AN INDEX FOR ONLINE MICROBUSINESSES

July 2021

What Drives Microbusiness Formation and Growth?

An Analysis with a New Index for Online Microbusinesses[1]

Leila Bengali, Anderson School of Management, UCLA

William Yu, Anderson School of Management, UCLA

June 28, 2021

Abstract

Small businesses have been hit hard by the recession that began in 2020. At the same time, there has been a substantial increase in business formation. Do these patterns appear in microdata? What local factors make some areas more likely than others to see small business formation and growth? What does this formation and growth mean for local economies and labor markets? We provide a microbusiness index that can address these questions. We focus specifically on very small businesses with an online presence. Using proprietary data from GoDaddy, one of the leading providers of internet domain names, we built a new index of microbusinesses. This index is monthly and down to the county level. We then used the index to study what local factors increase the likelihood that microbusinesses will form and grow and also how online microbusinesses impact local economies. We found evidence that access to broadband, a skilled labor force, training, and capital are factors that may support microbusiness formation and growth. We also found that microbusiness formation and growth may boost local economic activity measured in government labor market surveys.

Keywords: small business, microbusiness, ecommerce, economic index, human capital, broadband, access to capital, local labor markets

JEL classification codes: C43, M13, D24, I23, E22

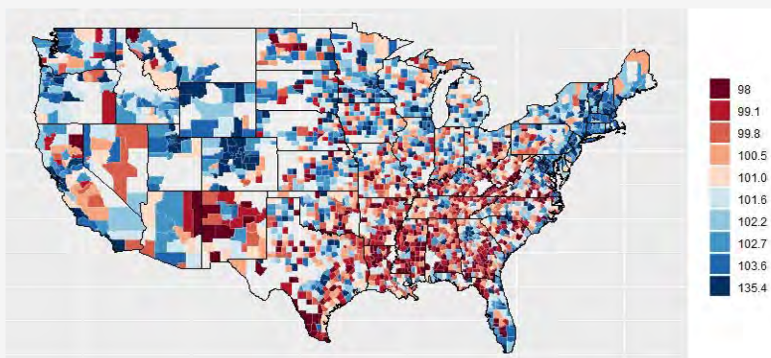
1. *Corresponding author. E-mail: leila.bengali@anderson.ucla.edu. Anderson School of Management, UCLA, 110 Westwood Plaza, Los Angeles, CA 90095. E-mail: william.yu@anderson.ucla.edu. Phone number: 310-825-7805. UCLA Anderson Forecast, 110 Westwood Plaza, Los Angeles, CA 90095.* *This research was funded by a grant from GoDaddy Inc. We are grateful to Ed Leamer, Jerry Nickelsburg, and Leo Feler for their helpful comments. We also thank the members of the Venture Forward team from GoDaddy for their assistance with data access and for their discussions that helped us better understand the data and GoDaddy's customers.*

Executive Summary

Small businesses have been hit hard by the recession that began in 2020. At the same time, there has been a substantial increase in business formation and growing research interest in small businesses. We want to understand specifically how small businesses with an online presence ('microbusinesses') fared during the pandemic, how they are distributed across the U.S., and what local factors contribute to their formation and growth. We provide the first microbusiness index, at the national, state, CBSA, and county level at a monthly frequency, that can address these questions.

To create this index, we used proprietary data from GoDaddy, the leading provider of internet domain names with a U.S. market share of about 50%. The data include information about GoDaddy customers (the business owners) and characteristics about their online businesses. The GoDaddy data suitably capture online microbusinesses because (1) GoDaddy's customers are mostly using the domains for commerce, (2) the majority of these businesses have ten or fewer employees, and (3) only 14% conduct business operations exclusively in-store. Our aim was to create an index that measures online microbusinesses and that picked up the facets of online microbusinesses that correlate with local labor markets and economic activity. To do this, we used statistical techniques to combine characteristics about the microbusiness owners and their businesses into an index. The figure below shows the microbusiness index across counties in the U.S. Counties that have high (blue) or low (red) values of the index tend to be adjacent to other counties that have high or low values of the index. Broadly, counties in the Southeast tend to have lower values of the index and counties in the Northeast, the West, and the mountain states tend to have higher values of the index.

Microbusiness Index Across Counties, March 2021



Note: Counties in grey have no available data for the index. Blue colors indicate higher values.

Sources: GoDaddy and UCLA Anderson Forecast

Having formed a microbusiness index, we used this index to study what local factors contribute to microbusiness success and how microbusinesses interact with local economies. Our main findings are summarized as follows:

- **Digital infrastructure:** In a companion paper, Yu and Bengali (2021) showed that counties with a higher fraction of residents that have broadband access tend to have stronger labor market outcomes.[2] We found broadband facilitates the formation, and enhances the success, of online microbusinesses.

- **Education and access to skills training:** We studied the relationship between human capital and access to skills training and microbusinesses. We used educational attainment to capture human capital and the number of postsecondary institutions and the number of community colleges to capture access to skills training. We found that even after controlling for various county characteristics, at least some of the measures of human capital and skills training are positively related to online microbusinesses.

- **Access to capital:** We used two methods to study the relationship between access to capital and microbusinesses: a case study and a regression analysis. The case study is inconclusive about the effect of access to capital on microbusinesses. In the regression analysis, we estimate the relationship between access to capital and microbusinesses after controlling for local county characteristics. We used the number of banks, the fraction of banks that are community banks, and the number of Paycheck Protection Program loans up to 150K to capture access to capital. The regression analysis provides some evidence that formal lending institutions help online microbusinesses, but also suggests that these online microbusinesses need to access capital in a different way or in smaller quantities than do other businesses.

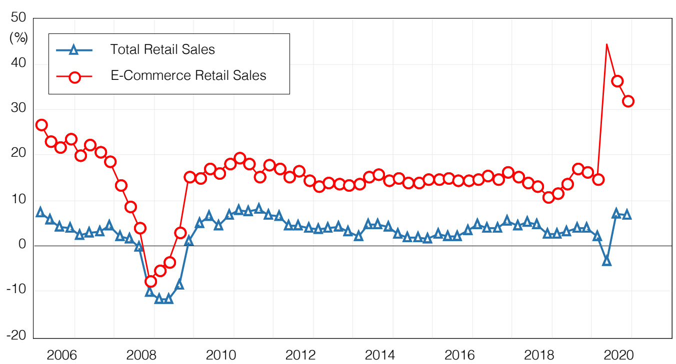
- **Microbusinesses and the local economy:** We found that the presence of online microbusinesses coincided with stronger labor market outcomes. After controlling for other factors, a county with more online microbusinesses per capita tended to have stronger labor markets as measured by the unemployment rate and the employment to population ratio. This finding, combined with survey evidence from GoDaddy, is suggestive that online microbusinesses contribute to economic growth at the local level.

This research has established relationships between selected inputs to business success by demonstrating that our microbusiness index can be used to examine such relationships. We view the work in this paper as first steps in a larger research agenda on online microbusinesses about what leads to microbusiness success at a local level.

2. *Yu, William, and Leila Bengali. (2021). "Digital Infrastructure, the Economy and Online Microbusiness: Evidence from GoDaddy's Microbusiness Data." UCLA Anderson Forecast Quarterly Report. June. pp. 81-99.*

The COVID-19 pandemic devastated millions of lives and disrupted economies across the world. The global pandemic has also caused an acceleration of two business trends: the adoption of e-commerce and the move to remote work. Figure 1 shows the year-over-year growth of total and e-commerce retail sales. E-commerce sales growth rates, which were already higher than total sales growth rates prior to the pandemic, skyrocketed in the pandemic. The e-commerce giant Amazon and other big tech companies that provided online platforms and services became the winners in this global health crisis. Overall, business formation was markedly higher in 2020 than in prior years (U.S. Census Business Formation Statistics, Fazio et al., 2021); however, little is known about the trends for small businesses. Businesses with fewer than five employees employ over 5.8 million workers in the U.S.[3] In addition, there were 26.5 million sole proprietors nationwide.[4] Even less is known about small businesses equipped with online platforms. Was there growth in small online business formation? Were those businesses more active as economic activity moved online in response to the pandemic? How are these small businesses geographically distributed across the U.S.? What local factors explain the geographic variation in small business formation and success? What impacts do these small businesses have on local economies?

Figure 1. Year-over-year Growth Rates of Total U.S. Retail Sales and E-Commerce Sales



Source: U.S. Census Bureau

There is a growing literature studying the impact of COVID-19 on small businesses (Alekseev et al., 2020; Barrlett, et al., 2020; Bartik, 2020; Chetty et al., 2020; Kim et al., 2020). These papers add to existing work about how economic conditions differentially affect small businesses relative to larger ones (e.g. Crouzet and Mehrotra, 2020). For example, Kim et al. (2020) using a financial account dataset from JPMorgan Chase Institute containing 380,500 businesses and 333,000 business owners found that small business revenues and owners' consumption both declined by roughly 40% compared to the pre-pandemic level. Chetty et al. (2020) built a public database that tracks small business transactions and revenue and found that Paycheck Protection Program (PPP) loans had small impacts on boosting employment. To the best of our knowledge, there is little research on online small businesses in the literature, and our paper helps fill the gap. Mossberger et al. (2020 and 2021), and Mossberger and Tolbert (2021) are exceptions. They study the link between online small businesses and various measures of economic vitality and generally find positive relationships between online small businesses and prosperity, income, and the unemployment rate. This is an important gap to fill because many businesses had to turn to e-commerce when the pandemic began. For example, in a Facebook survey of business owners, Alekseev et al. (2020) find that the pandemic led 61% of businesses to increase their online presence.

We want to understand how small online businesses fared during the pandemic and what factors contribute to their formation and growth. In order to do this, we used data from GoDaddy, the leading provider of internet domain names with a U.S. market share of about 50%. According to a GoDaddy survey,[5] GoDaddy's customers are mostly small businesses or nonprofits, 55% of which are sole proprietorships and 37% of which are small businesses with ten or fewer employees. We refer to this group of small online businesses as 'microbusinesses.' We combined these data with American Community Survey data to create an index that measures microbusiness prevalence and activity over time and across regions in the U.S. Our index is necessary to study online microbusinesses because these microbusinesses are not well represented by existing surveys of labor market and business activity. Moreover, county and CBSA (Core-Based Statistical Area)-level labor market information is available only at a lag. Our index is also designed to be contemporaneous so that the index can be used as a nowcast indicator of local economic activities at detailed geographic levels.

Our index helps address the void in knowledge about this under-studied type of small business. The index allows us to answer questions about what local factors contribute to microbusiness success and about how these microbusinesses contribute to local labor market conditions. We found that counties with more broadband, human capital (both access and educational attainment), and funding opportunities tend to have more microbusinesses. In addition, we found that the presence of online microbusinesses coincides with stronger labor market outcomes. This finding, combined with the aforementioned survey evidence from GoDaddy, is suggestive that online microbusinesses contribute to economic growth at the local level.

The paper contributes to the literature in four respects: (1) We analyzed a unique dataset provided by GoDaddy to create a new indicator of online microbusinesses. The data are much more comprehensive, with 11 million microbusiness owners, than those in the current research on small businesses. (2) Our index can track microbusiness formation, activity, and local labor market conditions in real time and across regions. This represents a valuable improvement over existing measures that are not

3. *According to BLS Business Employment Dynamics, 2020.*

4. *They are so-called self-employed individuals, or gig-workers (Census non-employer Statistics, 2018).*

5. *This is based on a survey that GoDaddy conducted in July 2020. The survey was conducted independently from the research project described in this paper. The survey was sent to a randomly selected subset of its customers. The number of respondents is 2,330. References to the GoDaddy customer survey later in this paper all refer to this survey.*

available with a similar frequency or geographic granularity. (3) The geographic granularity of our index allows us to answer policy-relevant questions about what factors help support the success of microbusinesses. (4) We showed that there is a strong link between microbusinesses and local economic outcomes.

The rest of the paper is organized as follows. Section 1 describes the methods and data used to create the index. Section 2 covers the empirical patterns in the index over time and across space. Sections 3 and 4 present analyses that used the index to discuss how local factors (e.g. human capital, skills training, access to capital) predict where microbusinesses are prevalent and where they succeed. Section 5 discusses the macroeconomic implications of online microbusinesses, and Section 6 provides concluding thoughts.

1. Methods: Creating a Microbusiness Index

What is not measured cannot be studied. As such, our first task was to create an index of online microbusinesses at a reasonable level of temporal and geographic detail. Our aim was to create an index that measures online microbusinesses, and that picked up the facets of online microbusinesses that correlate with local labor markets and economic activity.

A. Data

To accomplish these goals, we used proprietary data provided by GoDaddy. GoDaddy is one of the leading providers of domain names (with a market share of about 40%) and business website services with over 11 million customers who have over 40 million online microbusinesses in the U.S. and 20 million customers globally. We use data that were not restricted by how that domain name is used (e.g. for online retail, for informational purposes, or for email) or whether the domain name was linked to a publicly accessible website at the time the data were pulled from the database. Given the size of their business, GoDaddy's data provide a comprehensive picture of online microbusinesses. The reason we call them microbusinesses is because (1) GoDaddy's customers are mostly using the domains for commerce with 75% of domain name owners using their websites for businesses, and (2) these businesses are small: 55% are sole proprietorships and an additional 37% are small businesses with one to ten employees. The businesses are 'online' microbusinesses because according to GoDaddy's survey only 14% conduct business operations exclusively in-store.

We obtained monthly data aggregated to the zip code level from April 2020 through March 2021. Some variables are available

Table 1. GoDaddy Variables

Customers / Microbusiness owners	The number of unique individuals who have purchased a domain name from GoDaddy
Microbusinesses	The unique number of GoDaddy online microbusinesses. An online microbusiness, or microbusiness, is a unique domain name concept. For example, ABCbusiness.com and .net and .org are three domain names, but one microbusiness. One GoDaddy customer may have multiple microbusinesses.
The fraction of microbusinesses with a website	The fraction of all GoDaddy microbusinesses in which the microbusiness's domain name also has a publicly accessible website
WAM fraction	The fraction of microbusinesses that use GoDaddy's website design and marketing service (a service called Websites and Marketing, or WAM)
The average web traffic index to microbusiness websites	A measure of website traffic to the microbusiness's website (relative to web traffic to Google)
The average footprint index for microbusiness websites	An index giving an indication of the number of visitors to a microbusiness website, the turnover, and the microbusiness's size
The average heartbeat for microbusiness websites	A measure that reflects the frequency and amount of website updates by the microbusiness owner
The fraction of microbusinesses connected to a GoDaddy SSL	Connection to SSL allows connections between computers to be secure and is often used for online payments. GoDaddy is not the only provider of SSL services, and microbusiness owners can use other sources to make an SSL connection
Average microbusiness age	How long the microbusiness has been in existence, in days

Source: GoDaddy

sporadically prior to April 2020, but for the sake of consistency, we limited our attention to data after April 2020 except where explicitly specified. The data were aggregated to county and CBSA levels by using zip code crosswalks available from the U.S. Department of Housing and Urban Development.[6] The list of variables and definitions is in Table 1. Due to data availability limitations, most of the variation used in the analysis is in the cross-section. Note that our data include the universe of GoDaddy microbusinesses in addition to the targeted subset described in Mossberger et al. (2020), and Mossberger and Tolbert (2021).

B. Index Creation

To measure microbusiness activity across the nation and over time, we created three indices: (1) Simple index: we assigned a weight of one to the microbusinesses per capita variable and a weight of zero to all other variables. This index assumes that the density of microbusinesses (the number of microbusinesses per 100 residents) is sufficient and other variables are trivial to measure microbusinesses; (2) Even-weight index: we assigned each of the K variables included in the index a weight of 1/K. This index assumes that it is valuable to include all relevant variables to better understand microbusinesses, but that each are equally valuable, important, and informative. (3) Baseline Index: we assigned each of the K variables a weight based on statistical techniques to capture various facets of online microbusinesses. We selected weights that maximize the correlation between these variables and local economic activity. As such, the Baseline Index could be used to understand online microbusinesses and the local economy in a more holistic way.

Variable Selection for the Baseline Index

The vitality of online microbusinesses can be comprised of three components: receptivity, reception, and activity. Receptivity is the physical and intellectual infrastructure needed to access and use the Internet. Reception is the number of microbusinesses and their owners as a percentage of the population of each locale. Note that the number of microbusinesses is larger than the number of owners because an owner could have more than one microbusiness. This component captures the extensive margin—the number of and change in the microbusiness formation. Activity is the frequency and intensity with which microbusiness websites are updated by their owners and used by their customers. These components capture the intensive margin—how active or successful microbusinesses are.

To capture receptivity, we used data from the American Community Survey 2019 five-year estimates at the county, CBSA, and state levels. The variables we used include (1) a weighted index of educational attainment developed by the UCLA Anderson Forecast and created from American Community Survey data on educational attainment.[7] This is called the City Human Capital Index (CHCI). Also included were: (2) the fraction of residents with broadband internet subscriptions, and (3) the fraction of residents with computer access. To capture reception, we used the number of GoDaddy customers and the number of GoDaddy microbusinesses per capita. To capture activity, we used the fraction of microbusinesses with a website, the fraction of microbusinesses that use GoDaddy's web design and marketing service, the average web traffic index, the average footprint index, the average heartbeat index, the fraction of microbusinesses connected to SSL, and the average microbusiness age. See Table 1 for a description of these variables.

Figures 2 – 9 present the variation in the cross section for these variables. The key message from these maps is that while the variables may be related, they have different geographic patterns. The maps do not all look exactly the same, meaning that each variable makes a unique contribution to the index. Still, there are associations between these variables. For example, Yu and Bengali (2021) showed that counties with a higher fraction of residents that have broadband access tend to have stronger labor market outcomes. They found broadband facilitates the formation, and enhances the success, of online microbusinesses.[8]

As another example, the geographic distribution of microbusinesses that use the website design and marketing service (WAM service, Figure 6) differs from the distribution of web traffic to microbusiness sites (Figure 7). For example, the inland areas of New York, Vermont, New Hampshire, and Pennsylvania have high web traffic but lower use of the WAM services, while coastal California tends to have lower web traffic and lower use of the WAM services. These two components, WAM and web traffic, capture different aspects of microbusiness activity. WAM service reflects business owners just starting to build an online presence who may not have the time, knowledge, or resources to build a website from scratch. Web traffic reflects demand for online shopping.

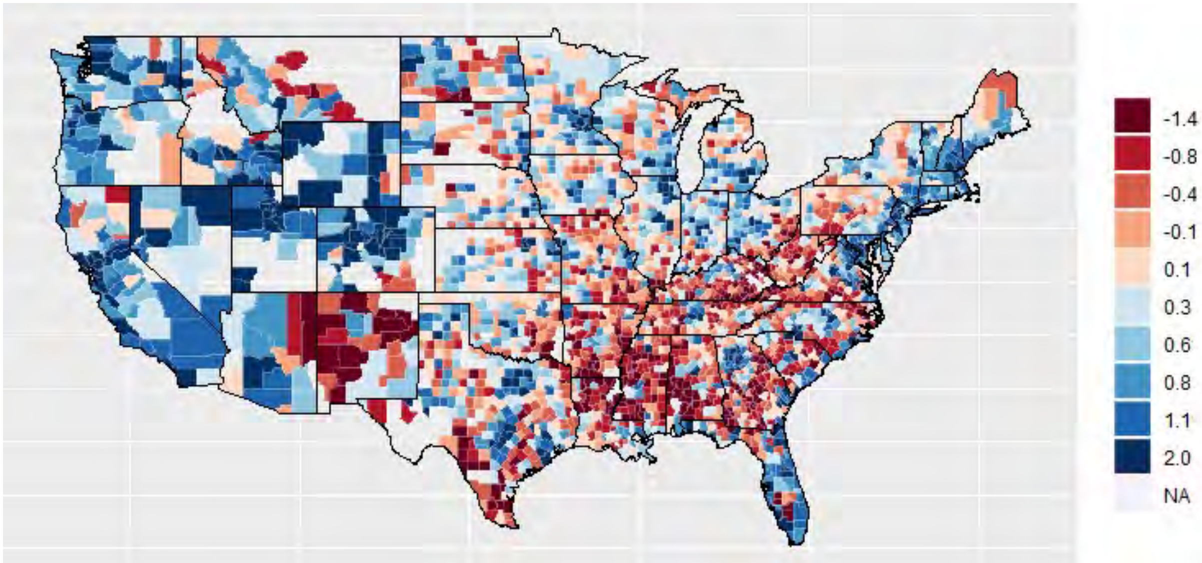
---

6. *When a zip code is assigned to more than one county or CBSA, the crosswalk file indicates the fraction of addresses in that zip code that correspond to each county or CBSA. In such cases, we used these fractions to proportionally allocate our zip code data to counties or CBSAs.*

7. *For details, see https://www.anderson.ucla.edu/centers/ucla-anderson-forecast/projects-and-partnerships/city-human-capital-index*

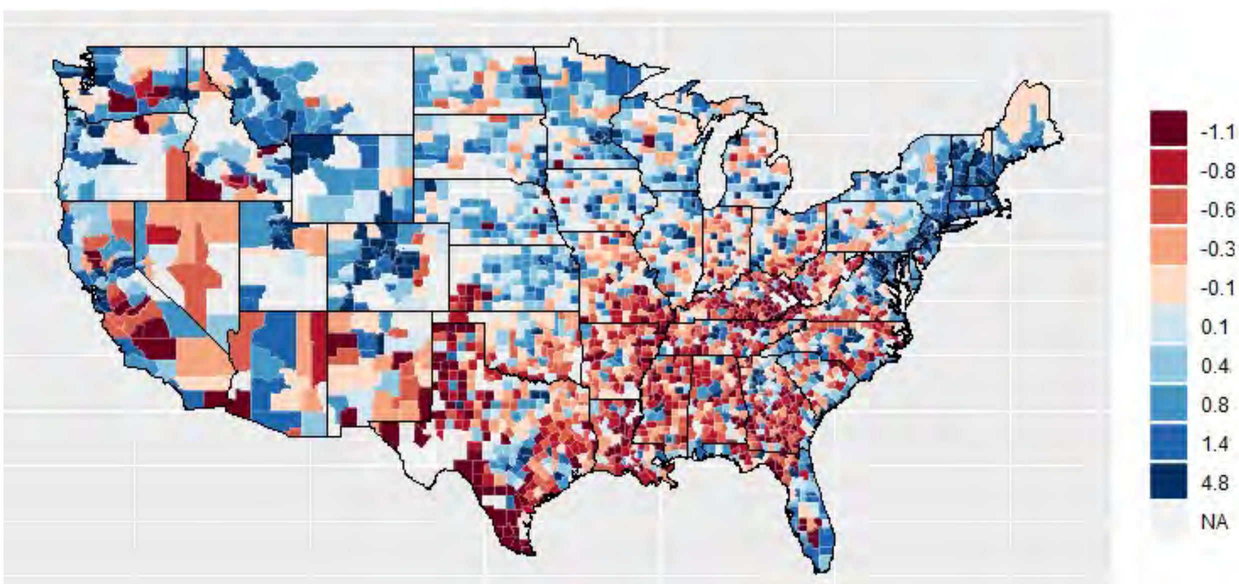
8. *See Appendix A for original percentages of broadband and computer connectivity, selected correlation charts, and the regression results.*

Figure 2. The Fraction of Residents with Computer Access



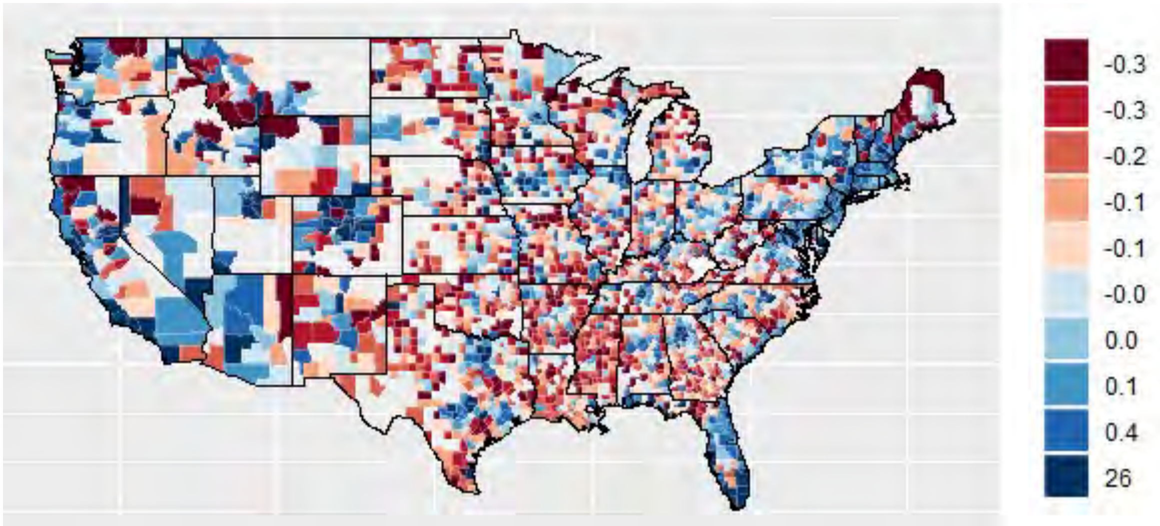
Note: Map shows normalized values. Blue colors indicate higher values.
Source: 2019 Five-year American Community Survey

Figure 3. City Human Capital Index (CHCI)



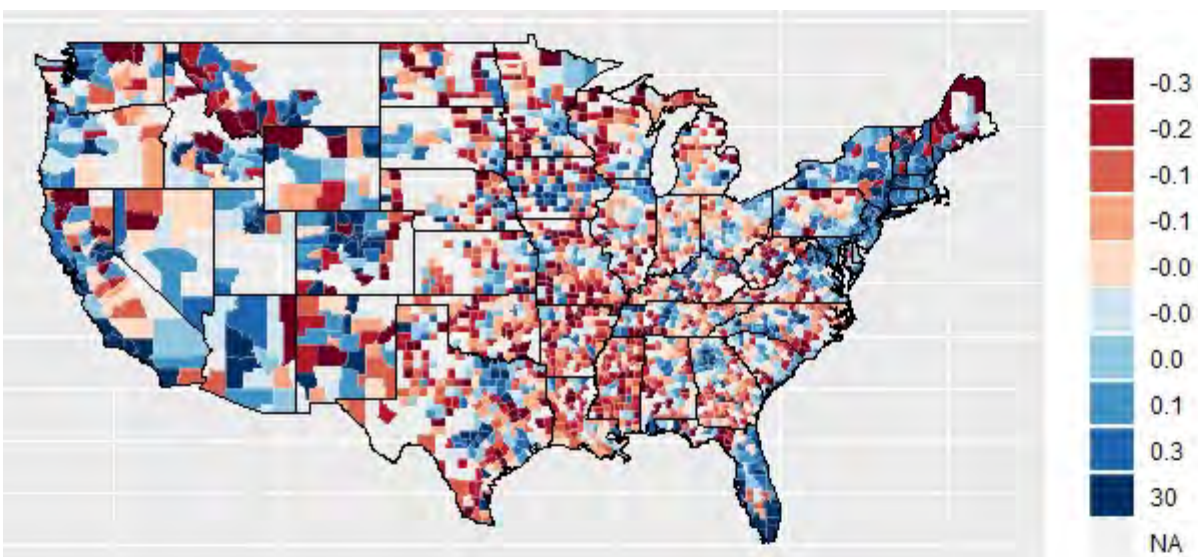
Note: The City Human Capital Index is a weighted average of educational attainment of adult residents. Map shows normalized values. Blue colors indicate higher values.
Source: 2019 Five-year American Community Survey

Figure 4. Density of Microbusinesses



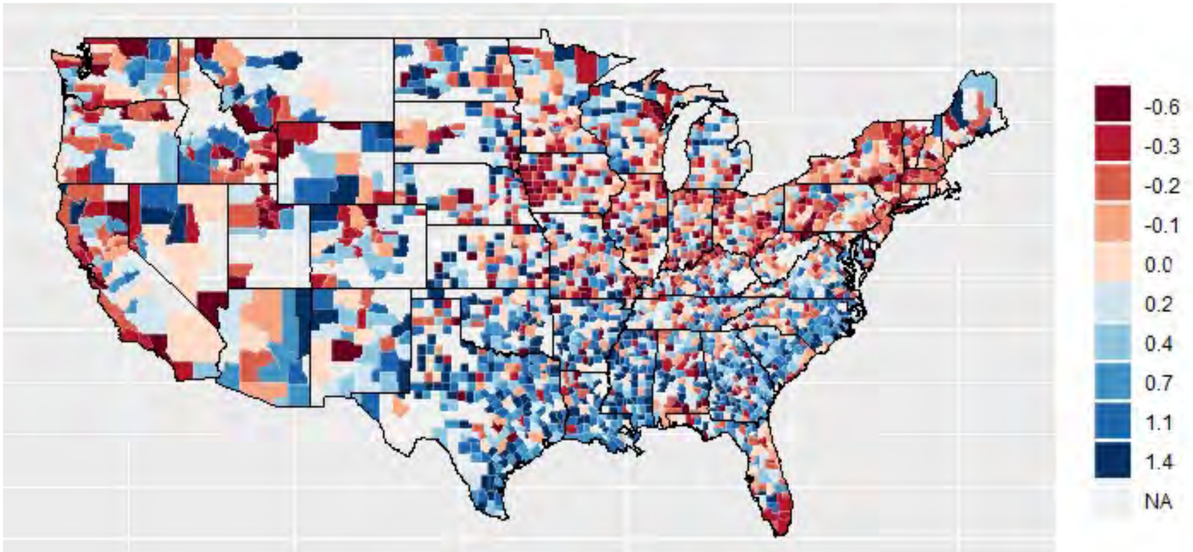
Note: Defined as the number of GoDaddy microbusinesses over the county population. A GoDaddy customer might own more than one microbusiness. Map shows normalized values. Blue colors indicate higher values. Data are for November 2020
Source: GoDaddy

Figure 5. Density of Microbusiness Owners



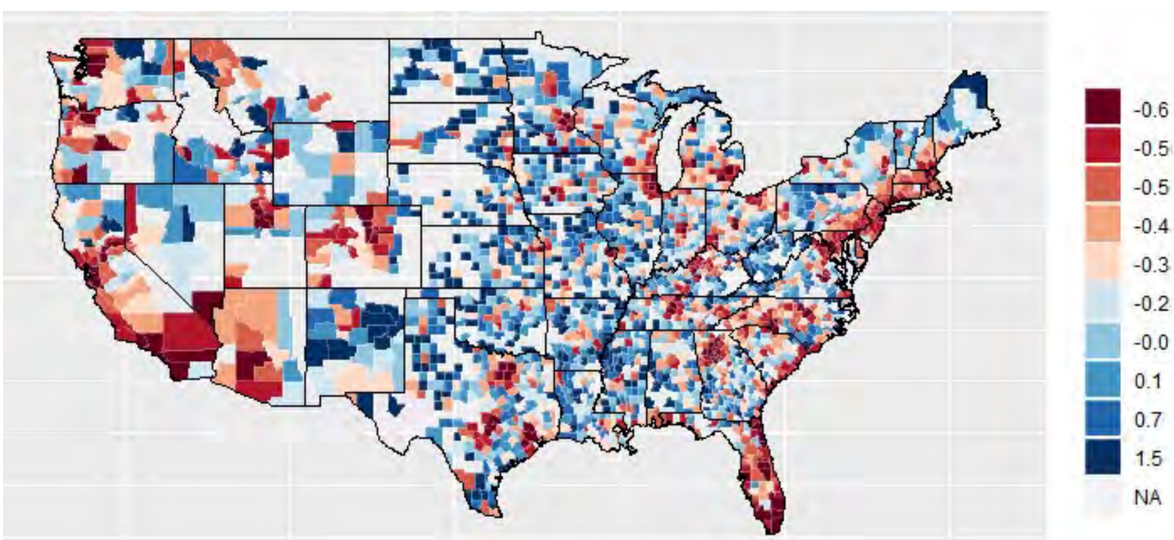
Note: Defined as the number of GoDaddy customers over the county population. A GoDaddy customer might own more than one microbusiness. Map shows normalized values. Blue colors indicate higher values. Data are for November 2020
Source: GoDaddy

Figure 6. Fraction of WAM Microbusinesses



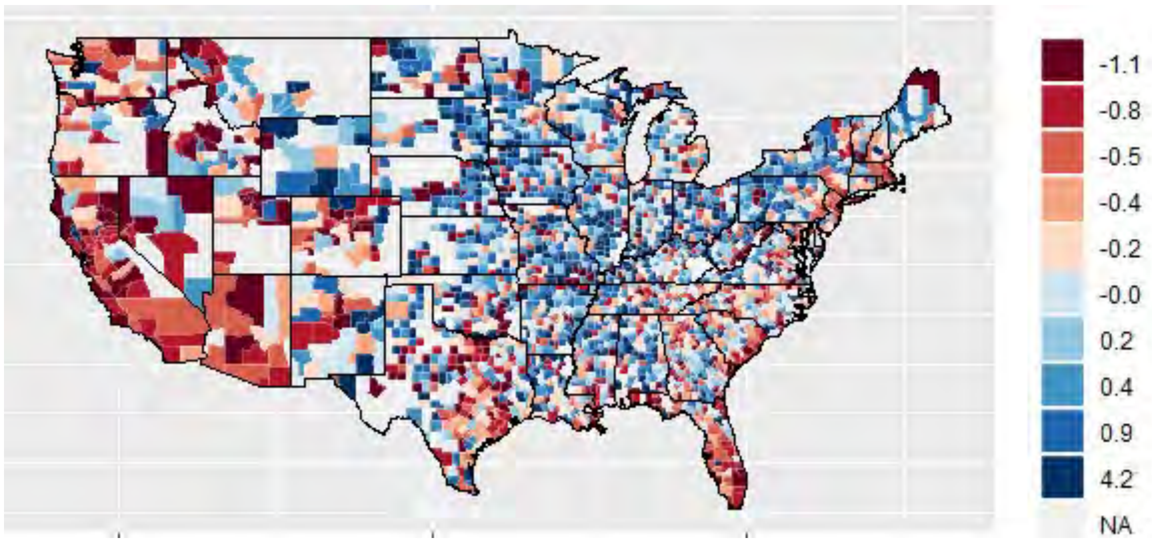
Note: Defined as the fraction of microbusinesses that use GoDaddy's web design and marketing service (WAM). Map shows normalized values. Blue colors indicate higher values. Data are for November 2020
Source: GoDaddy

Figure 7. Average Traffic Index



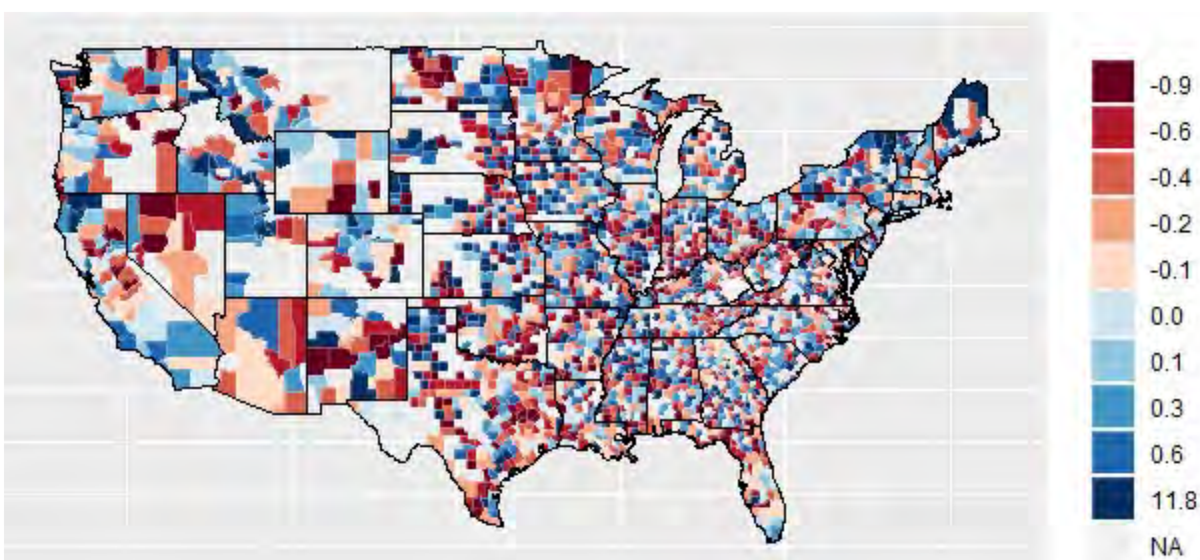
Note: Measured as website traffic to microbusiness's websites relative to web traffic to Google. Map shows normalized values. Blue colors indicate higher values. Data are for November 2020.
Source: GoDaddy

Figure 8. Fraction Connected to Website



Note: Defined as the fraction of all GoDaddy microbusinesses in which the microbusiness's domain name also has a publicly accessible website. Map shows normalized values. Blue colors indicate higher values. Data are for November 2020.
Source: GoDaddy

Figure 9. Average Footprint Index



Note: The footprint index for a microbusiness is an indication of the number of visitors, the turnover, and the microbusiness's size. Map shows normalized values. Blue colors indicate higher values. Data are for November 2020.
Source: GoDaddy

Creating Index Weights

Having selected the variables, we next determined how to combine them into a composite index by taking a weighted average of the selected variables. To determine the weights, we returned to our motivation for creating the index. The index would capture trends in online microbusinesses as well as contemporaneous movements in local labor markets. In order to achieve this goal, we used the selected variables' ability to explain existing measures of local labor market and small business activity to determine the weights.

Specifically, we wanted to determine the underlying relationship between local labor markets and online microbusinesses. To do this, we estimated a series of models using different measures of local labor market activity and different geographic levels of aggregation. Having estimated all of these models, we then compared and combined the estimates. All of the models were regressions of the form

$$y_{it} = \beta_0 + \beta_1 X_{it} + \alpha_t + e_{it},$$

where i indexes geographic area, t indexes time in months, y is one of several labor market and small business formation measures, X is a vector of the selected index variables, α_t are date fixed effects to flexibly capture national time trends, and e_{it} is the error term. We restricted the estimation data to April 2020 through October 2020 as the training set (in-sample). This left November 2020 to March 2021 as the testing set to test how well the index achieves our goal to pick up contemporaneous movements in local labor markets out of sample. We ran these regressions using data that were first normalized across all geographic areas (i) and dates (t). To ensure that the data used in estimation accurately reflected microbusiness activity, we imposed some restrictions.[9] Any geographic area in which the number of microbusinesses ever more than doubled or shrunk by more than half from one month to the next was dropped from the analysis used to calculate index weights. This was done based on guidance from GoDaddy. The number of microbusinesses is generally stable from month to month. Large fluctuations likely indicate third party bulk purchases or sales of GoDaddy domain names and thus do not reflect microbusiness activity. We also dropped any geographic area in which the number of microbusinesses was ever below the 5th percentile. This restriction was made because the way in which GoDaddy can map businesses to locations does not guarantee a perfect match to the business's true location. Areas with very few microbusinesses likely reflect these matching imperfections. The key parameter from these regressions is the vector of regression coefficients, β_1. Since both dependent and explanatory variables in the regressions were normalized, the value of β_1 indicates the relative importance of each variable in explaining variation of dependent variables. Therefore, it can be used directly for computing optimal weights for the index.

Table 2 summarizes the dependent labor market variables and geographic levels we used for our twelve regressions. From these twelve regressions, we tried seven methods to calculate optimal weights for the index. These included excluding coefficients from these twelve regressions based on sign or significance, a

Table 2. Regressions Used to Determine Index Weights

Geographic level (i)	Dependent variable (y)	Notes
County	Employment as a fraction of the population	
	Unemployment rate	
	Employment growth (month-over-month percent change)	
CBSA	Employment as a fraction of the population	
	Unemployment rate	
	Employment growth (month-over-month percent change)	
State	All new business applications per capita	
	High-propensity new business applications per capita	
	New business applications with planned wages per capita	
	Employment in establishments with fewer than five employees as a fraction of the population	Estimation time period is Jan. - Mar. 2020 due to data availability
	Average weekly wage in establishments with fewer than five employees	Estimation time period is Jan. - Mar. 2020 due to data availability
	Employment growth (month-over-month percent change)	

Sources: U.S. Bureau of Labor Statistics, Local Area Unemployment Statistics (employment, unemployment rate); U.S. Census Bureau (population); Business Formation Statistics, U.S. Census Bureau (business applications; see https://www.census.gov/econ/bfs/definitions.html for definitions of types of business applications); U.S. Bureau of Labor Statistics, Quarterly Census of Employment and Wages (employment statistics for small establishments).

9. *Weights created without these restrictions are similar to those created with the restrictions.*

stepwise regression, and a principal component approach. See Appendix B for details about the seven methods and the selection process. Table 3 shows the weights on each of the selected variables for these seven methods. The method we selected for the Baseline Index sets coefficients with a sign opposite of the predicted sign to zero. This imposes some subjective judgement based on our ex-ante expectations about the coefficients. We selected this index as our baseline because of its best fit in the testing set (see Appendix B for details). After dropping the opposite-signed coefficients, each of the twelve vectors of coefficients were then normalized so that the non-zero entries summed to one. To create the final weight for each variable, we took the simple average across these twelve vectors.

With the weights in hand, the process of creating the composite Baseline Index (or 'composite index') was straightforward. We: (1) took the full set of normalized data (April 2020 – March 2021), (2) truncated each of the selected index variables at the 95th percentile, and then (3) created the composite index by multiplying the weights in the 'baseline' column of Table 3 with corresponding index variables. The truncation in (2) was based on guidance from GoDaddy. As noted earlier, we knew from detailed data discussions with GoDaddy that extreme fluctuations in the recorded number of microbusinesses and microbusiness owners from month to month likely do not reflect true changes in microbusiness activity. In addition, outliers in the other index variables likely reflect transitory shocks that do not accurately capture the underlying construct our index is meant to capture. An example of this is a spike in web traffic due to a microbusiness being mentioned in the news. The method of truncation used here differs from that described earlier to calculate the weights. The reason is that we wanted to limit the influence of outliers without reducing the number of areas for which we can calculate the index. The method used to determine the weights drops outliers, rather than truncating them. In addition to the composite index, we created sub-indices for reception, receptivity, and activity using the variables that pertain to each category multiplied by their respective weights. We then re-scaled all indices (the composite and the three sub-indices separately) and centered them to average 100 in April 2020.

Table 3. Index Weights

	Index Weights								
Variable Component	Naïve/ simple	Even Weighted	Weighted (baseline)	Weighted (+ broadband* computer interaction)	Weighted (cross-section estimation)	Weighted (stepwise regression)	Weighted (allow opposite signed-coefficients)	Weighted (significant coefficients only; allow opposite signs)	Weighted (PCA)
chci (city human capital index)	0.0%	7.1%	10.8%	7.4%	11.1%	10.0%	53.5%	38.7%	10.7%
computer	0.0%	7.1%	19.4%	13.5%	19.7%	10.0%	29.1%	52.7%	9.9%
broadband	0.0%	7.1%	1.2%	8.1%	1.2%	4.7%	-14.7%	-41.4%	10.1%
computer*broadband interaction	0.0%	0.0%	0.0%	0.5%	0.0%	0.0%	0.0%	0.0%	0.0%
microbusiness owners per 100 people (density)	0.0%	7.1%	4.3%	1.9%	4.7%	10.0%	6.6%	-12.6%	3.4%
microbusinesses per 100 people (density)	100.0%	7.1%	5.3%	9.8%	4.7%	8.8%	3.7%	15.1%	3.4%
microbusiness owners density, m-m % change	0.0%	7.1%	4.6%	4.7%	5.0%	0.0%	33.0%	24.1%	22.1%
microbusiness density, m-m % change	0.0%	7.1%	4.7%	4.8%	6.9%	3.8%	-2.5%	2.2%	21.6%
footprint	0.0%	7.1%	4.6%	4.1%	5.6%	10.0%	3.2%	5.8%	0.0%
heartbeat	0.0%	7.1%	5.7%	4.5%	2.9%	0.0%	7.2%	6.4%	0.0%
web traffic	0.0%	7.1%	19.7%	21.0%	17.5%	0.0%	0.8%	4.3%	1.7%
average microbusiness age	0.0%	7.1%	2.3%	5.6%	2.9%	8.8%	-22.3%	-16.7%	0.0%
fraction connected to SSL	0.0%	7.1%	5.2%	4.0%	4.9%	15.0%	-6.3%	-2.5%	8.3%
fraction connected to website	0.0%	7.1%	8.7%	6.4%	10.0%	15.0%	-7.6%	12.9%	8.8%
fraction using WAM service	0.0%	7.1%	3.7%	3.7%	2.9%	3.8%	16.4%	11.0%	0.0%

Note: The detailed descriptions are in Appendix B.

Testing the Indices

To confirm that the Baseline Index successfully captures contemporaneous local labor market conditions, we assessed the index's explanatory power using data from November 2020 – March 2021, dates which were not used to calculate the variable weights. We compared the performance of the simple index and the even-weight index to the performance of the Baseline Index to get a sense of how robust the Baseline Index is to different ways to calculate weights.

Table 4 and Table 5 present sample regression results showing how the Baseline Index, the simple index, and the even-weight index correlate with local labor market conditions. The regressions are of the form

$$y_{it} = \beta_0 + \beta_1 y_{it-1} + \beta_2 Index_{it} + \alpha_t + g_s + e_{it},$$

where i indexes geographic area (either county, CBSA, or state), t indexes time in months, y is a measure of labor market activity (such as the unemployment rate or employment as a fraction of the population), α_t are time fixed effects, g_s are state fixed effects, and $Index_{it}$ is either the Baseline Index, the simple index, or the even-weight index. We used a short panel of data from November 2020 – March 2021. We included the lagged dependent variable as an explanatory variable because these types of labor market data are highly persistent.

Table 4. Microbusiness Indices and Employment to Population (County Level, Panel)

	Dependent variable: employment to population (epop)		
Independent variables	(1)	(2)	(3)
Lag(epop)	0.996***	0.997***	0.997***
	(0.001)	(0.001)	(0.001)
Baseline index	0.003***		
	(0.001)		
Even-weight index		0.008***	
		(0.002)	
Simple index			0.001***
			(0.0002)
Constant	-0.309***	-0.880***	-0.173***
	(0.052)	(0.172)	(0.043)
Observations	13,233	13,233	13,340
Adjusted R^2	0.995	0.995	0.995

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. All have time fixed effect and state fixed effect. Sample period is from November 2020 to March 2021.*

Table 5. Microbusiness Indices and the Unemployment Rate (County Level, Panel)

	Dependent variable: unemployment rate (urate)		
Independent variables	(1)	(2)	(3)
Lag(urate)	0.890***	0.894***	0.895***
	(0.003)	(0.003)	(0.003)
Baseline index	-0.005***		
	(0.0005)		
Even-weight index		-0.011***	
		(0.002)	
Simple index			-0.001***
			(0.0002)
Constant	0.805***	1.407***	0.428***
	(0.058)	(0.181)	(0.037)
Observations	13,238	13,238	13,345
Adjusted R^2	0.941	0.941	0.941

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. All have time fixed effect and state fixed effect. Sample period is from November 2020 to March 2021.*

We checked whether these indices provide explanatory power above and beyond y_{it-1} because y_{it-1} alone is highly predictive of y_{it}.[10] Purely cross-sectional single month versions of the model that omit date fixed effects show very similar results. The indices are validated by two results. First, the index correlates with current local labor market conditions. When the Baseline Index increases by one index point, the employment to population ratio tends to increase by 0.003 percentage points (Table 4 Equation (1)) and the unemployment rate tends to fall by 0.005 percentage points (Table 5 Equation (1)). For the simple index, this value is 0.008 (-0.01 for the unemployment rate) percentage points and for the even-weight index and 0.001 (-0.001 for the unemployment rate). That is, even after controlling for labor market conditions in the prior month, the regression coefficients are statistically significant and of the predicted signs. Second, while there are slight differences in performance, the three indices generally perform similarly in terms of their ability to explain local labor market conditions. Given these findings, why not just use the simple index? As we will discuss in the next section, exclusively using the reception component fails to capture patterns in online microbusinesses that are picked up by the activity and receptivity components of the composite Baseline Index.

Finally, we compared our index to an existing index of small business activity. There are a limited number of existing indices of small business activity. To start, online microbusinesses are distinct from small businesses in general, so no exactly comparable index exists. For the existing small business indices, most are only available nationally, and the ones that are available at smaller geographic levels are available only for a subset of geographies. The most closely related index is the Paychex-IHS Markit Small Business Employment Watch,[11] an index based on data from businesses that use Paychex (a human resources software and service provider). Since the online microbusinesses that we study are small and may not be large enough to need human resources software, we view our index as complementary to the Paychex-IHS Markit index. Table 6 provides some evidence that the two indices contain different information.[12] We ran regressions of two dependent variables on our index: 1) employment as a fraction of the population, and 2) new business applications with planned wages per 1000 residents. The results indicate that our index has additional explanatory power when competing with the Paychex-IHS Markit index: in columns (1) and (5), the coefficients on both the Baseline Index and the Paychex-IHS index are statistically significant.[13] These results are an indication that these two indices likely measure different aspects of local economies, and that the Baseline Index provides additional information about local labor market conditions.

Table 6. Microbusiness Index Vs Paychex-IHS Index (State Level, Panel)

	Dependent variables							
	epop employment to population ratio (%)				biz apps with wages per 1000 residents			
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Lag(epop)		0.935***	0.915***	0.912***				
		(0.015)	(0.029)	(0.032)				
Lag(biz apps)						0.931***	0.909***	0.902***
						(0.018)	(0.040)	(0.041)
Baseline index	0.400***	0.034**		0.008	0.002**	0.0001		0.001
	(0.077)	(0.015)		(0.036)	(0.001)	(0.0002)		(0.001)
Paychex index	0.491***		0.022	0.031	0.010***		0.002	0.002*
	(0.169)		(0.061)	(0.074)	(0.002)		(0.001)	(0.001)
Constant	-35.098*	-7.153***	-6.017	-7.607	-1.115***	-0.019	-0.157	-0.267*
	(21.115)	(1.186)	(6.094)	(9.109)	(0.272)	(0.020)	(0.095)	(0.149)
Observations	200	612	200	200	200	612	200	200
Adjusted R^2	0.424	0.928	0.895	0.894	0.533	0.877	0.870	0.870

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. All have time fixed effect and state fixed effect. Sample period is from April 2020 to February 2021.*

10. *This persistence is confirmed by the estimated coefficients on the lags of the y variables. The coefficients are statistically significant and close to one.*

11. *https://www.paychex.com/employment-watch/#!/*

12. *The regressions in the table used data from April 2020 to February 2021, due to data availability from the Paychex-IHS Markit index.*

13. *We ran other models that include lagged dependent variables as robustness checks to account for persistence in the dependent variables. In the table of results, we include results using the Paychex-IHS index for comparison. In column (2) when the employment to population ratio is the dependent variable, our index's coefficient remains statistically significant; however, this is not the case when business applications is the dependent variable (column (6)). In models that include both indices and also lagged dependent variables (columns (4) and (8)), the coefficient on our index falls in magnitude with no statistical significance. In column (8) when business applications per capita is the dependent variable, the Paychex-IHS Markit index coefficient remains statistically significant, though the magnitude is substantially reduced. This suggests that the Paychex-IHS Markit index is better for explaining variation in business applications with planned wages. This makes sense given that the data underlying the index come from businesses that use Paychex for payroll and HR.*

2. Empirical Evidence: The Microbusiness Index Over Time and Space

What does the index tell us about how online microbusinesses are distributed across space and how they have performed during the recovery from the recession that began in 2020? This section shows time series and cross-sectional patterns in both the composite index and the sub-indices that measure receptivity, reception, and activity.

Figure 10 shows the Baseline Index for the U.S. as a whole, along with the even-weight and simple indices for comparison. The Baseline Index time series shows a rise in May 2020, followed by a slight dip in June, then a steady rise through September. The index falls in October and November, but since then has steadily risen. This pattern largely fits with the economic stopping and starting since April 2020 due to the pandemic, recession, and recovery. Initially, businesses may have invested in their online presence, leading to a spike in May. Businesses started to open up in the summer, but some areas reinstated restrictions as coronavirus cases rose in the fall, which could explain the index's decline from September through November. With more consumers looking to make their holiday purchases online, we see a rise in the index since November to reflect seasonality. The even-weight index follows a similar pattern, though is more muted. The simple index shows steady growth through the sample period, a reflection that the number of GoDaddy online microbusinesses has been growing over time. This steady growth highlights that the simple index misses important patterns that are captured by the Baseline Index.

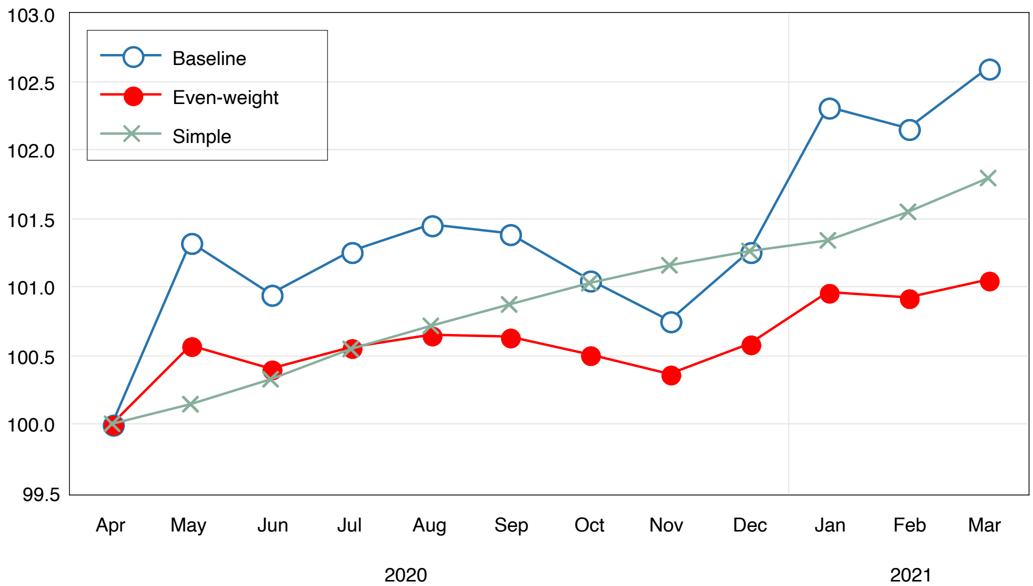
To get a more nuanced picture, we look at the receptivity, reception, and activity sub-indices. The patterns suggest that much of the variation in our composite index, at least over this historically unusual time period, is driven by the intensive margin of online microbusiness owners' and their customers' use of business websites. First worth noting is that the three components capture different information about online microbusinesses. These components are related, but not perfectly so. Table 7 shows the correlations across counties for November 2020, which vary in magnitude and sign, evidence that each component is indeed unique. Looking over time, Figure 11 shows reception, which captures the extensive margin – the growth rate and density of online microbusinesses. Reception smoothly rose in the summer, fell in the fall/early winter and then rose again. This pattern gives an indication of when businesses decided to take their operation online. Receptivity is based on variables that are only updated at an annual frequency, so it is flat through the sample period by construction. The variation in the activity index (which captures the intensive margin – how intensively and frequently business owners and their customers use the business's website) is what seems to be driving the main patterns we see in the composite index.

Table 7. Correlations between Composite and Sub-indices

	Composite	Receptivity	Reception
Receptivity	0.858		
Reception	0.367	0.391	
Activity	0.180	-0.343	-0.378

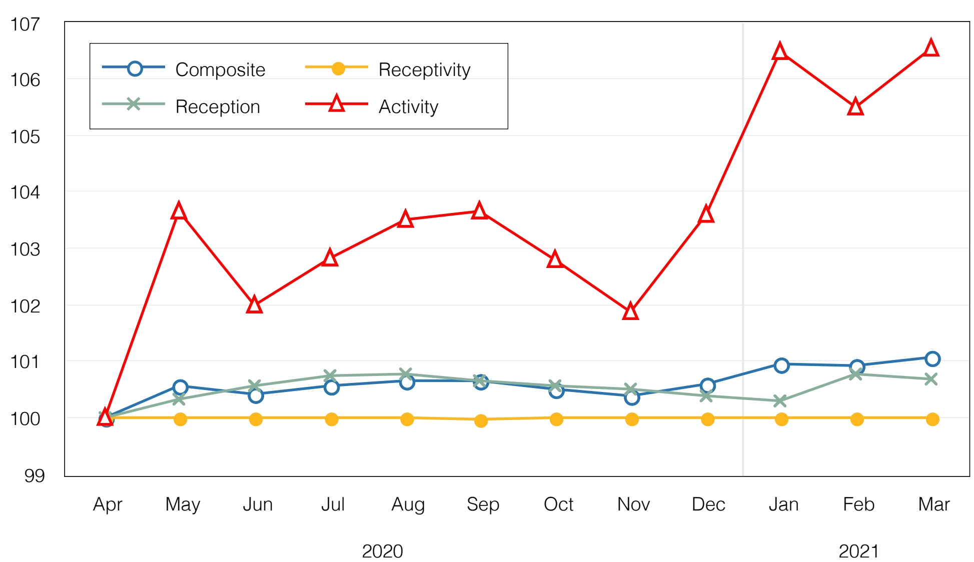
Sources: GoDaddy and UCLA Anderson Forecast

Figure 10. Microbusiness Indices Time Series, U.S.



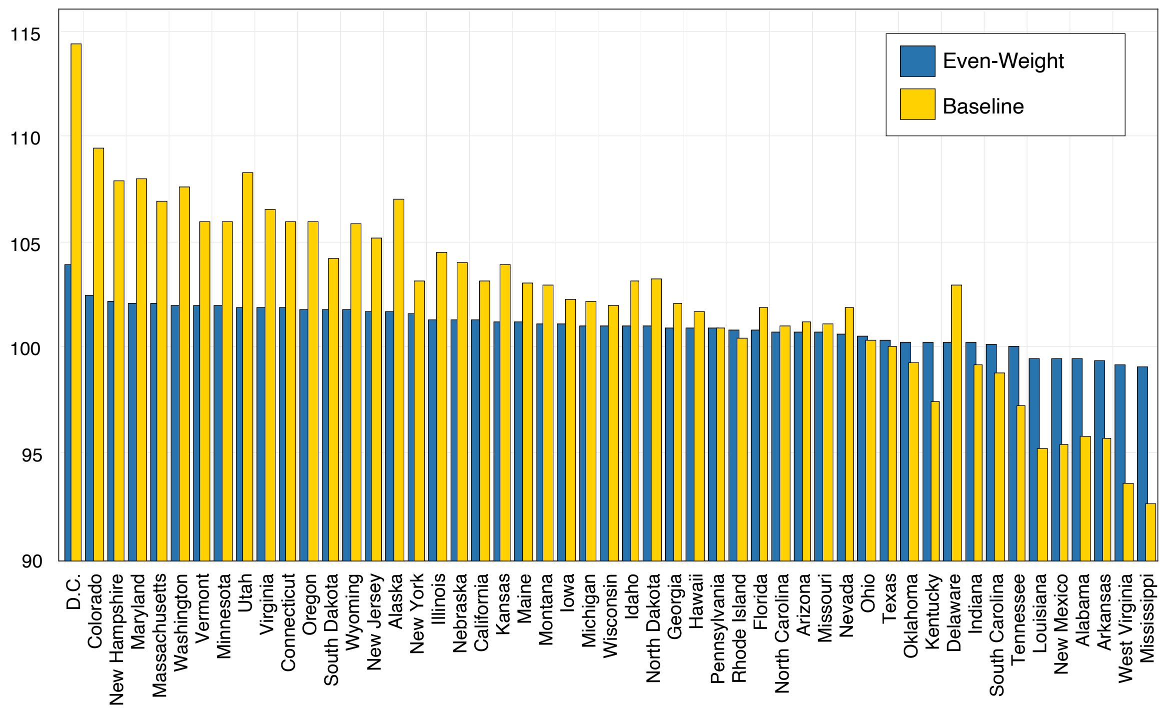
Sources: GoDaddy and UCLA Anderson Forecast

Figure 11. Microbusiness Sub-Indices Time Series, Even-Weight Index, U.S.



Sources: GoDaddy and UCLA Anderson Forecast

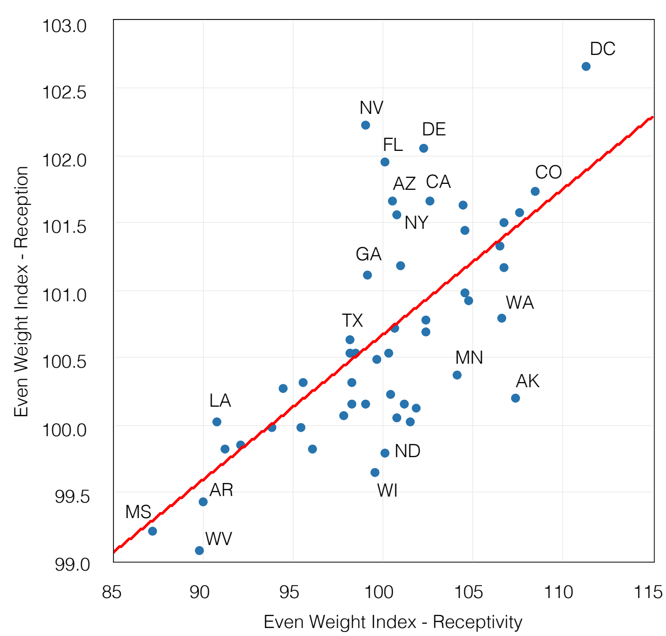
Figure 12. Microbusiness Index by State, Ranked by Even-Weight Index, March 2021



Sources: GoDaddy and UCLA Anderson Forecast

These time series hide regional variation. Looking at the Baseline and even-weight indices across states at a snapshot in time in March 2021, Figure 12 shows that D.C. has the highest index value and Mississippi has the lowest. Although the Baseline Index shows more variation across states than does the even-weight index, the ordering by and large is the same for both measures. Figure 13 shows how the sub-indices of receptivity and reception by state correlate and rank, where higher values of the receptivity sub-index indicate higher digital infrastructure and human capital, and higher values of the reception sub-index indicate more online businesses. We found a positive correlation (the red line) between receptivity and reception. The figure shows that D.C., a dense urban city and the national capital, has the highest index values while Mississippi and West Virginia have the lowest. States that are above the red line (average regression line), such as Florida, California, New York, Nevada, and Arizona, have better reception than the national average given their receptivity. On the other hand, those below the red line, such as Alaska, Minnesota, and Wisconsin, have relatively weaker reception given their receptivity. This suggests that while the states below the red line have the human capital and digital infrastructure capacity to support more microbusinesses, we do not see as many microbusinesses in those areas as we would expect based on the correlation between reception and receptivity across states. Explaining why we see this pattern is an area of future research.

Figure 13. Correlation Between Receptivity and Reception of Even-Weight Microbusiness Index by State, March 2021

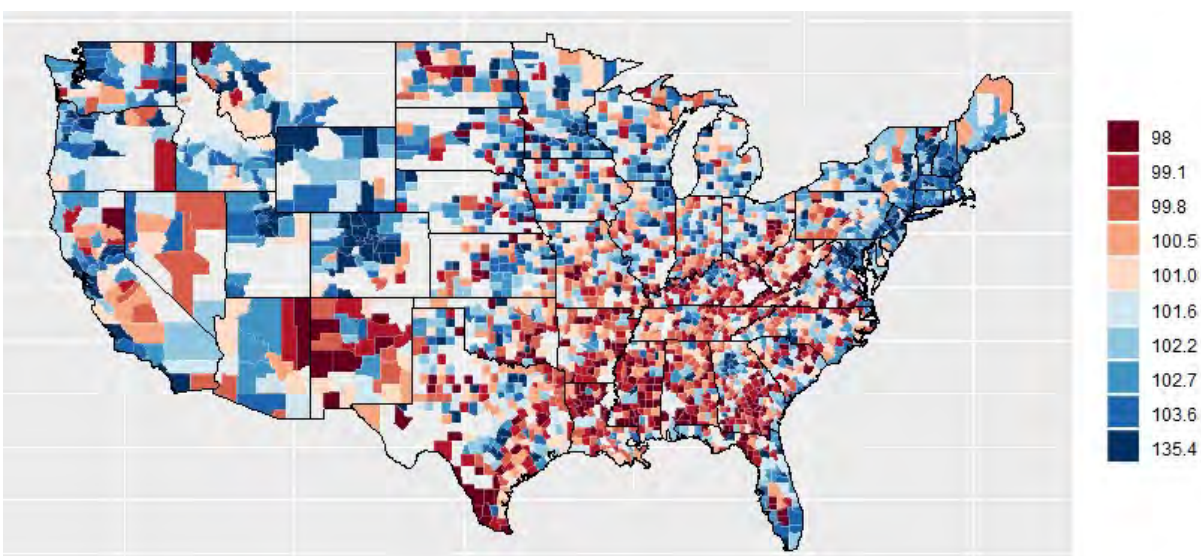


Sources: GoDaddy and UCLA Anderson Forecast

Exploring further detail, Figure 14 shows the variation across counties for March 2021. As suggested by the state rankings, there is spatial clustering – counties that have high or low values of the index tend to be adjacent to other counties that have high or low values of the index. Broadly, counties in the Southeast tend to have lower values of the index and counties in the Northeast, the West, and the mountain states tend to have higher values, though within states there is variation. For example, coastal Oregon has higher index values than does the eastern side of the state. Similarly, coastal California shows more strength in online microbusiness than does central California.

We can also use county-specific time series of the index to understand how the pandemic and recession that began in 2020 affected online microbusinesses. To illustrate the dynamics of the Baseline Index at the local level, we arbitrarily chose six major counties shown in Figure 15. First, San Francisco and New York had higher microbusiness index values than the other four counties as they are major tech hubs. While most counties followed a time series trend similar to the national average trend (Figure 10), we can tell the difference among them. As an example, San Francisco is widely reported to have been hit hard economically, and this can be seen by looking at the Baseline Index for San Francisco. From April 2020 to March 2021, San Francisco had an index value increase of one (126 to 127), lower than the national average increase of 2.5 (100 to 102.5), New York's three (119 to 122), Chicago's three (112 to 115), and Los Angeles's three (107 to 110).

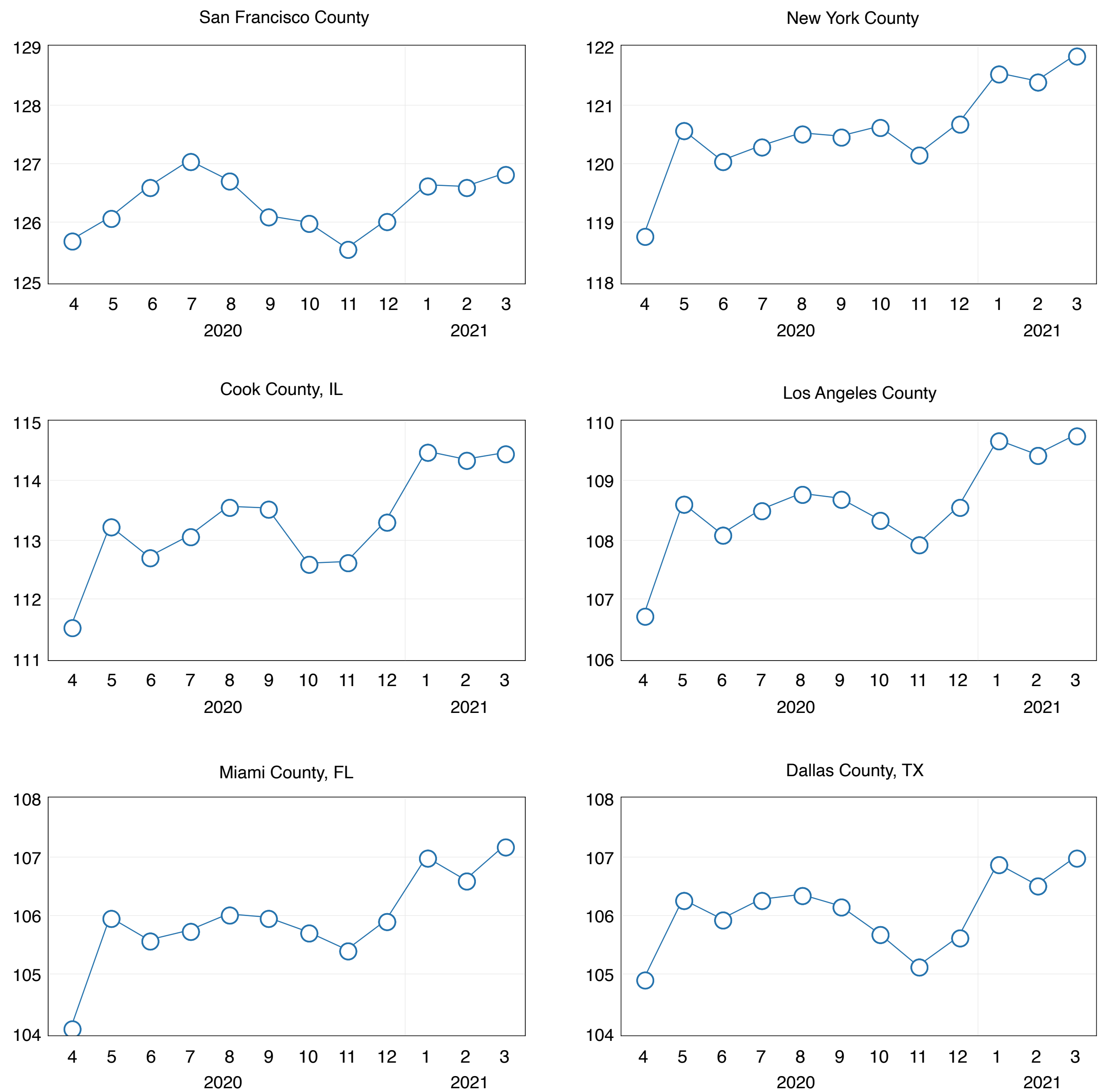
Figure 14. Microbusiness Index Across Counties, Even-Weight Index, March 2021



Note: Counties in grey have no available data for the index. Blue colors indicate higher values. Data are for March 2021.

Sources: GoDaddy and UCLA Anderson Forecast

Figure 15. Microbusiness Index, Baseline Index, Selected Counties



Sources: GoDaddy and UCLA Anderson Forecast

3. Understanding Microbusiness Success: Human Capital and Skills Training

Having formed a summary measure of online microbusinesses and a sense of the time series and cross-sectional patterns, we turned to the question of what local characteristics and policy factors help explain online microbusiness prevalence and growth. Mazzarol et al. (1999) summarize five main environmental factors that affect business formation:[14] (1) social—e.g. the impact of networks and the support of sociopolitical leaders with cultural acceptance, (2) economic—e.g. capital availability, (3) political—e.g. the support of public agencies, (4) infrastructure development—e.g. education system, incubator organizations, information accessibility, and (5) market emergence factors—e.g. niche emergence and technological innovation. In addition to the online platform and digital infrastructure mentioned in the previous sections, we focus in this section on human capital (both access to educational opportunities and educational attainment) and in the next section on access to capital. Both of these factors vary substantially across space so having a measure of online microbusinesses at detailed geographic levels can make use of this variation to estimate the relationship between human capital and microbusiness success. Figure 3 gives an example of the geographic variation in human capital.

A. Descriptive Statistics

Using simple correlations, we found a positive relationship between various measures of online microbusinesses and human capital. [15] To capture human capital, we used county-level data on the CHCI and also used the number of post-secondary institutions by county (total and the number of community colleges).[16]

Table 8. Correlations between Online Microbusinesses and Human Capital, September 2019

	Microbusiness density	Microbusiness owner density
# of postsecondary schools (all types)	0.108	0.065
# of community colleges	0.059	0.033
City Human Capital Index (CHCI)	0.200	0.143

Note: County level cross section from September 2019

Table 8 shows simple correlations from before the pandemic in September 2019. The table shows that counties that had higher levels of human capital and more postsecondary institutions also tended to have more microbusinesses and microbusiness owners. The magnitudes of these correlations are similar when using data about microbusinesses from after the pandemic hit (March 2021), but now we can add our microbusiness index (Table 9). We do not consider the receptivity index since this sub-index does not measure online microbusinesses. Again, we see positive correlations. For the composite index and the reception index, this is not surprising since both of these indices include variables that measure the number of microbusinesses and microbusiness owners per capita. The activity index, which captures how often and how intensely microbusiness websites are used and updated, is negatively correlated with measures of human capital: counties with more active microbusinesses tend to be those with fewer postsecondary institutions and lower levels of educational attainment. This could indicate that skills training and access to a skilled labor force is less important for online microbusiness activity, but is important for business formation. These correlations are generally statistically significant, as confirmed in the univariate regression results in Table 10.

Table 9. Correlations between Microbusiness Index and Human Capital, March 2021

	Microbusiness density	Microbusiness owner density	Composite index	Reception index	Activity index
# of postsecondary schools (all types)	0.116	0.060	0.254	0.265	-0.094
# of community colleges	0.061	0.029	0.170	0.179	-0.058
CHCI	0.202	0.129	0.837	0.461	-0.127

Note: County level cross section from March 2021

14. *See Fayolle and Gailly, 2015; Martin et al., 2013; Rupasingha and Wang, 2017.*
15. *In this and our other analysis, we removed outliers: we flagged geographic areas below the fifth percentile in number of microbusinesses and geographic areas where the number of microbusinesses more than doubled month to month or fell by more than half month to month. Any geographic area that was flagged as an outlier in any month was removed from the analysis. Analysis that did not remove outliers in this manner yielded similar results in terms of magnitudes, signs, and significance.*
16. *Postsecondary institution data are from the National Center for Education Statistics, Integrated Postsecondary Education Data System (https://nces.ed.gov/ipeds/use-the-data), which collects data about all postsecondary institutions that participate in federal financial aid programs. Community colleges are defined as public two-year institutions that offer an associate's degree as the highest degree offered. Data are a cross-section as of September 2019.*

Table 10. Skills Training and Microbusiness Index, Univariate Regressions

Panel A

Independent variables	Dependent variables				
	Microbusiness density	Microbusiness owner density	Composite index	Reception index	Activity index
	(1)	(2)	(3)	(4)	(5)
City Human Capital Index	0.334***	0.077***	0.838***	0.088***	-0.085***
	(0.031)	(0.012)	(0.011)	(0.003)	(0.013)
Constant	-38.627***	-8.229***	-12.442***	88.676***	117.318***
	(4.316)	(1.584)	(1.465)	(0.452)	(1.776)
Observations	2,670	2,670	2,646	2,670	2,646
Adjusted R^2	0.041	0.016	0.701	0.212	0.016

Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. The sample period is March 2021.

Panel B

Independent variables	Dependent variables				
	Microbusiness density	Microbusiness owner density	Composite index	Reception index	Activity index
	(1)	(2)	(3)	(4)	(5)
# postsecondary schools (all types)	0.270***	0.050***	0.358***	0.072***	-0.089***
	(0.045)	(0.016)	(0.027)	(0.005)	(0.018)
Constant	6.632***	2.285***	101.809***	100.591***	105.857***
	(0.346)	(0.126)	(0.205)	(0.039)	(0.141)
Observations	2,670	2,670	2,646	2,670	2,646
Adjusted R^2	0.013	0.003	0.064	0.070	0.008

Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. The sample period is March 2021.

Panel C

Independent variables	Dependent variables				
	Microbusiness density	Microbusiness owner density	Composite index	Reception index	Activity index
	(1)	(2)	(3)	(4)	(5)
# community colleges	1.329***	0.233	2.246***	0.451***	-0.510***
	(0.422)	(0.153)	(0.253)	(0.048)	(0.171)
Constant	6.861***	2.333***	101.971***	100.622***	105.803***
	(0.355)	(0.129)	(0.214)	(0.040)	(0.145)
Observations	2,670	2,670	2,646	2,670	2,646
Adjusted R^2	0.003	0.0005	0.029	0.032	0.003

Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. The sample period is March 2021.

B. Regression Results

We wanted to understand the extent to which these simple correlations are due to the microbusiness information captured by our index and to what extent they are due to other local characteristics. To address this, we ran a number of multivariate regression models that controlled for county characteristics. We included population because counties with more people will naturally have more businesses and more schools. We included a measure of labor market strength (the employment divided by the total population) because this may affect how favorable business conditions are. Moreover, the employment to total population ratio captures the proportion of the population that is working, as opposed to the total population, which includes minors and retirees. Since we have found that computer and broadband access matter for online microbusinesses, we added these controls (measured as the percent of households with computer or broadband access). Median income was included because this could reflect both owners' ability to self-fund a new business and also the ability of local customers to spend money on goods and services. We controlled separately for the CHCI and the number of postsecondary institutions because the CHCI captures educational outcomes (what skills people have) whereas the number of postsecondary institutions captures the availability of skills training. Finally, we controlled for state factors using state fixed effects to account for otherwise unobservable differences across states that could affect business taxes and the ease or difficulty of starting and running a business.

The results, summarized in the table below (Table 11) use different measures of online microbusinesses: the number of microbusinesses per 100 residents (density), the composite index, the reception index, and the activity index. Even after controlling for various demographic factors, human capital and access to skills training are positively related to online microbusinesses in most specifications, though the coefficients are not always statistically significant. For example, a county with one more postsecondary institution is expected to have 0.3 more microbusinesses per 100 residents and a composite microbusiness index that is larger by 0.05 index points. The results are qualitatively similar in terms of coefficients using per capita values for the number of postsecondary institutions and community colleges (Table 12).

Table 11. Skills Training and Microbusiness Index, Regression Results (in Levels)

	Dependent variables			
	Microbusiness density	Composite index	Reception index	Activity index
Independent variables	(1)	(2)	(3)	(4)
# of postsecondary schools (all types)	0.311**	0.049*	0.021	0.086
	(0.150)	(0.030)	(0.015)	(0.060)
# of community colleges	-0.484	0.226*	0.005	0.487*
	(0.686)	(0.135)	(0.069)	(0.273)
Employment to population ratio (%)	0.208***	0.056***	0.019**	0.047
	(0.074)	(0.015)	(0.007)	(0.029)
City Human Capital Index	0.186***	0.350***	0.046***	-0.079***
	(0.056)	(0.011)	(0.006)	(0.022)
Median household income	0.0001***	-0.00001	0.00002***	-0.00004**
	(0.00005)	(0.00001)	(0.00000)	(0.00002)
Population	-0.00000	-0.00000*	0.00000	-0.00000**
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
% of household with computer	0.112	0.947***	0.007	0.036
	(0.140)	(0.028)	(0.014)	(0.056)
% of household with broadband	-0.174	0.048**	0.020*	-0.025
	(0.108)	(0.021)	(0.011)	(0.043)
Constant	-31.760***	-29.945***	91.093***	116.037***
	(7.309)	(1.441)	(0.738)	(2.906)
Observations	2,667	2,643	2,667	2,643
Adjusted R^2	0.065	0.902	0.285	0.103

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. Sample period is in March 2021.*

Table 12. Skills Training and Microbusiness Index, Per Capita Variables

Independent variables	Dependent variables			
	Microbusiness density	Composite index	Reception index	Activity index
	(1)	(2)	(3)	(4)
# postsecondary schools (all types) per 100 residents	-27.824	100.125***	27.793*	149.926**
	(156.646)	(30.804)	(15.802)	(62.229)
# community colleges per 100 residents	-0.026	23.372	6.922	50.991
	(245.487)	(48.400)	(24.764)	(97.776)
Employment to population ratio (%)	0.219***	0.058***	0.020***	0.050*
	(0.073)	(0.015)	(0.007)	(0.029)
City Human Capital Index	0.215***	0.342***	0.044***	-0.089***
	(0.057)	(0.011)	(0.006)	(0.023)
Median household income	0.0001***	-0.00000	0.00002***	-0.00003*
	(0.00005)	(0.00001)	(0.00000)	(0.00002)
Population	0.00000**	0.00000	0.00000***	-0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
% of household with computer	0.110	0.951***	0.008	0.043
	(0.140)	(0.028)	(0.014)	(0.056)
% of household with broadband	-0.180*	0.044**	0.018*	-0.031
	(0.108)	(0.021)	(0.011)	(0.043)
Constant	-34.677***	-29.393***	91.164***	116.803***
	(7.365)	(1.447)	(0.743)	(2.924)
Observations	2,667	2,643	2,667	2,643
Adjusted R^2	0.064	0.902	0.286	0.104

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. The sample period is March 2021.*

To better understand whether there is scope for a causal relationship between human capital and access to skills training and microbusinesses, we looked at how online microbusinesses fared during the recovery from the pandemic, looking at changes from April 2020 to March 2021. The results, in Table 13 suggest that in counties with more pre-pandemic postsecondary institutions, community colleges, or higher educational attainment do not experience a larger increase in the microbusiness index value. When the dependent variable is the change in the number of microbusinesses, we do see that counties with more postsecondary institutions tended to experience a larger increase in the number of microbusinesses, but find that the relationship with community colleges and human capital is negative (and statistically significant for the human capital index). In sum, the results in this section generally indicate that areas with better access to skills training, education, and an educated workforce are also areas with more microbusinesses.

Table 13. Skills Training and Microbusiness Index, Change April 2020 - March 2021

	Dependent variables: Change from Apr 2020 to Mar 2021			
	Change of # of microbusiness	Change of composite index	Change of reception index	Change of activity index
Independent variables	(1)	(2)	(3)	(4)
# of postsecondary schools (all types)	440.556***	-0.002	0.005	-0.013
	(51.086)	(0.021)	(0.019)	(0.045)
# of community colleges	-359.305	0.031	0.035	0.117
	(233.806)	(0.095)	(0.085)	(0.206)
% point change of employment to population ratio	58.293	0.018	-0.036	0.102*
	(63.452)	(0.026)	(0.023)	(0.056)
City Human Capital Index	-45.314**	-0.003	0.001	-0.007
	(19.252)	(0.008)	(0.007)	(0.017)
Median household income	0.049***	0.00001	-0.00000	0.00002
	(0.015)	(0.00001)	(0.00001)	(0.00001)
Population	0.008***	0.00000	-0.00000	0.00000
	(0.001)	(0.00000)	(0.00000)	(0.00000)
% of household with computer	-29.835	0.032	-0.006	0.085**
	(47.570)	(0.020)	(0.017)	(0.042)
% of household with broadband	-56.669	-0.030**	0.011	-0.086***
	(36.946)	(0.015)	(0.013)	(0.033)
Constant	8,967.854***	2.407**	0.596	4.870**
	(2,492.721)	(1.018)	(0.904)	(2.201)

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level.*

4. Understanding Microbusiness Success: Access to Capital

Business knowledge and skills are necessary but may not be sufficient to start a new business or online business. Would-be business owners often need capital to start and expand their business. GoDaddy's own customer survey found that 25% of business owners ranked access to capital as the biggest challenge to starting up their business. Alekseev et al. (2020) found that, particularly during the pandemic, the "biggest challenge for 38.9% of businesses was accessing capital." Chodorow-Reich et al. (2020) found in a study of loans of one million dollars and above that small firms tended to have less favorable formal loan terms than did larger firms.

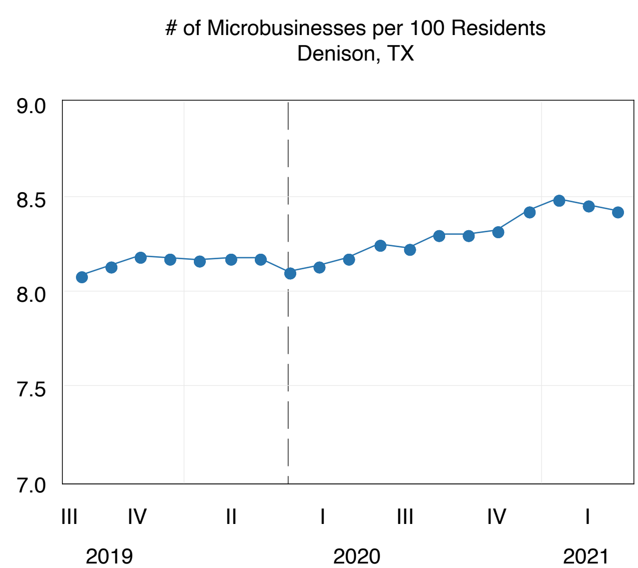
The online microbusinesses we study may not need large sums of money, but do nonetheless need some funds to begin operations. The aforementioned GoDaddy survey found that 72% of microbusiness owners needed less than $25,000 to start their business. In this section, we use two methods to explore how access to capital helped online businesses grow and thrive. The first section discusses two case studies from Denison, TX and Gilbert, AZ, two cities that partnered with GoDaddy to create support and funding programs for online microbusinesses. The second section takes a comprehensive look at how the presence of banks (including community banks) and loans from the Paycheck Protection Program (PPP) correlate with measures of online microbusinesses.

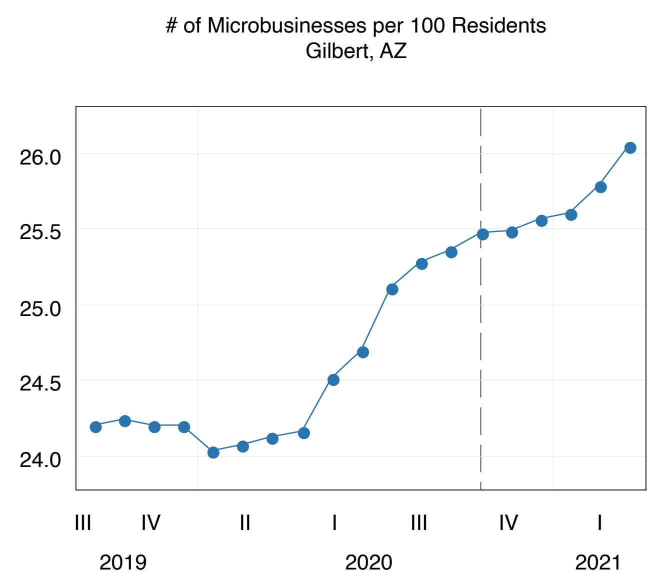
A. Case Studies: Denison, Texas and Gilbert, Arizona

Across the country, state and local governments created programs intended to help local businesses during the pandemic. Denison, Texas and Gilbert, Arizona made use of GoDaddy's expertise on online microbusinesses to create programs to help small local businesses use online platforms (from any source, not necessarily just from GoDaddy). Denison is a small city with about 25,000 residents north of Dallas on the border with Oklahoma. Beginning in April 2020 the city put aside $200,000 to make grants up to $6,000 to help local businesses create an online presence with marketing and web design assistance. Gilbert, Arizona, a city of over 270,000 people in the greater Phoenix area, had a larger program with a broader scope. Starting in October 2020, the program put aside $18 million in CARES Act funding to help small businesses, including online microbusinesses.[17]

Since these programs helped give small businesses access to capital, they offer a way for us to study how access to capital affects online microbusinesses. If access to capital is a binding constraint that prevents otherwise viable business ideas from moving forward, we would expect to see an increase in the number or density of microbusinesses after these programs began. Figure 16 shows the number of microbusinesses per 100 residents over time in Denison and in Gilbert.[18] The vertical dash line shows the start date of each city's program. The graphs show that after the programs began, microbusiness density increased; however, based on this graph alone, we cannot say how much

Figure 16. Density of Online Microbusinesses Over Time, Denison TX and Gilbert AZ





Sources: GoDaddy and UCLA Anderson Forecast

17. *These programs may have provided more than just access to capital. Technically, we cannot separately identify the effect of each program component with the data we have.*
18. *These figures were created by aggregating up from the zip codes that comprise Denison and Gilbert. In addition, because Denison's program began in April 2020, we could not conduct the analysis in this section with our index because April 2020 is the first month in which the index is available. There would be no 'pre-period' to serve as a comparison.*

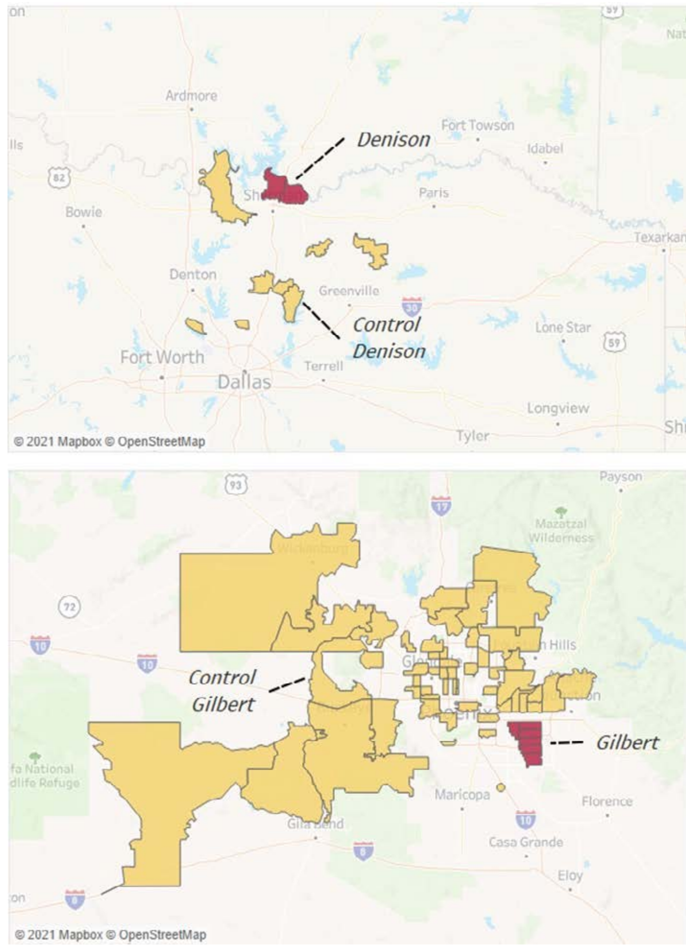
of the post-program patterns can be attributed to the aid provided by the two cities. The reason is that we do not know what would have happened in the absence of the programs – we do not know the counterfactual. An ideal way to answer this question is to have an experiment in which Denison and Gilbert were randomly chosen out of a set of cities to be given these small business programs. In such a case, we could use the outcomes in those 'untreated' cities to tell us about the counterfactual.

A method called 'synthetic controls' allowed us to approximate this situation (Abadie et al. 2010, 2011). We constructed a 'control Denison' and a 'control Gilbert' using a weighted average of a set of other areas (called the 'donor' areas). The synthetic control method selected the weights so that the 'control Denison' and the 'control Gilbert' had characteristics similar to those of the actual Denison and the actual Gilbert.[19]

Using the methods described in Abadie et al. (2010, 2011), we created a synthetic control Gilbert and a synthetic control Denison. Figure 17 shows real Gilbert and Denison (red color) and the zip codes that the algorithm selected to comprise the control for Denison and Gilbert (yellow color). Once we had these synthetic controls, we could simply compare the time series of microbusiness density for the real and the synthetic version of each city and see if there were more microbusinesses in the real cities than in the synthetic control cities after the programs began.

The result is in Figure 18. Focus first on the thick black lines, which plot the difference between online microbusiness density in the real Denison (or Gilbert) and its synthetic counterpart. Values above zero indicate that the real city had a relatively higher microbusiness density than the control version of the city. Starting with Denison, roughly before April 2020, microbusiness density in Denison and control Denison followed a similar trajectory (the difference between the two hovers around zero). This tells us that the synthetic control method was able to construct a con-

Figure 17. Map of Denison and Gilbert Zip Codes and Controls

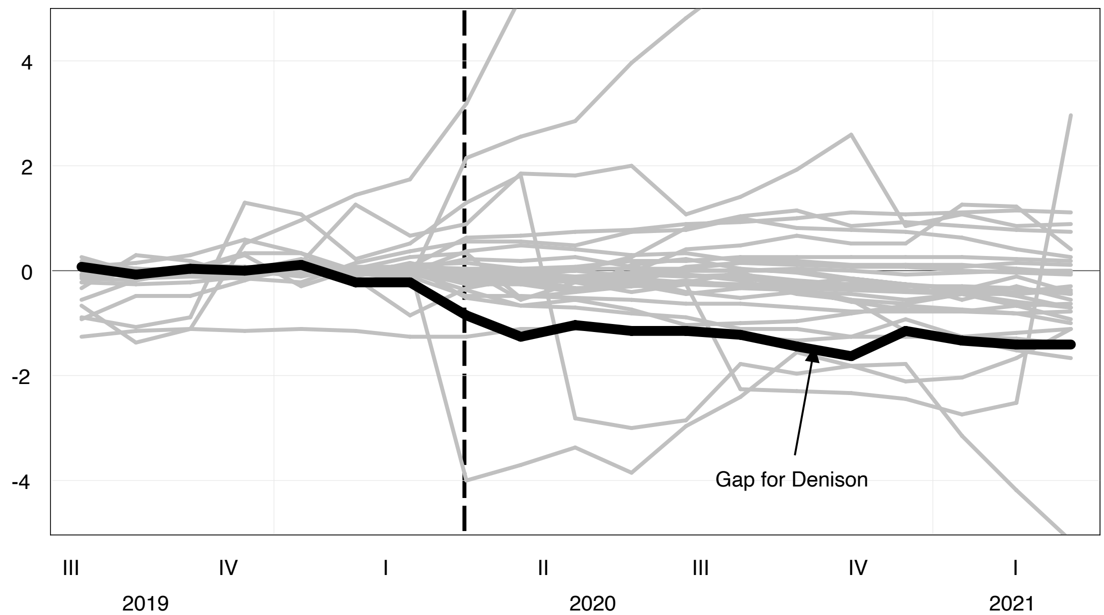


Source: UCLA Anderson Forecast

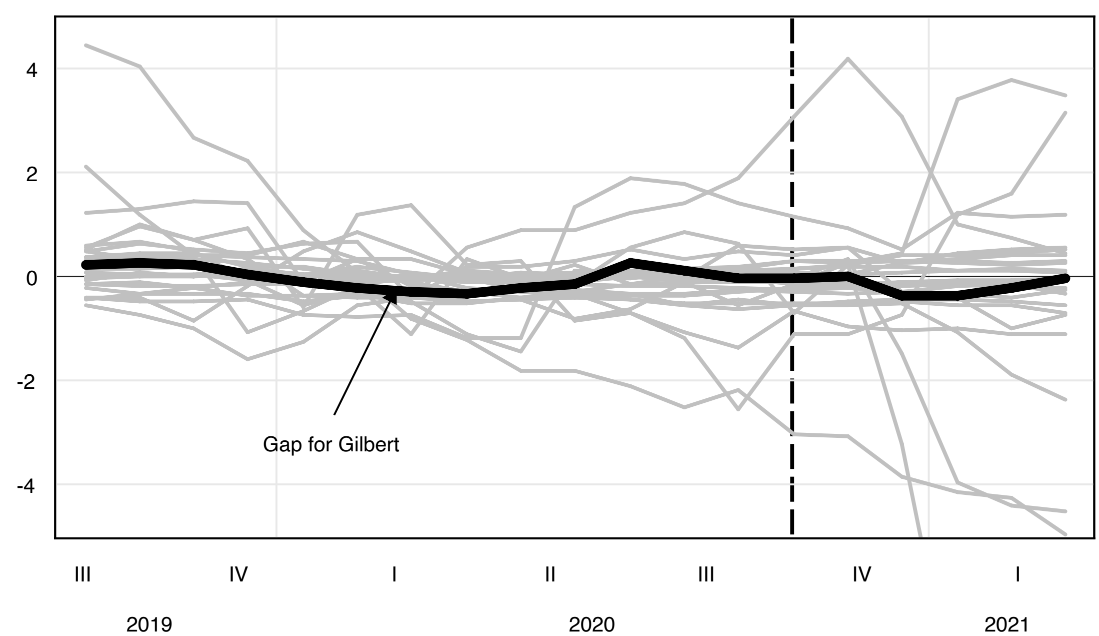
19. *The characteristics used to create the synthetic controls were taken from the time before the programs began. We used the average value of various demographic characteristics from the pre-treatment period (September 2019 – March 2020 for Denison and September 2019 – September 2020 for Gilbert) that included the fraction of households with a computer, the fraction of households with a broadband subscription, the City Human Capital Index, median household income, population density, the density of microbusinesses, and the density of microbusiness owners. In the synthetic control method, the researcher also has control over the pool of potential donor areas that make up the 'control Denison' and 'control Gilbert.' Since we were looking at two cities rather than two counties, our donor pool consisted of zip codes. For Denison, we allowed the donor pool to be zip codes in the counties that are adjacent to the county in which Denison is located. For Gilbert, we allowed the donor pool to be zip codes that are in Gilbert's same county (Maricopa County, which contains Phoenix). In both cases, we wanted the potential donors to be subject to the same state laws and policies and to be (relatively) geographically close by. We also wanted to minimize the chance that there would be spillovers of the policy, where businesses located in neighboring areas got assistance from the programs. In the case of Denison, we only let potential donor areas come from zip codes outside Denison's county. In the case of Gilbert, this was trickier because the adjacent counties are much more sparsely populated than is Maricopa county and may not be good candidates for a control. So, we restricted the donor zip codes to come from Maricopa county but did not allow donor zip codes to be those that directly abut Gilbert.*

Figure 18. Synthetic Control Results

Panel A: Denison, TX



Panel B: Gilbert, AZ



Sources: GoDaddy and UCLA Anderson Forecast

trol Denison that resembled the real Denison. Starting in April 2020 when Denison's program started, we see that microbusiness density in Denison fell relative to control Denison. In Gilbert, microbusiness density remained relatively flat after the program began relative to the 'control Gilbert.' This analysis does not indicate that there was a measurable benefit from the programs in Denison and Gilbert. To see this, consider the light grey lines in the figure. These lines tell us what range of results to expect just based on chance. Because the thick black lines lie within the range of the grey lines, the effects of these programs are likely to be indistinguishable from zero. To the left of the dashed lines, the grey lines are generally close to zero, but to the right they spread out substantially. This means there is enough noise in the data to prevent us from making strong conclusions about the effects of the Denison and Gilbert programs.[20]

There are a number of possible interpretations for this result. One is that the control areas themselves could have had their own small business assistance programs which could have encouraged online microbusiness growth in the control zip codes, therefore masking the benefits of the Denison and Gilbert programs. There also could have been anticipatory business formation, where owners took their businesses online or formed businesses in anticipation of the city assistance programs. In this case, we would see microbusiness growth before program implementation, not after. Another possibility is that these cities created support programs because local microbusinesses were already struggling more than were microbusinesses in surrounding areas. If this were the case, then we would not expect to see microbusiness growth after these programs started, relative to the control areas.

These two case studies are very specific cases and cover a very unique time (the pandemic). Since we cannot draw conclusive evidence from these two case studies, we employ another method to understand the relationship between microbusinesses and access to capital. In this method, we use regressions to look more broadly across all counties to study how the prevalence of bank branches (including community banks) and Paycheck Protection Program (PPP) loans explain variation in microbusinesses.

B. Access to Capital: Banks, Community Banks, and the Paycheck Protection Program

When a business owner needs capital to start a new business or to expand an existing one, the owner can turn to a number of sources. Some are more traditional and structured (such as banks) and some are more informal (such as personal savings, family, and online platforms such as Kickstarter). Though banks, and the loans they give such as PPP loans, are just a part of the story, by studying the relationship between access to these formal lenders and online microbusinesses, we can gain a better understanding of what types of access to capital contribute to online microbusiness success.

C. Descriptive Statistics

We gathered data on FDIC insured bank branch locations in the U.S. as of September 2020.[21] We classified each branch as being a 'community bank' or 'not a community bank' using the FDIC's 2019 classification of banking institutions. We then collapsed the data to give us a count of the number of community bank and non-community bank branches in each county in the U.S. We gathered data on the number of loans and on average loan size for loans up to $150K and loans above $150K by county from the PPP program.

In simple cross-sectional correlations from before the pandemic, we see that counties with more access to capital (measured by the number of banks and the number of community banks) had more online microbusinesses and microbusiness owners per 100 residents (Table 14). Interestingly, the fraction of banks that are community banks, which might be more inclined to or specialized in making small loans to small businesses, is negatively correlated with the number of online microbusinesses and microbusiness owners per 100 residents in a county.

These correlations are quite similar when using a cross-section in March 2021. Table 15 shows a correlation table from March 2021 that includes our index and selected sub-indices as well as variables that measure the amount of PPP support that businesses in each county received. Correlations are similar when

Table 14. Access to Capital and Online Microbusinesses, September 2019

Correlations between online microbusinesses and access to capital

	Microbusiness density	Microbusiness owner density
# of banks	0.125	0.077
# of community banks	0.093	0.058
Frac community banks	-0.076	-0.047

Note: County level cross section from September 2019.

20. *To assess this, we conducted an exercise where we: 1) took the pool of donor zip codes, 2) chose one at random, 3) created a synthetic version of that randomly chosen zip code from the remaining donor zip codes, and 4) saw how the density of online microbusinesses looked in that zip code relative to its synthetic control version. We then repeated this many times. This process created the rest of the lighter lines in Figure 18. If the difference between microbusiness density in Gilbert and Denison and their synthetic counterparts was within the range created by the process of repeated random sampling, then the effects of the programs in Denison and Gilbert are likely to be indistinguishable from zero. The figure shows that this is the case for both cities: the change in microbusiness density in the real Gilbert and Denison is within the range of what we might expect to see just by chance, though in the case of Denison, this change in microbusiness density is towards the low end of what one might expect by chance.*

21. *We restricted to branches with a branch opening date on or before February 1st, 2020. We made this restriction because we did not want to capture branches that opened in response to the pandemic, and due to the nature of the data, we could not observe branches that may have permanently closed due to the pandemic.*

Table 15. Access to Capital and Online Microbusinesses, March 2021

Correlations between online microbusinesses and access to capital

	Microbusiness density	Microbusiness owner density	Composite index	Reception index	Activity index
# of banks	0.133	0.071	0.317	0.311	-0.117
# of community banks	0.099	0.053	0.331	0.260	-0.051
Fraction of community banks	-0.079	-0.042	-0.328	-0.281	0.174
# of PPP loans up to $150K	0.108	0.058	0.221	0.250	-0.112
Total value PPP loans up to $150K	0.117	0.063	0.238	0.264	-0.121

Note: County level cross section from March 2021.

using all PPP loans and when restricting to only small PPP loans up to $150K. The correlations between our composite and reception index and the number of banks or community banks are positive, just like those using the density of microbusinesses. We again see a negative correlation with the fraction of banks that are community banks. There is a negative correlation between the activity index and the number of banks and community banks, and a positive correlation with the fraction of community banks. It could be that counties with fewer banks generally have smaller brick and mortar business districts so businesses have to compensate by being more active online. With respect to the PPP variables, we find that counties with more PPP loans up to $150K (by quantity or dollar value) also have more online microbusinesses and business owners per 100 residents. They also have higher values of the composite and reception indices and lower values of the activity index. These simple correlations can be confirmed using univariate regressions and the relationships are statistically significant (Appendix Table C1). Some of the counterintuitive results in these simple correlations could be driven by other local factors. We now turn to regression analyses that included controls to better estimate the relationship between access to capital and microbusinesses.

D. Regression Results

To account for other factors, we ran regression models that included a county's population, a measure of labor market strength, the percent of households with computer or broadband access, educational attainment using the CHCI, and state factors using state fixed effects. We included measures of both banks and PPP loans because we wanted to account for the correlation between these variables. We did so by including the total number of bank branches, the percent of branches that are a community bank, and the total number of PPP loans up to $150K in each county. PPP loans were processed by banks, so we would expect to, and do indeed, see a positive correlation between loans and banks.

Using the number of online microbusinesses per 100 residents, counties with more banks and a higher fraction of community banks do have more online microbusinesses: the coefficients are positive, but only marginally statistically significant for the number of banks (Table 16).[22] For example, a county with one more bank is expected to have 0.04 more microbusinesses per 100 residents and a county where the percent of community banks increases by one percentage point is expected to have 0.02 more microbusinesses per 100 residents. The sign on the PPP loan variable is positive as well, but is not statistically significant.[23] When the dependent variable is the composite or reception index, the coefficient on the total number of banks is positive and either significant (5% level using the reception index) or marginally significant (10% level using the composite index) and the coefficients on the number of loans and the fraction of community banks are not significant. For example, a county with one more bank is expected to have a reception index that is higher by 0.004. The lack of statistical significance for the PPP loan variable and the fraction of community banks could indicate that the hurdles to access loans and institutions intended for small businesses are too high for the microbusinesses we study.

With the activity index as the dependent variable, we do not see any statistically significant relationships on the bank and PPP variables. These results are suggestive that access to capital has a stronger tie to the extensive margin (the number of online microbusinesses, as measured by the composite index, the reception index, and the number of microbusinesses) than to the intensive margin (how active the microbusinesses are, as measured by the activity index). Also, because we do not see a significant coefficient on the PPP variable in any of the models, this suggests that perhaps the PPP loan program may not have been widely able to reach the online microbusinesses we study (by design or in implementation). Additional results in Alekseev et al. (2020) support this interpretation: only 25% of businesses they surveyed on Facebook reported having "access to formal sources of financing through a loan or line of credit from a finan-

22. *We also ran regressions using per capita variables for the number of banks and PPP loans. The coefficients are generally not statistically significant (Table 17).*

23. *One possible concern is that there is multi-collinearity between the PPP and the bank variables. The coefficients on these variables are significant in univariate regressions (Appendix Table C1) but often not in the multivariable regressions. We ran versions of the multivariable regressions in Tables 16 – 18 with either the bank or the PPP variables, but not both. The coefficients on these alternative specifications are similar in magnitude, sign, and significance to those in the regressions in Tables 16 – 18.*

Table 16. Access to Capital and Microbusiness Index, Regression Results in Levels

	Dependent variables			
Independent variables	Microbusiness density (1)	Composite index (2)	Reception index (3)	Activity index (4)
# of banks	0.037*	0.007*	0.004**	0.011
	(0.020)	(0.004)	(0.002)	(0.008)
% community banks	0.018	-0.003	-0.002	0.004
	(0.015)	(0.003)	(0.002)	(0.006)
# of PPP loans up to $150K (in thousands)	0.044	-0.008	0.002	-0.023
	(0.064)	(0.013)	(0.006)	(0.025)
City Human Capital Index	0.203***	0.349***	0.043***	-0.073***
	(0.057)	(0.011)	(0.006)	(0.023)
Median household income	0.0001***	-0.00001	0.00002***	-0.00005***
	(0.00004)	(0.00001)	(0.00000)	(0.00002)
Population	-0.00001	-0.00000	-0.00000	-0.00000
	(0.00000)	(0.00000)	(0.00000)	(0.00000)
% of households with computer	0.229	0.952***	0.011	0.050
	(0.142)	(0.028)	(0.014)	(0.056)
% of households with broadband	-0.223**	0.041*	0.017	-0.036
	(0.109)	(0.021)	(0.011)	(0.043)
Employment to population ratio (%)	0.204***	0.060***	0.022***	0.045
	(0.074)	(0.015)	(0.007)	(0.030)
Constant	-39.852***	-29.457***	91.373***	115.172***
	(7.835)	(1.536)	(0.789)	(3.105)
Observations	2,654	2,630	2,654	2,630
Adjusted R^2	0.068	0.902	0.290	0.101

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. Sample period is in March 2021.*

Table 17. Access to Capital and Microbusiness Index, Per Capita Variables

	Dependent variables			
Independent variables	Microbusiness density (1)	Composite index (2)	Reception index (3)	Activity index (4)
# banks per 100 residents	-8.570	0.876	-4.824**	14.845
	(22.486)	(4.495)	(2.274)	(9.082)
% community banks	0.014	-0.003	-0.003	0.002
	(0.016)	(0.003)	(0.002)	(0.006)
# PPP loans up to 150k per 100 residents	-0.069	-0.006	-0.026	-0.057
	(0.173)	(0.035)	(0.017)	(0.071)
City Human Capital Index	0.247***	0.354***	0.052***	-0.072***
	(0.056)	(0.011)	(0.006)	(0.022)
Median household income	0.0001***	-0.00001	0.00002***	-0.00004**
	(0.00004)	(0.00001)	(0.00000)	(0.00002)
Population	0.193	0.948***	0.002	0.051
	(0.143)	(0.028)	(0.014)	(0.057)
% of households with computer	-0.216**	0.042*	0.019*	-0.037
	(0.109)	(0.021)	(0.011)	(0.043)
% of households with broadband	0.233***	0.061***	0.031***	0.038
	(0.076)	(0.015)	(0.008)	(0.030)
Employment to population ratio (%)	-43.459***	-29.949***	90.753***	114.957***
	(7.746)	(1.517)	(0.783)	(3.065)
Observations	2,654	2,630	2,654	2,630
Adjusted R^2	0.065	0.901	0.283	0.101

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. The sample period is March 2021.*

cial institution." One implication for policy is that 'regular' small business loan programs might be targeting businesses that are larger than the microbusinesses we study and thus that programs targeted to very small businesses might provide a better fit for microbusiness owners. Fazio et al. (2021) find spikes in new business formation after the passage of the CARES Act and the subsequent Relief Supplemental Appropriations Act of 2021, suggesting that generally easing liquidity constraints may help business formation.

We also tried these same models estimated with the dependent variables as changes from early in the pandemic (April 2020) until March 2021. If we see that counties with more banks, a higher percentage of community banks, and more PPP loans had a larger increase in the number of online microbusinesses or the microbusiness index, that would be stronger evidence that these banking and loan factors have a causal effect on online microbusinesses. These results are in Table 18. Looking at changes in the number of online microbusinesses, there are positive and significant coefficients on the number of banks and the percent of community banks and no statistical significance for the PPP loan variable. That said, counties with one more bank or where the percent of community banks is one percentage point higher (pre-pandemic) tended to experience a larger increase in the number of online microbusinesses by 73 or 19 businesses respectively. With the composite index, we do not see any statistically significant coefficients; however, looking at the reception index, we see that counties with a higher percentage of community banks have more growth in the reception index. When the percentage of community banks increases by one percentage point, the reception index grew by an extra 0.005 index points. This suggests that if or when online microbusinesses get formal loans, community banks may play a role. The results show that the activity index has a negative relationship with the percentage of community banks, which could reflect that microbusinesses that have less access to capital make up for the deficit in financing with more online activity. Perhaps the lack of access to capital makes it harder for them to have a brick and mortar store, so they must rely on their online store.

Table 18. Access to Capital and Microbusiness Index, Regression Results in Changes (April 2020 - March 2021)

	Dependent variables: Change from Apr 2020 to Mar 2021			
	Change of # of microbusiness	Change of composite index	Change of reception index	Change of activity index
Independent variables	(1)	(2)	(3)	(4)
# of banks	73.456***	0.0001	0.0004	-0.001
	(6.911)	(0.003)	(0.003)	(0.006)
% community banks	18.977***	-0.001	0.005***	-0.009**
	(5.137)	(0.002)	(0.002)	(0.005)
# of PPP loans up to $150K (in thousands)	-27.732	-0.006	-0.006	-0.008
	(21.737)	(0.009)	(0.008)	(0.019)
City Human Capital Index	-52.367***	-0.003	0.004	-0.010
	(19.407)	(0.008)	(0.007)	(0.017)
Median household income	0.009	0.00001	-0.00000	0.00002
	(0.014)	(0.00001)	(0.00001)	(0.00001)
Population	0.003**	0.00000	0.00000	0.00000
	(0.001)	(0.00000)	(0.00000)	(0.00000)
% of households with computer	20.234	0.032	0.001	0.076*
	(47.996)	(0.020)	(0.017)	(0.043)
% of households with broadband	-68.306*	-0.035**	0.008	-0.093***
	(36.931)	(0.015)	(0.013)	(0.033)
% point change employment to population ratio	55.415	0.026	-0.028	0.110*
	(63.483)	(0.026)	(0.023)	(0.056)
Constant	6,587.621**	2.684**	-0.544	6.873***
	(2,660.383)	(1.092)	(0.968)	(2.364)
Observations	2,654	2,610	2,654	2,610
Adjusted R^2	0.499	0.008	-0.004	0.011

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level.*

Access to capital in some form is a prerequisite for business formation. The results in this section provide some evidence that formal lending institutions help online microbusinesses, but also suggest that these online microbusinesses need to access capital in a different way or in smaller quantities than other businesses to get their businesses going. Given the size of these businesses and the fact that many are sole proprietorships, we suggest that the difficulty, time cost, and complication of obtaining formal loans and learning about loan programs may be a hurdle to online microbusiness owners and potential owners.

5. Using Microbusinesses to Understand the Macroeconomy

We have formed an index and used its geographic variation to understand what local factors may contribute to microbusiness success. The next question is, once online microbusinesses are formed, is there evidence that they contribute to local economic activity at a level that is detectable in standard government data? In this section, we explore what the index can tell us about the relationship between microbusinesses and the economy.

First, we investigate if there is a correlation between GoDaddy's online microbusinesses and the local economy. We ran a panel regression using data from June 2018 until March 2021 of county unemployment rates on a set of variables that control for specific time and state characteristics, the density of microbusinesses, and COVID-19 new cases and deaths per capita (Table 19 Equation (1)). We found that the density of microbusinesses was significantly negatively correlated with the unemployment rate. In other words, a county with a higher concentration of online microbusinesses tends to also have a lower unemployment rate.

In Table 19 Equation (2), we changed the dependent variable to another measure of labor market strength: the employment to population ratio. We again see evidence that there is a significantly positive association between online microbusiness density and the employment rate. Note that these two regressions neither prove nor disprove that there is a causal relationship from online microbusinesses to local economic activity. In Equation (3), we investigated causality with a dynamic relationship. The dependent variable is the change in a county's employment between two time periods and the explanatory variable of interest is the change in the number of microbusinesses. In order to control for the variation in county size, we added county population and replaced COVID-19 new cases and deaths per capita with the simple count of new cases and deaths in each county. We also controlled for trends and persistence in employment by including a lag of the employment change. We found a significant and positive relationship between changes in the number of online microbusinesses and changes in local employment.

We view this as stronger evidence that online microbusinesses contributed to employment growth and strengthened local labor markets. These results make sense given what we know about GoDaddy's online microbusinesses from the survey conducted by GoDaddy to learn about their customers. About 25% of respondents indicated that their microbusiness is their main source of income and about the same fraction works 41 or more hours per week on their microbusiness. In addition, almost 50% say their microbusiness is their main source of employment.

We repeated the analysis but added the even-weight index[24] as an explanatory variable and limited the sample period to April 2020 – March 2021, the time period for which we can calculate the index. The reason is that we wanted to see whether the index captures variation in labor market outcomes, such as the unemployment rate, that is not explained by microbusiness density alone. We can get a sense that microbusiness density and the even-weight index capture different information by comparing Figure 14, which shows the variation in the composite index across counties in March 2021, and Figure 4, which shows microbusiness density. This comparison indicates that the even-weight index captures information about online microbusinesses that is not fully reflected in the number of microbusinesses per capita. In both measures, the coasts tend to both have higher values of the index and higher microbusiness density, but the Midwest and mountain states have higher index values despite having relatively lower reception (number of microbusinesses per capita).

In our more formal regression tests, if the even-weight index is useful for explaining the variation in local labor markets above and beyond what we can learn from microbusiness density (simple index), we would expect to see that the coefficient on the index is statistically significant. We find that this is the case for Equations (4) and (5), but is not the case for Equation (6). When the dependent variable is the unemployment rate or the employment to population ratio, the coefficient on the even-weight index is statistically significant and of the predicted sign: counties with a higher index value tend to have lower unemployment rates and higher employment to population ratios even after controlling for microbusiness density. The coefficients on the even-weight index variable are also larger in magnitude than are those on the density variable. When the dependent variable is the change in employment, we do not find that changes in the index help explain changes in employment above and beyond the change in the number of microbusinesses. One possible explanation is that the even-weight index better explains cross-sectional patterns than time series patterns. Still, these results generally support the idea that the microbusiness density variable (part of the reception component of the composite index) misses some information about online microbusinesses and highlights the importance of including other facets of online microbusinesses in the even-weight index, such as the activity component.[25]

24. *We use the even-weight index instead of the Baseline Index because the latter is a construct to maximize the correlation with the local economic activity.*

25. *Models that included the three sub-indices of receptivity, reception, and activity in place of microbusiness density and the composite index are qualitatively similar, showing that the activity and receptivity sub-indices have additional explanatory power over the reception index. These results are available from the authors upon request.*

Table 19. Microbusiness Indices, Microbusiness Density, and Labor Market Variables (County Level, Panel with Time and State Fixed Effects)

	Dependent variables					
	Unemployment rate (%)	Employment to population ratio (%)	Employment difference over two period	Unemployment rate (%)	Employment to population ratio (%)	Employment difference over two period
Independent variables	(1)	(2)	(3)	(4)	(5)	(6)
Density of microbusiness	-0.038***	0.459***		0.007***	-0.029***	
	(0.002)	(0.006)		(0.001)	(0.002)	
Change of microbusiness			0.000***			0.240***
			(0.000)			(0.032)
Monthly new Covid-19	-13.135***	-1.589	0.000***	-3.319	2.403	-0.228***
case rate	(2.068)	(6.320)	(0.000)	(2.676)	(5.242)	(0.009)
Monthly new Covid19	832.48***	-1256***	-0.004***	833.58***	-1292***	7.151***
death rate	(61.83)	(189)	(0.001)	(81.07)	(159)	(0.454)
Population			0.000***			0.006***
			(0.000)			(0.000)
Lag of dependent variable			0.109***			-0.005
			(0.004)			(0.002)
Even-weight index				-0.119***	0.914***	
				(0.007)	(0.014)	
Change of even-weight index						-6.116 (21.51)
Constant	3.644***	39.841	0.272	24.451***	-52.36***	583***
	(0.063)	(0.192)	(0.195)	(0.712)	(1.395)	(166)
Month fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
State fixed effect	Yes	Yes	Yes	Yes	Yes	Yes
Sample period	June 2018 to March 2021			April 2020 to March 2021		
Observations	67,622	67,597	67,592	31,760	31,748	28,993
Adjusted R^2	0.639	0.37	0.092	0.605	0.482	0.242

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level.*

6. Conclusions

One of the many trends accelerated by the COVID-19 pandemic was the shift to e-commerce. Businesses moved their operations online or expanded their online footprint to survive. Perhaps as a result of the recession that accompanied the pandemic, 2020 saw a notable rise in the number of new business formations and a growing interest in research related to small businesses. Given the importance of conducting business online and based on our research, we found that understanding trends not just in small businesses but specifically in online microbusinesses is a valuable input to descriptors of local economic conditions. The microbusiness index that we developed and introduced in this paper is an important step. Using this index, we show that access to capital and a skilled workforce are associated with microbusiness success and that there is a link between microbusinesses and local economic outcomes. We view the work in this report as first steps in a larger research agenda on online microbusinesses. First, we plan to continue to update our microbusiness index as new data become available. Creating a longer time series of index data will yield a more robust index that can account for seasonal patterns as well as longer, slow moving trends as we continue to refine and test the index. Second, this research has established relationships between selected inputs to business success (human capital and funding) by demonstrating that our microbusiness index can be used to examine such relationships. This paves the way to use the index to answer other policy questions about what leads to microbusiness success at a local level, such as whether affordable health care options for small business owners encourages entrepreneurship. Another avenue is to expand on our existing work by testing whether the relationships we find between local factors and microbusinesses is causal. We encourage other researchers to use our index in their own work, and to that end we hope to release the index at various geographic levels at regular intervals.

REFERENCES

Abadie, A., Diamond, A., & Hainmueller, J. (2010). “Synthetic Control Methods for Comparative Case studies: Estimating the Effect of California’s Tobacco Control Program.” Journal of the American statistical Association, 105(490), pp. 493-505.

Abadie, A., Diamond, A., & Hainmueller, J. (2011). “Synth: An R Package for Synthetic Control Methods in Comparative Case Studies.” Journal of Statistical Software, 42(13).

Alekseev, Georgij, Safaa Amer, Manasa Gopal, Theresa Kuchler, JW Schneider, Johannes Stroebel, and Nils C. Wernerfelt. (2020). "The Effects of COVID-19 on U.S. Small Businesses: Evidence from Owners, Managers, and Employees." NBER Working Paper No. 27833.

Barrlett, Robert P., and Adair Morse. (2020). "Small Business Survival Capabilities and Policy Effectiveness: Evidence from Oakland." NBER Working Paper No. 27629.

Bartik, W. Alexander, Marianne Bertrand, Zoe Cullen, Edward L. Glaeser, Michael Luca, and Christopher T. Stanton. (2020). "How Are Small Businesses Adjusting to COVID-19? Early Evidence From A Survey." NBER Working Paper No. 26989.

Chodorow-Reich, Gabriel, Olivier Darmouni, Stephan Luck, and Matthew C. Plosser. (2020) “Bank Liquidity Provision Across the Firm Size Distribution.” NBER Working Paper No. 27945.

Chetty, Raj., John N. Friedman, Nathaniel Hendren, Michael Stepner. (2020). "The Economic Impacts of COVID-19: Evidence from a New Public Database Built Using Private Sector Data." Opportunity Insights Working Paper.

Crouzet, Nicolas, and Neil R. Mehrotra. 2020. "Small and Large Firms over the Business Cycle." American Economic Review, 110 (11): pp. 3549-3601

Fayolle, Alain, and Benoit Gailly. (2015). "The Impact of Entrepreneurship Education on Entrepreneurial Attitudes and Intention: Hysteresis and Persistence." Journal of Small Business Management. 53, pp. 75-93.

Fazio, C. E., Guzman, J., Liu, Y., & Stern, S. (2021). “How is COVID Changing the Geography of Entrepreneurship? Evidence from the Startup Cartography Project.” NBER Working Paper No. 28787.

Kim, Olivia, Jonathan A. Parker, and Antoinette Schoar. (2020). "Revenue Collapses and the Consumption of Small Business Owners in the Early Stages of the COVID-19 Pandemic." NBER Working Paper No. 28151.

Marin, Bruce C., Jeffrey J. McNally, and Michael J. Kay. (2013). "Examining the Formation of Human Capital in Entrepreneurship: A Meta-Analysis of Entrepreneurship Education Outcomes." Journal of Business Venturing. 28, pp. 211-224.

Mazzarol, Tim., Thierry Volery, Noelle Doss, and Vicki Thein. (1999). "Factors Influencing Small Business Start-ups: A Comparison with Previous Research." Journal of Business Venturing. 28, pp. 211-224.

Mossberger, Karen and Caroline J. Tolbert. (2021) “Digital Citizenship and Digital Communities: How Technology Matters for Individuals and Communities.” International Journal of E-Planning Research 10 (3), pp. 19-34.

Mossberger, Karen, Caroline J. Tolbert, and Scott LaCombe (2021). “Choosing the Future: Technology and Opportunity in Communities.” Oxford University Press.

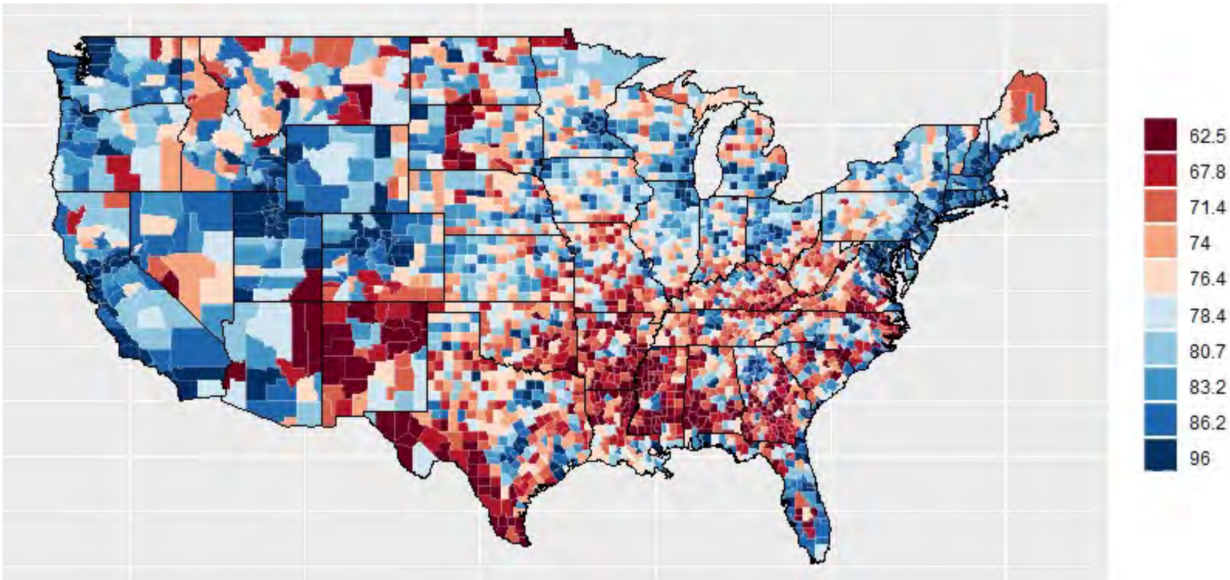
Mossberger, Karen, Caroline J. Tolbert, and Scott LaCombe. (2020) “A New Measure of Digital Participation and Its Impact on Economic Opportunity.” (https://img1.wsimg.com/cdn/Other/All/FOS-Intl/1/en-US/db5fd8b8-1e0f-4eaf-9ba2-da0dd19a5ea9/White_Paper_VF-final-website.pdf)

Rupasingha, Anil, and Kyungsoon Wang. (2017). "Access to Capital and Small Business Growth: Evidence from CRA Loans Data." Annals of Regional Science. 59, pp. 15-41.

Yu, William, and Leila Bengali. (2021). "Digital Infrastructure, the Economy and Online Microbusiness: Evidence from GoDaddy’s Microbusiness Data." UCLA Anderson Forecast Quarterly Report. June. pp. 81-99.

Appendix A:

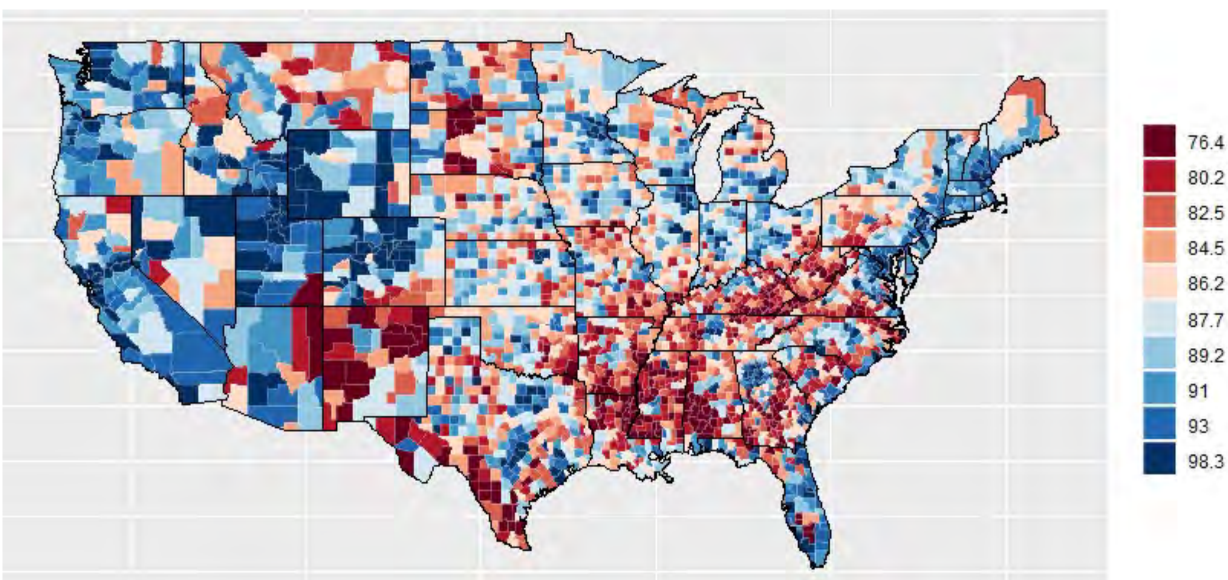
Figure A1. Fraction of Households with A Broadband Internet Subscription (%)



Note: Blue colors indicate higher percentages.

Source: 2019 Five-year American Community Survey

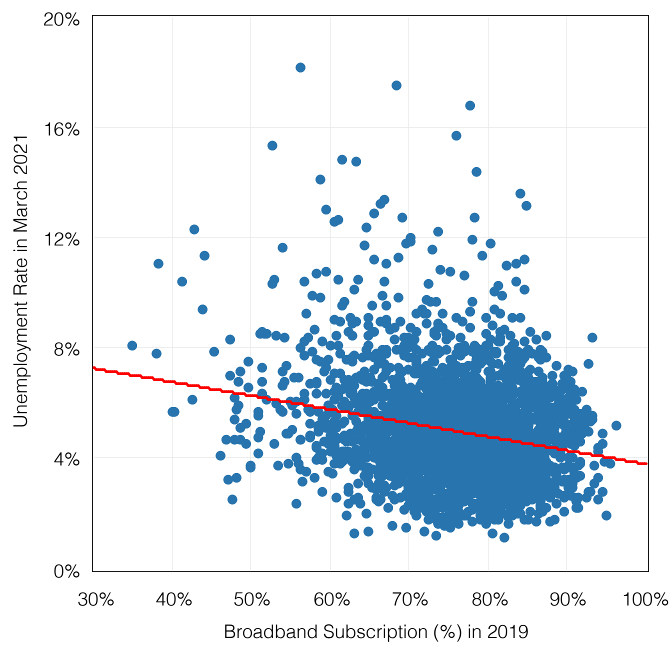
Figure A2. Fraction of Households with A Computer (%)



Note: Blue colors indicate higher percentages.

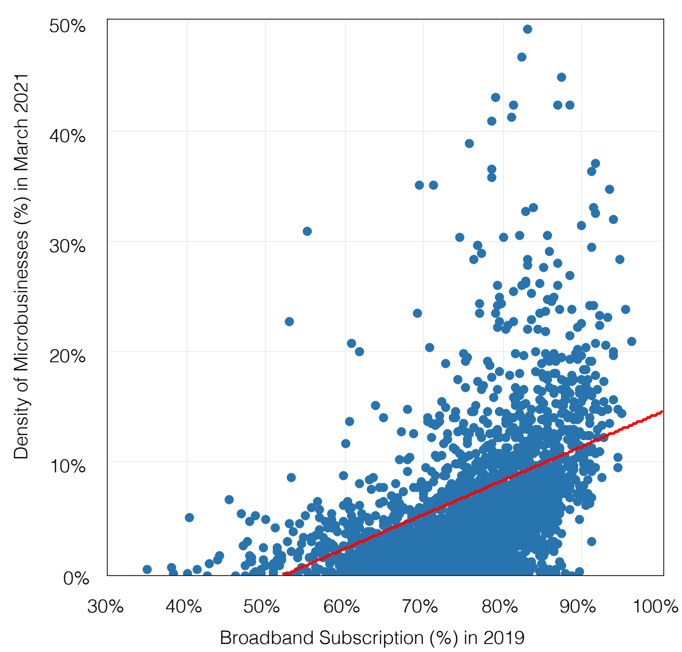
Source: 2019 Five-year American Community Survey

Figure A3. Correlation Between Broadband Subscriptions and the Unemployment Rate in March 2021 by County



Sources: 2019 Five-year American Community Survey and Bureau of Labor Statistics

Figure A4. Correlation Between Broadband Subscriptions and the Density of Online Microbusinesses by County



Sources: GoDaddy, UCLA Anderson Forecast, and the 2019 Five-year American Community Survey

Table A1. Multivariate Regressions

	Dependent variable: Unemployment rate (%) - March 2021								
Explanatory variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Broadband	-0.049		-0.023	-0.012	-0.024	-0.023	-0.016		-0.007
	(0.004)		(0.007)	(0.005)	(0.005)	(0.005)	(0.005)		(0.003)
Computer		-0.068						0.011	
		(0.006)						(0.006)	
Human capital index			-0.018	-0.000	-0.001	-0.010	-0.009	-0.015	0.002
			(0.005)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.003)
Median Income			-0.009	0.0181	0.02	0.015	0.015	0.009	-0.008
($ Thousands)			(0.004)	(0.000)	(0.000)	(0.000)	(0.000)	(0.003)	(0.002)
Unemployment rate (%)				0.931	0.822	0.808	0.804	0.819	0.762
- Feb 2020				(0.021)	(0.022)	(0.021)	(0.021)	(0.021)	(0.017)
Unemployment rate (%)					0.066	0.057	0.057	0.053	0.102
- Apr 2020					(0.005)	(0.005)	(0.005)	(0.005)	(0.004)
Population						0.843	0.834	0.809	0.508
(Million)						(0.081)	(0.080)	(0.080)	(0.051)
Population density						0.000	0.000	0.000	0.000
						(0.000)	(0.000)	(0.000)	(0.000)
Cumulative Covid-19							-0.052	-0.056	-0.021
case rate							(0.010)	(0.010)	(0.008)
Cumulative Covid19							0.000	0.000	0.000
death rate							(0.000)	(0.000)	(0.000)
Constant	8.744	10.831	9.703	1.408	1.848	3.332	2.673	1.699	0.098
	(0.306)	(0.486)	(0.561)	(0.469)	(0.458)	(0.446)	(0.499)	(0.588)	(0.349)
State fixed effect	No	No	No	No	No	No	No	No	Yes
Observations	2865	2865	2865	2865	2865	2864	2865	2864	2864
Adjusted R2	0.05	0.05	0.058	0.446	0.474	0.521	0.533	0.532	0.826

Notes: In parentheses are standard errors.

Appendix B: Detailed Index Creation Methods

From the coefficients obtained from the regressions summarized in Table 2, we used the coefficients, β_1, from each regression to form several sets of weights. In one set, we used all the coefficients regardless of sign or significance, in another set we set opposite-signed coefficients to zero (either using all of the regressions or only those run in the cross-section), and in another set we set coefficients that were not statistically significant to zero. We then re-scaled the coefficients to sum to one.

We also created a set of weights using stepwise regressions and principal component analysis, again re-scaling the resulting coefficients to sum to one. In the stepwise method, we used county level data and the employment to population ratio, the unemployment rate, and the employment growth rate as dependent variables. For each model, the index variables that were not dropped get the same weight and we also calculated a set of weights after dropping variables with signs opposite of what was predicted. The final stepwise weights were the average of each index variable's weight across the regressions (six total, three with and three without removing opposite-sign coefficients).

The principal components weights were constructed with the first three principal components of the set of index variables. Using the data sample from April to November 2020, we found that the first principal component loads heavily on the receptivity component because all these three variables do not vary in the sample period. The second principal component picks up the reception component mostly, in particular for the percentage changes of microbusinesses and customers. The third principal component loads more on three variables: the fraction of microbusinesses connected to a GoDaddy SSL, the fraction of microbusinesses connected to a website, and the average traffic index. We averaged the factor loadings of these three major principal components. For those variables that had negative signs (opposite sign), we set their weights to zero and then rescaled the weights so that they sum up to 100%. Including the simple index weight (which assigned a weight of one to the density variable and zero to all others) and the even weight (which was a simple average of the index variables), we had nine sets of candidate weights (see Table 3).

With the sets of weights in hand, the process of creating the candidate indices was straightforward. We took the full set of normalized data, truncated each of the index variables at the 95th percentile to limit the influence of outliers without reducing the number of areas for which we could calculate the index, and then created the various candidate indices as weighted averages of the normalized, truncated variables. We then re-scaled all indices and centered them to average 100 in April 2020.

To select our preferred index, we based our choice on which index was superior according to our criteria of goodness of fit of the index (how well the index explained local labor market conditions and small business activity). We tested all of the candidate indices using data from November 2020 – March 2021 (the test set). Generally, the various candidate indices were highly correlated, including the simple index and the even-weight index that did not use regression techniques to determine variable weights. One interpretation is that all of the candidate indices are picking up similar trends in online microbusiness activity and that the specific choices we made regarding the various regression methods (e.g. what dependent variables to use and what model specifications to use) did not affect the index substantially.

To test the candidate indices, we ran regressions of the form $y_{it} = \beta_0 + \beta_1 y_{it-1} + \beta_2 Index_{it}$ where t indexes time in months and i indexes counties, CBSAs, or states. The dependent variables included the employment to population ratio, the unemployment rate, employment growth (month over month percent change; no lag was included in this case), and new business applications (for state-level regressions only). Lags of the dependent variables were included to capture the persistence of these economic indicators.

For each regression, we calculated the residual sum of squares to capture the model's fit and ranked the candidate indices based on this metric (with a lower rank indicating better fit). The index with the lowest mean rank across all regressions in our tests was selected as our preferred index, which is the index that set opposite-signed coefficients to zero, which we call the Baseline Index. This index is the one described in the main text of the paper. While this index is currently our preferred index, we will continue to test and refine the index over time as a longer time series of data become available. Even in the current data, the candidate indices performed similarly (similar model fit measured by sum of squared residuals), so that as more data become available, other indices, such as the even-weight index, could prove to be better. Another reason why it is especially important to refine this index as time progresses is that the current data sample covers only the recession that began in 2020, which has been an unusual economic time period.

Appendix C

Table C1. Access to Capital and Microbusiness Index, Univariate Regressions

Panel A

	Dependent variables				
Independent variables	Microbusiness density	Microbusiness owner density	Composite index	Reception index	Activity index
	(1)	(2)	(3)	(4)	(5)
# banks	0.032***	0.006***	0.045***	0.009***	-0.011***
	(0.005)	(0.002)	(0.003)	(0.001)	(0.002)
Constant	6.317***	2.220***	101.286***	100.503***	105.984***
	(0.356)	(0.130)	(0.208)	(0.039)	(0.145)
Observations	2,670	2,670	2,646	2,670	2,646
Adjusted R^2	0.017	0.005	0.100	0.097	0.013

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. The sample period is March 2021.*

Panel B

	Dependent variables				
Independent variables	Microbusiness density	Microbusiness owner density	Composite index	Reception index	Activity index
	(1)	(2)	(3)	(4)	(5)
# community banks	0.117***	0.023***	0.237***	0.036***	-0.024***
	(0.023)	(0.008)	(0.013)	(0.003)	(0.009)
Constant	6.094***	2.178***	100.281***	100.403***	105.891***
	(0.401)	(0.146)	(0.232)	(0.045)	(0.164)
Observations	2,670	2,670	2,646	2,670	2,646
Adjusted R^2	0.009	0.002	0.109	0.067	0.002

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. The sample period is March 2021.*

Panel C

	Dependent variables				
Independent variables	Microbusiness density	Microbusiness owner density	Composite index	Reception index	Activity index
	(1)	(2)	(3)	(4)	(5)
% community banks	-0.043***	-0.008**	-0.108***	-0.018***	0.039***
	(0.011)	(0.004)	(0.006)	(0.001)	(0.004)
Constant	9.681***	2.850***	108.774***	101.757***	103.517***
	(0.681)	(0.245)	(0.390)	(0.076)	(0.272)
Observations	2,657	2,657	2,633	2,657	2,633
Adjusted R^2	0.006	0.001	0.107	0.079	0.030

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. The sample period is March 2021.*

Panel D

Independent variables	Dependent variables				
	Microbusiness density	Microbusiness owner density	Composite index	Reception index	Activity index
	(1)	(2)	(3)	(4)	(5)
# PPP loans up to150K (thousands)	0.145***	0.028***	0.179***	0.039***	-0.060***
	(0.026)	(0.009)	(0.015)	(0.003)	(0.010)
Constant	6.755***	2.304***	102.022***	100.622***	105.861***
	(0.342)	(0.125)	(0.205)	(0.039)	(0.139)
Observations	2,670	2,670	2,646	2,670	2,646
Adjusted R^2	0.011	0.003	0.049	0.062	0.012

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. The sample period is March 2021.*

Panel E

	Dependent variables				
	Microbusiness density	Microbusiness owner density	Composite index	Reception index	Activity index
Independent variables	(1)	(2)	(3)	(4)	(5)
dollar value (in mil.) all PPP loans up to 150K	0.005***	0.001***	0.006***	0.001***	-0.002***
	(0.001)	(0.0003)	(0.001)	(0.0001)	(0.0004)
Constant	6.712***	2.295***	101.969***	100.614***	105.881***
	(0.342)	(0.125)	(0.204)	(0.038)	(0.139)
Observations	2,670	2,670	2,646	2,670	2,646
Adjusted R^2	0.013	0.004	0.056	0.069	0.014

*Notes: In parentheses are standard errors. *** denotes statistical significance at the 1% level, ** at the 5% level, * at the 10% level. The sample period is March 2021.*