remain suppressed due to data quality concerns (with cells set to `"(S)"`) or confidentiality reasons (with cells set to `"(D)"`). This is problematic when combining the two smaller size classes. Suppressed values would introduce substantial measurement error when combining to create measures for small and large firms. Therefore, I keep only the counties that have complete data for both size classes (firm employment of 1-19 and 20-499) and all years between 2010 and 2018. This procedure removes a total of 51 counties, primarily in Montana, Nebraska, and Texas.[40] To construct variables for large firms, I simply subtract small-firm statis-

[40]Counties with insufficient data to construct small-firm statistics have the following FIPS codes: 6091, 8033, 8053, 8079, 8111, 13101, 13265, 16033, 29227, 30011, 30033, 30037, 30045, 30069, 30075, 30103, 30109, 31005, 31007, 31009, 31075, 31085, 31091, 31103, 31113, 31165, 31183, 32009, 32029, 35021, 38007, 46061, 48033, 48045, 48081, 48101, 48137, 48155, 48261, 48263,

tics from the totals derived from the bds2018_cty file.

In a similar fashion, I incorporate data that classify firms by age, using the bds2018_cty_fage file. For each county-year, these data report the number of firms, establishments, and employment by several firm-age groups. Firms are placed into age groups based on the difference between their birth year (i.e., first year in which they reported positive employment) and the current year. These include the following: start-ups, firms that have been in operation for 1-5, 6-10, and 11+ years, plus firms that began operations before 1975, the start of the LBD data series.

In constructing variables for young and mature firms, some county-year cells remain suppressed in the public-use files, as outlined above with the bds2018_cty_fsize file. To ensure that only counties with complete data for the 2010-18 period are used, I drop incomplete counties from the panel. This procedure removes 85 counties that have censored new-entrant firm data for at least one of the years spanning the 2010-18 period.[41] More than 60 percent of these incomplete counties are in Texas, Nebraska, Montana, and South Dakota. I define young firms as those that have been in operation for less than six years, including new entrants. The remaining set of firms, those that been in operation for at least six years, are defined as mature firms. I subtract young-firm statistics from the totals derived from the bds2018_cty file.

B.6 *Business Formation Statistics*

To better understand how greater connectivity impacts early-stage business formation, I turn to the Business Formation Statistics (BFS). The BFS data report the number of requests for Employer Identification Numbers (EINs) (i.e., business applications) at the county level for the 2005-2019 period. The Census Bureau receives information from the Internal Revenue Service (IRS) weekly regarding Form SS-4, the form which individuals submit to obtain an EIN. This form can be found at https://www.irs.gov/pub/irs-pdf/fss4.pdf. These data, along with other aggregations and frequencies, are available at the following link: https://www.census.gov/econ/bfs/index.html.

48269, 48301, 48311, 48345, 48393, 48433, 48443, 48447, 49009, 49031, and 55078.

[41] Counties with incomplete young/mature firm data have the following FIPS codes: 6003, 8011, 8025, 8053, 8111, 13007, 13061, 13125, 13265, 13307, 16025, 16033, 20083, 21039, 21063, 21165, 21201, 26083, 28009, 28055, 29197, 30011, 30033, 30037, 30051, 30055, 30069, 30075, 30103, 30109, 31005, 31007, 31009, 31061, 31069, 31075, 31085, 31091, 31103, 31113, 31115, 31117, 31133, 31149, 31165, 31171, 31183, 32009, 35021, 35033, 38007, 38085, 38087, 38095, 46017, 46031, 46063, 46069, 46071, 46075, 46085, 46095, 48011, 48033, 48045, 48079, 48081, 48101, 48155, 48173, 48261, 48263, 48271, 48275, 48311, 48345, 48359, 48393, 48413, 48443, 48447, 49009, 49031, 53023, and 55078.

The BFS data offer another view new-firm births. To be sure, not all business applications turn into employer businesses. As outlined in Bayard et al. (2018), across all business application types (i.e., from sole proprietors, partnerships, and corporations), 13.6 percent business applications transition into employer businesses within four quarters; nearly 16 percent transition within eight quarters.

B.7 *Quarterly Census of Employment and Wages*

To provide an additional glimpse at the county-level employment and establishment responses to greater broadband connectivity, I incorporate data from the Quarterly Census of Employment and Wages (QCEW) program. The data can be accessed at the following link: `https://www.bls.gov/cew/downloadable-data-files.htm`. The QCEW data provide a comprehensive tabulation of the number of establishments, their employment levels, and wages paid to employees covered under the State Unemployment Insurance (UI) program.

Using the QCEW data, I incorporate county-year data measuring the number of establishments, average annual employment, total annual wages, average weekly wages, and average annual wages into my main panel dataset. Due to confidentiality censoring, four counties remain incomplete. These include Ringgold County, Iowa (FIPS=19159), Adams County, Iowa (FIPS=19003), Sullivan County, Pennsylvania (FIPS=42113), and Loving County, Texas (FIPS=48301). To maintain a balanced panel, I drop these counties from the data. I also address a peculiarity in the data for Shannon County, South Dakota (formerly Oglala Lakota County). For three data series (i.e., annual establishments, annual employment, and annual total wages) in 2015, it appears that the true values have been split between FIPS=46102 and FIPS=46113.[42] I simply add these together before combining with the main panel dataset.

B.8 *SOI Tax Stats - Migration Data*

To understand the migration patterns occurring between counties during the 2010-18 sample period, I turn to the IRS's SOI Tax Stats program. Available for tax-filing years 1991-2019, the IRS data document the number of households reporting year-to-year address changes from federal income tax returns. The data are available at the following link: `https://www.irs.gov/statistics/soi-tax-stats-migration-data`. For inflows (i.e., households moving *into* a county) and outflows (i.e., households moving *out* of a county), the IRS data define the total number of income tax returns, the number of income tax returns from intra-state and inter-state migration,

[42] For example, to generate the true measure of annual establishments for Shannon County, South Dakota in 2015, I add 59 (from FIPS=46102) to 59 (from FIPS=46113) to get 118.

as well as the number of returns for non-movers. In addition, the data also report the adjusted gross income (AGI) for each of these categories.

While the IRS data do allow for (and report) adjusted gross deficits, a value of −1 for data measuring the number of returns or adjusted gross income signals data suppression. This usually occurs in smaller counties. Importantly, county-year cells must have at least 20 returns to be identified in the data. In total, 87 counties have at least one year in which data is missing for one or more of the following variables: the total number of returns for out-migrants, the total number of returns for in-migrants, the AGI for out-migrants, and the AGI for in-migrants.[43]

B.9 *Antenna Structure Registration*

In this paper, data documenting the number of cell towers by county is rendered from the FCC's Antenna Structure Registration (ASR) database. As per federal law, all antenna (cell tower) structures must be registered with the Federal Aviation Administration (FAA) if taller than 200 feet or located near an airport. The ASR data report detailed characteristics of each registered cell tower along with its longitude and latitude coordinates. Importantly, the data report when the cell tower was constructed and, if applicable, when it was dismantled. Using these data, I create a county-year panel of cell towers and incorporate this into the main panel dataset.

[43] Counties with incomplete IRS migration data have the following FIPS codes: 6003, 8017, 8033, 8053, 8061, 8079, 8111, 16025, 16033, 20023, 20033, 20071, 20075, 20089, 20157, 20183, 20187, 20199, 28055, 29227, 30011, 30019, 30033, 30037, 30051, 30055, 30069, 30075, 30079, 30103, 30109, 31005, 31007, 31009, 31015, 31057, 31075, 31085, 31091, 31103, 31113, 31115, 31117, 31149, 31165, 31171, 31183, 32009, 35011, 35021, 38001, 38007, 38033, 38037, 38047, 38083, 38087, 40025, 46017, 46021, 46049, 46055, 46063, 46069, 46075, 46089, 46095, 46097, 46105, 46107, 46119, 46137, 48033, 48101, 48155, 48261, 48263, 48269, 48301, 48311, 48345, 48393, 48443, 49009, 49031, 51091, and 56027.