ifh Working Paper No. 25/2020

Assessing the 'Digital Divide' and its Regional Determinants: Evidence from a Web-Scraping Analysis

Anita Thonipara[a], Rolf Sternberg[b], Till Proeger[a], Lukas Haefner[b,*]

[a] Institute for Small Business Economics at the Georg-August-University Goettingen, Heinrich-Düker-Weg 6, 37073 Göttingen, Germany
[b] Institute of Economic and Cultural Geography. Leibniz University Hannover, Schneiderberg 50, 30167 Hannover, Germany

Abstract:

Following the 'death of distance' postulate, digitization may reduce or even eliminate the penalty of firms being located in rural areas compared with those in urban agglomerations. Despite many recent attempts to measure digitization effects across space, there remains a lack of empirical evidence regarding the adoption of digital technologies from an explicit spatial perspective, i.e. comparing urban with rural areas. Using web-scraping data for a representative sample of 345,000 German firms, we analyze the determinants of homepage usage. Accordingly, we show that homepage usage - as a proxy for the degree of digitization of the respective firm - is highly dependent on location, whereby firms in urban areas are more than twice as likely to use webpages than those located in rural areas. Our county-level analysis shows that a high population density, young population, net gains in internal migration, high educational level and high firm-specific revenues have a positive and significant effect on the probability that firms conduct digital marketing using webpages. Access to broadband internet has a positive effect in rural areas. There are no differences between urban, suburban and rural areas in terms of webpage up-to-dateness as well as social media usage. We conclude that there is a substantial digital divide in online marketing and discuss policy implications.

JEL: D22, L22, L26

Keywords: digital divide; digitization; Germany; rural; Small and Medium-Sized Enterprises; urban; web-scraping

* Corresponding Author. haefner@wigeo.uni-hannover.de

The work of Lukas Haefner on this paper has been funded by the Lower Saxony Ministry of Science and Culture under grant number ZN3492 within the Lower Saxony "Vorab" of the Volkswagen Foundation and supported by the Center for Digital Innovations (ZDIN). The work of Anita Thonipara and Till Proeger has been co-funded by the European Regional Development Fund under the grant number ZW6-85017668.

1. Introduction

There is a public perception of a growing divergence in the adoption of digital technologies among urban and rural firms. Digitization is considered the major technological development driving future business models and thus regional and firm-level competitiveness. Acknowledging a growing divergence in digitization and competitiveness could warrant far-reaching innovation and digitization support for rural firms to prevent a further polarization in competitiveness. Such support is timely since the adoption of digital technologies can be of a cumulative nature: once a firm lacks the capacity to adapt to new digital developments, further developments are even more difficult to cope with, eventually causing a further growing polarization. This is problematic because the unequal distribution of competitiveness and wealth among firms and regions is already leading to social tensions in many countries (Dijkstra et al., 2019; Rodríguez-Pose, 2018). Digitization thus poses a risk to social cohesion when new digital technologies are primarily adopted in and benefiting thriving regions. Additionally, digitization yields positive and negative impacts at the same time. For example, workers performing routine tasks face a high risk of substitution through digital technologies, while workers executing complex tasks tend to be complemented (Akerman et al., 2015). This could amplify the threat for social cohesion when thriving regions benefit from digitization, while at the same time less-wealthy regions suffer from job degradation or job automation. Consequently, in order to be capable of benefiting from digitization, policy-makers at many geographical levels have created government programs to improve the quality and quantity of broadband access in all regions and all types of regions.[1]

Contrary to the threat of causing further polarization among firms and regions, digitization has long been seen as an opportunity to reduce inter-regional disparities, in particular from politics (BMI, 2019; Grimes, 2005). Especially in regions that have previously largely been excluded from markets or value chains, digital technologies could allow for a better inclusion of firms (Galloway et al., 2011). These potentially growing opportunities for firms were based on the expectation that the importance of physical proximity would decrease, eventually triggering the 'death of distance' (Cairncross, 1997). Once all regions were digitally connected through the internet, firms could digitally exchange knowledge, ideas, services or even products. Thus, firms in marginalized rural regions could overcome existing disadvantages through integration into previously-inaccessible networks and value chains, whereby these regions could consequently catch up to more advanced regions.

However, although most firms – including in rural regions – now have access to the internet, such digitization-induced catch-up processes have hardly been witnessed, prompting some authors to revise their expectations based on a more recent perspective (Cairncross, 2018). The absence of catch-up processes may be due to the digital divide among firms in rural and urban regions, which features two interconnected facets (Salemink et al., 2017). These facets can be conceptualized as the first- and second-level digital divide, as has been differentiated in (aspatial) research on the digital divide among individuals or social groups (Büchi et al., 2016; Scheerder et al., 2017). The first-level digital divide refers to access. While nowadays most firms in rural regions have access to the internet, the quality of the connection often trails behind the quality in urban regions (Briglauer et al., 2019b; Prieger, 2013) and thus limits the rural firms' opportunities of digital participation. The second-level digital divide refers to usage. The adoption of not only the internet but also other digital technologies in rural firms trails behind their urban counterparts, even if these technologies are available. Research, which is mostly carried out on households and not on firms, shows that socio-demographic characteristics of regions such as income, professional qualification, a young population and labor force are central determinants in fostering the usage of digital technologies (Blank et al., 2018; Billon et al., 2016; Prieger, 2013; Schleife, 2010). However, often these indicators are deficient in rural regions.

The digital divide among firms in urban and rural regions may still play an important role in hindering catch-up processes of rural regions to more prosperous urban regions or it may even exacerbate inter-regional disparities. While a spatially unequal accessibility of the internet infrastructure is easily observable (first-level digital divide), it is less straightforward to determine a digital divide in the actual adoption of digital technologies (second-level). However, exactly this adoption is needed to seize opportunities from digitization, not only for firms but also for regions and countries as a whole (Whitacre et al., 2014a & 2014b).

Determining whether there is a second-level digital divide among urban and rural firms is therefore an important research goal to substantiate or refute political calls for a novel support of rural firms aimed at preventing a decoupling of digitization dynamics. Nevertheless, empirical contributions face a number of obstacles, particularly in determining and compiling suitable indicators measuring digitization in firms. Since digitization is a very

[1] See e.g. the Digital Agenda for Europe for an attempt at the supra-national level of the EU (https://ec.europa.eu/digital-agenda/en), the "digital agenda" for an attempt at the national level (for Germany, see the Federal Government of Germany, 2018) or the 'Broadband for Bavaria' initiative for the sub-national level of The German Federal State of Bavaria (https://ec.europa.eu/digital- single- market/en/policies/ improving-connectivity-and-access; see also Briglauer et al., 2019a).

heterogeneous and industry-specific development, there are few common indicators that can be used to compare digitization across industries and regions.

To date, most indicators aimed at capturing digitization and determining the digital divide between regions refer to the internet. While initial attempts to capture the digital divide have focused on access, research has more recently considered the speed of the available internet infrastructure (Salemink et al., 2017; Townsend et al., 2013) or the number of broadband providers (Prieger, 2013). However, the digital divide has recently shifted from access (first-level digital divide) to usage and skills (second level), at least within advanced economies (Büchi et al., 2016; Evangelista et al., 2014). Accordingly, in order to measure the second-level digital divide, indicators for the actual usage and the necessary skills for usage of the internet are considered (Billon et al., 2016; Prieger, 2013; Scheerder et al., 2017). However, it is important to note that these variables mostly refer to households rather than firms.

Proceeding beyond the adoption of single indicators, some digitization indices have been elaborated that allow for a geographical digitization comparison and thus determining a potential digital divide. The Enterprise Digital Development Index developed by Ruiz-Rodríguez et al. (2017) uses data for Spanish regions (NUTS2-level) and focuses solely on the digitization degree of firms. It comprises four components: ICTs and connectivity (access) to the internet; the use of ICTs; e-commerce; and e-government. The most interesting indicators for this study are the indicators of the "use of ICTs" component: enterprises having a website or homepage; the use social networks; cloud computing services used over the internet; employees using computers; enterprises that pay to advertise on the internet; and finally, indicators of the "e-commerce" component: enterprises sending e-invoices B2BG, suitable for automated processing; and enterprises receiving e-invoices, suitable for automated processing. These indicators are measured as the percentage of enterprises in a region, except for "persons employed using computers", which is measured as the percentage of total employment. The European Commission (EC, 2020) has developed the Digital Economy and Society Index, which comprises data on countries rather than regions. It comprises five main dimensions: connectivity, human capital, use of internet, integration of digital technology, and digital public services. The "integration of digital technology" dimension captures the degree of digitization in firms. Its first sub-dimension of "business digitization" includes the indicators electronic information sharing, social media, big data, and cloud. Its second sub-dimension of "e-commerce" includes the indicators SMEs selling online, e-commerce turnover, and selling online cross-border. Katz and Koutroumpis (2013) have developed a digitization index that also enables comparing between countries, but not between regions. It comprises six components: affordability, infrastructure reliability, network access, capacity, usage, and human capital. However, it does not comprise any detailed information on firms' digitization degree.

In this paper, we suggest using a firm's decision to create and promote a webpage in the yellow pages, its respective up-to-dateness and social media references as a proxy for its degree of digitization. While these indicators primarily capture online marketing, we argue that its basic nature entails a firm's propensity to conduct further digitization measures. Put simply, if a firm geared at B2C services refrains from using a webpage as a near-costless measure of marketing, it is highly unlikely that this firm has invested in digitizing its internal workflows, let alone uses advances digital technologies in its production. We are thus confident that these indicators capture a firm's basic propensity to conduct digitization measures and hence enable us to test differences in digitization across regions. Moreover, it can be assumed that the usage of webpages and social media has a fairly similar function for firms, even in very different (consumer-oriented) sectors. Since most other digital technologies are used quite heterogeneously across sectors, webpage usage enables us to compare larger samples of firms.

Using this approach to measuring digitization, we run a web-scraping algorithm, which gives us roughly 345,000 firm entries of the yellow pages and about 104,000 webpages, which are further analyzed regarding social media references and up-to-dateness. Using the firms' postal codes, we link county-level data for various socio-economic indicators. This provides us with evidence regarding the occurrence and potential structural determinants of a digital divide among German SMEs.

This article's contribution to the literature is twofold. First, it adds a simple but widely-accessible measure of the digitization degree of regions, which does not build on the digital infrastructure but rather the actual usage of digital technologies. Indicators measuring the usage of digital technologies in firms are thus very scarce, especially at the regional level for Europe (Billon et al., 2016; Ruiz-Rodríguez et al., 2017). The proposed indicator comprises a large dataset and a wide range of sectors and is feasible for many countries without raising measurement and data accessibility problems that are inherent in indices. Second, by applying this indicator, new insights into the spatial digital divide and its regional determinants are presented. While we confirm an urban-rural digital divide in webpage usage, we do not find a divide in the technological specifics of these webpages. In accordance with existing literature, we find evidence that population density, the professional qualification of the workforce, a lower population age and (internal) immigration are important regional determinants that are positively associated with the regional digitization level. Contrary to existing findings, the wealth of regions (GDP) is not positively associated with webpage usage. Finally, our results tend to show new evidence that lacking digital infrastructure is still an important digitization constraint for some rural regions.

The remainder of this paper is structured as follows. Chapter two reviews the relevant literature for our research goal, before chapter three presents our method and dataset. Chapter four presents our results, while chapter five discusses our results and presents policy implications.

2. Literature review

2.1 Related research

2.1.1 Definitions of 'digital divide'

In principle, a digital divide refers to the different availability, usage and usability of digital infrastructure, which may lead to gaps between people and groups of people with specific attributes within a society. Such a divide may be related to various joint characteristics of the named people or groups of people, e.g. their age (old vs. young), income (rich vs. poor) or education (university degree vs. no degree). In our paper, we follow the most common interpretation in the academic literature (see Graham, 2019), which refers to spatial entities (territories or spatial units) when referring to a digital divide. These territories may be assigned to all possible spatial dimensions, such as the global (e.g. comparison of continents), national (comparing countries), regional (i.e. comparing sub-national ones like the 50 states in the US or the sixteen federal states in Germany) or the local one (e.g. comparing the quarters of a city). In addition to these spatial units based on administrative boundaries, the comparisons between urban and rural regions have attracted strong attention among scholars (see Rodríguez-Pose, 2018) in recent years, partially driven by specific policy programs and interests to ensure equal living conditions in all regions of a country (Briglauer et al., 2019a). These comparisons between urban and rural areas may be intra-national (all areas within a country), cross-national (all areas in various or all countries within the same continent) or cross-continental.

In the following description of the state-of-the-art research on the (spatial) digital divide, we distinguish between the first- and the second- level digital divide. While the former has already provoked a lot of empirical and some theoretical research, the latter has not. We will explicitly address urban-rural differences in our overview of the literature.

2.1.2 First-level digital divide: Digital infrastructure and its spatial effects

When assessing the potential effects of digital infrastructure in rural areas, it is relevant to note that the high-quality (i.e. in most cases, high-speed) digital infrastructure (e.g. super-fast internet) is more often and much earlier available in economically strong urban areas in a given country, which - ceteris paribus - leads to an increase of inequalities between urban and rural areas. The persistence of this divide is often ascribed to liberalized telecommunications markets and the fact that internet expansion for telecommunications providers is more profitable in urban rather than rural regions (Briglauer et al., 2019a; Grimes, 2003). In some cases, the spatial digital divide is even growing (Townsend et al., 2013), which is often the case particularly for novel technologies or next generation access (NGA), i.e. higher bandwidths (Briglauer et al., 2019b; Prieger, 2013). Politicians have recognized this problematic divide and many policies have aimed to narrow this gap (Briglauer et al., 2019a), although they often quickly become outdated by new technological developments (Salemink et al., 2017).

Regarding the spatial effects of digital infrastructure, there is some empirical evidence in the literature if inter-regional differences between sub-national regions are considered (Bertschek et al., 2015; Grimes, 2005, Malecki, 2003). However, these empirical results are often contradictory regarding the extent, type and direction of such effects (Jung/López-Bazo, 2020). This also holds true for the effects on rural vs. urban regions. Some studies show positive effects of digital infrastructure for rural areas (Townsend et al. 2017). To give an example, in many rural areas in developing countries, digital infrastructure allows time to integrate such regions in global value chains and enables a low-cost and reliable exchange with extra-regional and international customers (Krone/Dannenberg, 2018). However, only in some cases are these effects also positive in a relative sense, i.e. they are stronger than the positive effects of the same infrastructure in favor of urban regions (Fabritz, 2013; Bertschek et al., 2015; Ivus/Boland, 2015). Such relatively (compared with urban regions) strong effects of digital infrastructure would be a necessary condition if digitization is intended to be used as a regional policy instrument to reduce inequalities between urban and rural areas of a country, i.e. to achieve convergence instead of (further) divergence (Jung/López-Bazo, 2020; Celbis/de Crombrugghe, 2018). If urban and rural areas benefited from digital infrastructure to the same extent, then inter-regional disparities between rural and urban areas would not be reduced. By contrast, many studies argue in favor of (often economically strong) urban regions regarding the positive spatial effects of digitization (Camagni/Capello, 2005; Ciffolilli/Muscio, 2018). Digital infrastructure seems to have stronger positive impacts in urban rather than rural areas in terms of both quantitative as well as qualitative aspects (Malecki, 2002, Tranos, 2016; WWG, 2015). In particular, this seems to be true for young (or even new) and knowledge-intensive firms (see Mack et al., 2011; McCoy et al., 2018) and for higher bandwidth levels (Briglauer et al., 2019a).

2.1.3 Second-level digital divide: Regional determinants and spatial effects

As shown above, research on whether digitization (via the internet) mainly benefits urban or rural regions (and thus whether digitization leads to increased or reduced inter-regional disparities) yields inconclusive results. The availability of digital infrastructure (high quantity and/or high quality) is nowadays only a necessary condition for digitization's effects on the development of urban and rural areas (Evangelista et al., 2014; Tranos, 2012) and an increase or decrease in economic disparities between the two types of areas. If this first-level component meets higher or even the highest demands, the regional economic effects would nevertheless remain very limited if people living and/or working in these areas lack the necessary motivations and skills required for exploiting the potential offered by a very high standard infrastructure (second-level digital divide). In other words, the second-level component is decisive if a certain standard of the first-level component is provided (Evangelista et al., 2014). Studies confirm this view, showing that increased internet availability does not necessarily have positive (spatial) impacts, but rather that the usage matters (Mack/Faggian, 2013; Whitacre et al., 2014a, 2014b; Scheerder et al., 2017). Consequently, in order to answer the question asked above regarding whether digitization mainly benefits urban or rural regions, it appears more fruitful to determine a spatial digital divide in usage rather than accessibility.

A number of studies comparing urban and rural regions find a digital divide in using digital technologies, which does not stem solely from lacking access in rural regions (Moriset et al., 2012; Prieger, 2013; Salemink et al., 2017; Townsend et al., 2013). Thus, there must be some regional determinants apart from infrastructure availability that explain this second-level digital divide. Studying the impact of these determinants on the degree of digitization and their regional variation can thus help to better understand the digital divide.

The amount of theoretical or empirical research regarding the impact of regional determinants on the urban-rural second-level digital divide is rather small, particularly for firms. It seems obvious that the lack of highly-educated employees with profound digital skills is one reason for various problems of firms in rural areas, when trying to exploit the opportunities offered by digital technologies in such areas (Camagni/Capello, 2005; Grimes, 2005, Prieger, 2013; Salemink et al., 2017). These differences in terms of human capital in general – and digital skills in particular – between the two types of regions result from the higher attractiveness of urban areas for the young well-educated workforce (Blank et al., 2018; Schleife, 2010, Mack et al., 2011). Socio-demographic variables are often found to explain large parts of the regional degree of digital technology usage and thus explain the second-level digital divide (Billon et al., 2016; Blank et al., 2018; Prieger, 2013; Schleife, 2010). Accordingly, literature shows that it is not necessarily "rurality per se", a low population density or geographical distance but regional characteristics like low income and education levels or a high average age of rural inhabitants and employees that hinder digitization-induced catch-up processes of rural regions (Blank et al., 2018; Mills and Whitacre, 2003; Schleife, 2010).

2.2 Hypotheses

It has been argued that regional determinants strongly influence the adoption of digital technologies and accordingly affect the second-level digital divide. However, especially for Europe, there is a lack of regional data for the usage of digital technologies in firms (Billon et al., 2016; Ruiz-Rodríguez et al., 2017), which underlines the need for developing new indicators. The regional intensity of webpage usage in firms - addressed in our paper - serves as a proxy for the regional degree of digitization. We expect that it is only partially dependent on access to high-speed internet, whereas other regional determinants also play a major role. We formulate several hypotheses for regional determinants that may explain the regional degree of webpage usage of firms.

H1: Population density has a positive effect on webpage usage. The first hypothesis follows the diffusion theory, stating that technology and its adoption diffuses through the spread of information (spillovers), which mostly occurs in personal interactions. Since a higher population density encourages personal interaction (and the supply of technologies), technologies should be spread and adopted faster and more extensively in cities (Billon et al., 2016; Salemink et al., 2017). Network effects also encourage the adoption of technologies because the regional number of users of a technology should positively influence the number of new users of this technology in the same region (Li and Shiu, 2012; Schleife, 2010).

From another perspective, and with respect to the specific indicator used, rural firms traditionally acquire their orders from personal networks and regional reputation, which makes webpages less relevant. Urban regions have less established social networks, which makes more anonymous forms of marketing more relevant. Thus, increased population density should be reflected in a higher degree of online marketing.

H2: (Internal) migration into a region increases webpage usage. The majority of migrants in advanced countries are young adults who migrate for tertiary education or in early stages of their careers, as is the case for this article's study area. This group tends to have profound digital skills compared with older people, particularly through life-long learning ("digital natives"). Regions with a higher share of immigrants should thus have a stronger propensity to use digital technologies like webpages for digital marketing.

Following a second line of argumentation, regions with high levels of migration imply new inhabitants outside of traditional reputational social networks who require additional information on firms. Therefore, additional inhabitants are likely to drive webpage usage, and vice versa. This effect should also be reflected in the average sectoral revenue, which should be positively associated with inward migration due to the ensuing higher regional demand.

H3: GDP has a positive effect on webpage usage. In line with previous research, we expect GDP to be a strong predictor of regional digitization (Billon et al., 2016; Vicente and López, 2011). The GDP per capita of a given region is a measure of firm competitiveness, higher education of the workforce and higher overall purchasing power, all of which are likely to affect firm innovativeness, which should be reflected in higher levels of digitization (webpage usage). Furthermore, as digital technologies are not free of cost, their usage also depends on investment decisions (Schleife, 2010) and it should therefore increase with household income and GDP (Li and Shiu, 2012).

H4: The availability of high-speed internet positively affects webpage usage. The respective infrastructure enables firms to invest in digitization and customers to change consumption patterns, both of which are likely to positively affect webpage usage. With this hypothesis, we test the first-level digital divide, i.e. whether accessibility still is a barrier to digitization for some firms and regions.

H5: Higher levels of professional qualification in the workforce increase webpage usage. From a capabilities approach, it can be expected that higher shares of employees with professional qualification will increase webpage usage by firms and demand for online information by customers of a given region. This expectation is in line with previous findings in the literature. At the firm level, a number of studies have shown that professional qualification is a central determinant for the effective use of digital technologies (Akerman et al., 2015; Berger and Frey, 2016a, b). From a spatial perspective, digital technologies seem to be used effectively particularly in regions with high levels of skilled labor (Salemink et al., 2017; Schleife, 2010), which in turn benefit from digitization (Atasoy, 2013; Hasbi, 2020; Mack and Faggian, 2013; McCoy et al., 2018).

H6: Higher average age of a region's inhabitants is associated with lower webpage usage. Demographic variables have been shown to be strong predictors of the use of digital technologies, not only from a spatial perspective (Büchi et al., 2016; Scheerder et al., 2017). For regions, it has been shown that the age structure is an important determinant of internet use (Prieger, 2013; Schleife, 2010; Vicente and López, 2011). Older employees in firms tend to be less digitally skilled and less willing to accept organizational changes through digitization (Billon et al., 2016; Blank et al., 2018; Evangelista et al., 2014). Behavioral patterns associated with digitization are thus more strongly adopted among younger persons. Therefore, we expect that regions with a higher average age structure are less likely to foster high webpage usage.

H7: Competition among firms in the same sector increases webpage usage. A larger number of firms from the same sector in the respective region implies stronger competition, which can foster technology adoption (Zhu et al., 2003). Furthermore, a higher density of firms in the same sector increases the potential for knowledge spillover in the field of digitization, in line with the network argument from above (Li and Shiu, 2012; Schleife, 2010). Therefore, a higher density of same-sector firms is likely to induce a stronger webpage usage.

3. Method

3.1 Web-scraping

To obtain our data, we employ a web-scraping procedure, which was conducted in June 2018. Web-scraping can be described as a simple algorithm that downloads specific information from a predefined number of webpages. Once fed with a large number of webpage links, the algorithm can download the predefined information efficiently and it can be used to build a dataset encompassing the respective variables. To obtain information on firm webpages, we use the German yellow pages, which provides us with a large set of B2C firm contacts, their respective postal codes and - if entered - the webpage address. We choose a broad set of professions[2] that primarily represent firms from the craft sector, focusing on consumers to prevent analyzing e.g. industry suppliers, which might not have the same incentive to use digital marketing compared with consumer-oriented businesses. For firms

[2] The full set of professions is documented in Appendix 7.1.

that stating having a webpage address, we conduct a further web-scraping procedure, which provides us with the information concerning whether the webpage is actually online, when it was most recently updated and whether a social media reference (Facebook, Twitter, Instagram) is used. The information regarding whether a homepage has been updated is analyzed using three separate proxies. First, the “last update” variable indicates the year in which the webpage was last updated. This variable is only available for a small fraction of the sample, which is why we use two other proxies, namely “HTML5” (Hypertext Markup Language) and “HTTPS” (Hypertext Transfer Protocol Secure). HTML5 is the most recent language used for programming webpages and it stands for a modern appearance of the webpages. On the other hand, HTTPS indicates whether a company has invested in the website’s security, which could hence be another indicator of up-to-dateness.

This approach has the strong advantage of bypassing a number of problems associated with direct surveys or interviews regarding the state of digitization: low response rates, self-selection of digital pioneers, self-selection of larger firms, overly-optimistic self-assertions and regional or sectoral biases. Instead, a larger set of information can be used that directly measures firms’ behavior on a large scale across sectors and regions. The disadvantages of using the yellow pages as the source of firms’ webpages include the heterogeneous structure of the yellow pages in Germany, historically different cost models for entering firm information as well as potential errors during the web-scraping procedure itself, resulting in a loss of data. Furthermore, using the yellow pages might exclude digital pioneers that rely entirely on other platforms for finding customers. While the yellow pages claims that it remains the most important platform for finding professionals and that half of all Germans regularly use the yellow pages, we ultimately cannot assert that each firm has a strong incentive to use the yellow pages and enter its own webpage. Thus, we can conclude that our dataset presents a large and regionally-balanced set of firms, while acknowledging that it fails to present a fully representative picture of the respective sectors.

3.2 Sample specifics

Overall, our sample comprises 346.361 firm entries, of which 104,460 (.302) have entered a webpage in the yellow pages have was online and working in June 2018. We downloaded firm data from 44 different professions with quite heterogeneous numbers of firms and shares of webpage usage. An overview of the professions and respective number of observations is documented in appendix 7.2. The web-scraping procedure further yielded information on the existence of a Facebook/Twitter/Instagram reference in the homepage, as well as information on the most recent update and the use of the HTML5 and the HTTPS standard.

At a regional level, there was no limitation during the web-scraping procedure, and thus all firms of the respective professions were used, irrespective of their regional origin. We used the postal codes associated with the firm entries and matched this information to the respective counties. This approach leads to inaccuracies in some cases, whereby the associated error can be assumed to be uniformly distributed over all counties/postal codes. With firms clustered at the county level, we matched a number of additional county-level regional indicators, namely population density, the share of employees without a professional qualification, the share of employees with an academic education, the revenue of craft firms, gross domestic product per inhabitant, regional rates of immigration and emigration, the development of the population aged 65 and older, the number of employees in craft firms,[3] and finally the share of households with high-speed internet access.[4]

Using the web-scraping data as well as the additional regional indicators, we can analyze the regional disparities in webpage usage, social media usage and up-to-dateness of webpages and show potential drivers of the differences using regional indicators.

[3] More specific information on the variables is presented in appendix 7.2.

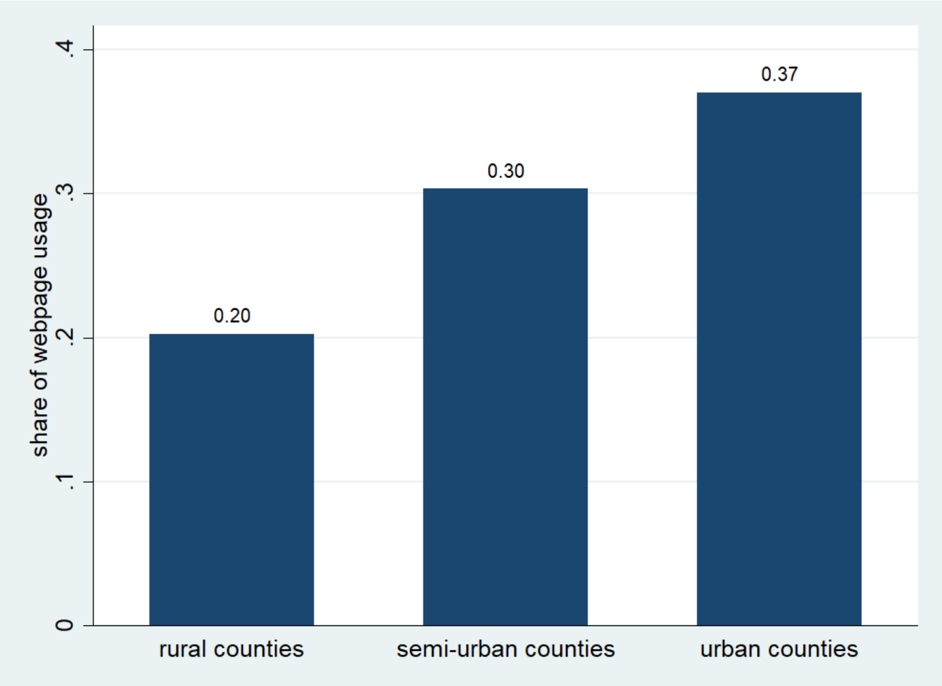
[4] To determine a region’s access to high-speed internet, we use a dataset from the TÜV Rheinland and the German Federal Ministry of Transport and Digital Infrastructure at the county level. The dataset gives us the share of households that can access an internet speed of 50 Mbit/s or higher and households with an internet speed of 30 Mbit/s or higher for 2017. We thus build a variable subtracting the share of households with 30 Mbit/s from those with 50 Mbit/s in each county. This provides us with a variable for the share of households in each county with access to 50 Mbit/s. For urban counties, there is essentially no variation in the data as essentially all households can access 50 Mbit/s, there is little variation for semi-urban counties and strong variation for rural counties.

4. Results

4.1 Descriptive results

The basic finding regarding webpage usage can be illustrated using the county-level data and the respective classification of all German counties into rural, semi-urban and urban.5 Table 1 shows how the shares of webpage usage differ between the three types of regions, with urban counties showing roughly a double share of webpage usage compared with rural counties. Obviously, rural firms have a substantially lower propensity to conduct digital marketing compared with urban firms.

Table 1. Share of webpage entries of firms classified as rural, semi-urban and urban counties.



Thus, there is an obvious and strong digital divide in terms of webpage usage that warrants further analysis and robustness checks. Furthermore, comparing firms' propensity to update their webpages, we can use the same broad classification of counties. Recall that the "last update" variable is common among older webpage formats and that it can only be obtained from a low share of webpages. Therefore, the two proxies HTML5 and HTTPS - which were introduced roughly four years ago - can be used to comment on the up-to-dateness of webpages. They can both be interpreted as follows: if a firm has conducted a substantial technical update of their webpage to the current format in the past four years, it has most likely used these formats, which have become the respective standard in webpage design (HTML5) and communication encryption (HTTPS). Table 2 provides an overview of the share of webpages featuring the respective variable according to the classification of the respective county.

[5] This classification is based upon the definition of the three regional types by the German Federal Institute for Research on Building, Urban Affairs and Spatial Development BBSR, which encompasses the following criteria (see also https://www.inkar.de/documents/Erlaeuterungen%20Raumbezuege19.pdf): **Urban regions**: regions in which at least 50% of the population is living in large and medium-sized towns and which includes a major city of at least 500.000 inhabitants or regions with a population density (excluding major cities) of at least 300 inhabitants per square kilometer. **Semi-urban**: Regions in which at least 33% of the population lives in large or medium cities with a population density of 150 to 300 inhabitants per square kilometer and regions with at least one major city with a population density (excluding major cities) of at least 100 inhabitants per square kilometer. **Rural regions**: regions in which less than 33% of the population is living in major or medium cities with a population density of under 150 inhabitants per square kilometer and regions with a major city but with a population density (excluding major cities) of under 100 inhabitants per square kilometer.

Table 2. Webpage update indicators.

	update 2018	HTML5	HTTPS
Urban	.247	.573	.296
Semi-urban	.236	.558	.287
Rural	.239	.551	.291

Interestingly, unlike the webpage usage itself, there are very small differences between urban, semi-urban and rural regions in terms of up-to-dateness of the webpages. While urban firms tend to have slightly higher shares on the three indicators, the differences are weak at best. From this, it can be interpreted that rural firms are less likely to use webpages, although the firms that do tend to keep their webpage similarly up-to-date as urban firms.

A similar descriptive analysis can be conducted regarding social media usage, as shown in table 3. Regarding social media, we find a similar pattern, whereby urban firms tend to have slightly higher shares of social media plugins, yet no substantial effects can be observed.

Table 3. Share of social media plugins in the webpages.

	Facebook	Instagram	Twitter
Urban	.309	.059	.179
Semi-urban	.297	.045	.171
Rural	.311	.047	.177

Thus, our core descriptive result is that firms in rural counties have a substantially lower propensity to use webpages. However, those rural firm using webpages show a similar disposition to keep their webpages updated and connected to social media. Building upon these descriptive results, we run two regressions aiming to determine driving factors for the strong difference in webpage usage by firms.

4.2 Regression analysis

We run a probit regression to determine regional drivers for homepage usage with the entry of a homepage as our dependent variable and the different regional indicators and professions as independent variables. We run five separate specifications: using all firms of all regions while excluding broadband access, using all firms and all variables and using the sub-samples for rural, semi-urban and urban firms with all variables. Table 4 presents the regression results.

Most importantly, the regression confirms that a higher population density has a significantly positive effect on webpage usage for the overall sample, both including and excluding broadband access, which can be interpreted as a digital divide between rural and urban firms and the respective markets. However, the interpretation of this result is rather speculative. The positive impact of population density could either be due to the higher diffusion of the technology and network effects in cities or it could be interpreted as a rational business decision such that rural firms prefer direct interaction with their customers. On the other hand, in urban markets - which are often characterized by anonymity - digital marketing is more effective than personal reputation. This might still imply a future issue for the respective firms' competitiveness in case of changing social structures, internal migration and consumer demand for digitization. However, it points to two separate logics regarding the firms' marketing.

Result 1: There is evidence in favor of H1, implying a digital divide in terms of digital marketing.

Building upon this result, we ask whether migration into a region drives webpage usage. This variable proves to be significantly positive for the overall sample. Again, there are two conceivable interpretations: first, anchored on the supply side of digital marketing, this result might support the notion that regions with many immigrants have a higher workforce of "digital natives" who adopt digital marketing instruments in the firms; and second, anchored on the demand side, the result might provide evidence in favor of an increasing importance of non-relational forms of marketing for firms to get in touch with new inhabitants. This second interpretation is supported by the fact that the positive effect in the overall sample is driven by the semi-urban regions. This can be explained by a common form of internal migration, namely by new families from inner cities into suburban villages, who often require craft services for building homes. At the opposite end, we find that emigration rates have a negative effect on webpage usage, which can be interpreted as confirming this interpretation.

Table 4. Probit regression to explain webpage usage * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

	All, excl. broadband internet	All	Rural	Semi-urban	Urban
Broadband internet (over 50 MB in 2017)		0.0002	0.0164***	-0.0068	-0.0143**
Population density	0.000172**	0.0002**	-0.0005	0.0001	0.0002
Share of employees with acad. education	0.0231**	0.0227**	0.0145	0.0423***	-0.0045
Share of employees without formal educ.	-0.0321*	-0.0329	-0.1020***	-0.0108	-0.164***
Turnover in craft sector	0.00611**	0.0061**	0.0119***	-0.0051	0.0076
GDP per inhabitant	-0.00759	-0.0076	0.0071	0.0020	-0.0063
Development of migration rate	0.00827	0.0081	-0.0022	-0.0106	0.0236*
Rate of reg. emigration	-0.0190**	-0.0187**	-0.0029	-0.0398***	-0.0044
Rate of reg. immigration	0.0174**	0.0171**	0.0006	0.0376***	0.0065
Development of inhabitants aged 65+	-0.0522***	-0.0521***	0.0438	-0.0414*	-0.1450***
Employees in craft sector overall	-0.0029	-0.0024	-0.0010	0.0547***	-0.0341
Apparatus construction	-0.0298	-0.0301	0.0496	-0.1400	0.0127
Optician	0.3030***	0.3031***	0.4392***	0.2790***	0.2780***
Concrete building	0.3410***	0.3410***	0.472o***	0.2800***	0.3230***
Well construction	0.3030***	0.3030***	0.2360**	0.6601***	0.2340**
Bookbinder	-0.0345	-0.0341	0.0337	-0.1741**	0.0107
Baker	-0.442***	-0.4421***	-0.4661***	-0.4100***	-0.5130***
Roofer	0.2271***	0.2270***	0.3580***	0.1730***	0.1950***
Electrician	-0.1220***	-0.1221***	-0.1810***	-0.1331***	-0.0967*
Butcher	-0.3910***	-0.3900***	-0.4020***	-0.3790***	-0.4150***
Floor tiler	-0.2000***	-0.2000***	-0.1630**	-0.2150***	-0.2270***
Hairdresser	-0.5340***	-0.5350***	-0.7010***	-0.4580***	-0.605***
Building cleaner	0.1080***	0.1080***	0.2810***	0.1260**	0.0389
Scaffolder	-0.0325	-0.0319	0.0841	-0.1340**	0.0039
Glazier	0.3780***	0.3780***	0.4460***	0.2990***	0.4280***
Goldsmith	-0.1240***	-0.1240***	-0.0725	-0.2910***	-0.0807*
Engraver	-0.0935	-0.0931	-0.0326	-0.2070*	-0.0563
Heating engineer	-0.0128	-0.0127	-0.0752	-0.0572	0.0640
Audiologist	0.3030***	0.3030***	0.5600***	0.3050***	0.1980**
Installer	0.4620***	0.4610***	0.2630***	0.3820***	0.3770***
Automotive mechatronics engineer	-0.0834**	-0.0831**	-0.0645	-0.0552	-0.1160**
Plumber	-0.06280*	-0.0628*	-0.0918	-0.1820***	-0.0072
Confectionist	-0.1530	-0.1530	-0.3890***	-0.4650***	0.0867
Furrier	-0.3160***	-0.3160***	-0.3760	-0.4530**	-0.2070*
Painter	0.0099	0.0100	-0.0108	0.0629	-0.0437
Mason	0.1640***	0.1640***	0.4440***	0.1330*	0.0887
Custom tailor	0.1610	0.1620	0.1670	-0.0281	0.2090*
Organ maker	-0.2400**	-0.2380**	-0.2280	-0.2600**	-0.3930**
Orthopedic technician	0.1350***	0.1340***	0.2330***	0.0831*	0.1370**
Interior decorator	-0.0765**	-0.0761**	0.0825	-0.2150***	-0.0699
Plumbing electrician	0.4910***	0.4910***	0.4380***	0.4720***	0.4590***
Heating engineering	-0.1370***	-0.1370***	-0.3060***	-0.0261	-0.1760***
Plumbing and heating installer	0.0250	0.0249	0.0080	-0.0843	0.0717
Smith	-0.2410***	-0.2410***	-0.2710***	-0.4210***	-0.1520***
Chimney sweeper	-0.6530***	-0.6530***	-0.6600***	-0.6460***	-0.7220***
Shoemaker	-0.5810***	-0.5810***	-0.4940***	-0.6270***	-0.5900***
Stone carver	0.0667*	0.0675*	0.2040***	-0.0139	0.1140
Roadmaking	-0.2070***	-0.2070***	0.0940	-0.4160***	-0.2320**
Plasterer	-0.1990***	-0.1990***	0.0847	-0.3010***	-0.1950***
Dry-cleaner	-0.4390***	-0.4390***	-0.2780***	-0.4790***	-0.5270***
Cabinetmaker	-0.0349	-0.0339	-0.0926*	-0.0775*	0.0316
Ceramist	-0.2980***	-0.2960***	-0.2050	-0.3160***	-0.3490**
Clocksmith	-0.2650***	-0.2650***	-0.4290***	-0.4060***	-0.1310*
Carpenter	-0.0298	-0.0285	-0.0106	-0.1060**	0.0741
Bicycle mechanic	.	.	.	.	.
Constant	-0.7740***	-0.8120**	-2.3370***	-0.7440	2.9630***
Observations	297,057	297,057	74,213	93,528	129,313

A further supportive indicator is that craft revenue has a significantly positive effect on webpage usage. Thus, regions with inwards migration and higher craft revenue feature a higher share of webpages.

Result 2: There is evidence in favor of H2, i.e. webpage usage is positively associated with migration (particularly in suburban regions) and higher sector-specific revenues.

It has been hypothesized that GDP per capita has a positive effect on webpage usage, which can be explained from a capability as well as an affordability perspective regarding firms and customers. We find no evidence in favor of this effect, whereby even less-wealthy regions with other properties - such as inward migration - have higher shares of webpage usage. This finding opposes previous studies showing that GDP is a strong predictor of digitization.

Result 3: GDP per capita is not associated with webpage usage, and thus H3 can be rejected.

Furthermore, we hypothesized that the availability of high-speed internet is likely to positively affect webpage usage by making investments in digitization more rewarding overall for firms, which should also affect webpage usage. For customers, it might increase internet usage overall and thus foster a stronger demand for online information on firms, thus prompting them to provide such information. For the entire sample, there is no effect for broadband access. However, when looking at rural regions, there is a significantly positive effect. There is no effect for semi-urban counties and a negative effect for urban counties. This result can be partly traced back to the respective broadband variable: there is only substantial variation in terms of households' access to internet with 50MBIT/second for rural counties. For semi-urban and urban counties, there are very few cases with an internet speed of less than 50MBIT/second. Thus, only the rural sub-sample should be interpreted. Accordingly, rural counties with high-speed internet are significantly more likely to have a higher share of webpages in craft firms. It goes without saying that this result says little about causality; rather, there might be a causal link between the existence of larger and innovative firms in particular rural counties, whose political pressure has led to broadband access and whose existence leads to the migration of high-skilled employees who require non-relational information on craft firms. Nevertheless, such causal links remain hypothetical.

Result 4: Broadband access in rural counties is positively associated with webpage usage, thus providing evidence in favor of H4.

Turning to result 5 regarding the level of professional qualification among employees at the place of residence in the regional workforce, we find that higher shares of academically-trained employees are associated with higher webpage usage, and vice versa higher shares of employees without professional qualification (at least in rural and urban counties) are associated with lower webpage usage. This hints at the knowledge-based nature of the current technological change through digitization: employees, firms and regions with sufficient education and knowledge can prosper, while those without are threatened by digitization. The result can also be interpreted in terms of the migration argument: younger, better-trained employees tend to be more mobile and work in expanding, more innovative firms. This leads to internal migration to the respective counties, in which higher demand for craft services and online information emerges. Less competitive regions tend to lose well-trained employees, which increases the relative weight of untrained employees.

Result 5: Higher levels of professional qualification in the regional workforce are associated with a higher webpage usage, which provides evidence in favor of H5.

Connected to this interpretation is a region's share of inhabitants of pension age (65+ in Germany). We find that counties with older populations have overall lower shares of webpage usage, which can be interpreted as resulting from the different affinity to using digital technologies among older generations, in the private as well as the occupational context. Moreover, it can be hypothesized that older persons have been living in the same region for a longer time and they have built up relational networks with craft firms during the course of their lifetime. Thus, online information on firms' services are less relevant since they have become repeat customers with specific firms over time.

Result 6: The share of inhabitants aged 65+ is negatively associated with webpage usage, which provides evidence in favor of H6.

Finally, we ask whether regional market competition motivates firms to invest more strongly in digital marketing. We therefore use the share of employees in craft firms per county as a proxy to determine the level of competition. There is no overall effect, yet semi-urban regions again exhibit a significantly positive effect. Thus, in semi-urban regions with a higher share of craft employees, there is a higher propensity to conduct digital marketing.

Result 7: Market competition - as measured by the share of craft employees in a given region - only has a positive effect on webpage usage in semi-urban counties. Hence, H7 is partially rejected.

Overall, we find a strong digital divide between rural and urban counties in terms of webpage usage. Looking at regional factors, we can show that regions with a younger, more skilled population and a higher internal migration rate display a higher webpage usage. This corresponds to the positive effect of higher sector-specific revenues and a positive effect of share of craftsmen in semi-urban counties. There is no effect of average GDP per capita, while broadband access is positively associated with webpage usage in rural counties.

5. Conclusion and policy implications

Interpreting the results of our dataset, we suggest that there are two separate systems of searching for craft firms from a customer perspective: the traditional long-term relationship with a specific firm, which is well-known in smaller, integrated communities and depends on its reputation and long-term relationships for customers. This mechanism is prevalent in rather stable regions with a lower internal migration dynamic and integrated social fabric. In this system, there is little incentive for digital marketing to acquire customers as the reputation in limited social networks and word-of-mouth recommendations lead to orders. This system depends on the long-term stability of social patterns and loses its effectiveness in disintegrated regions characterized by high internal migration, as younger employees in need of craft services locate in close proximity to their employers, often in suburban counties to accommodate family life. Suburban counties with this level of internal migration feature often have younger inhabitants on average, feature higher craft revenues and therefore have a higher share of craft employees. Due to the lack of long-term social reputations, other sources of information become important, which in recent times means more webpages. Therefore, we argue that the digitization of digital marketing is essentially customer-driven and that the respective intra-firm capabilities are developed once a larger set of customers requires other forms of information.

Interpreting the results of our dataset from a perspective of the digital divide and inter-regional disparities, we suggest another explanation: since some rural regions still seem to suffer from lacking broadband access, we find new evidence for the first-level digital divide, which could exacerbate urban-rural economic disparities since it still hinders firms in some rural regions from participating in digitization. There is also evidence in favor of a strong second-level digital divide in our sample, namely in usage differences that are not explained by lacking internet access. We argue that these differences can essentially be traced back to changing generational and social patterns in the respective regions. However, this does not rule out the notion that the reluctance to conduct digitization measures can amplify issues related to regional disparities: as all forms of digitization influence a firm's competitiveness, the lacking market pressure to open up to new technological developments might become a liability for regions as a whole. Thus, the market pull favoring younger, high-skilled employees by technologically-advanced firms might further increase the respective region's technological development and serve as a competitive advantage. Vice versa, well-established market structures by an aging population and the lack of novel market demands could delay the absorption of new technologies and therefore serve as a competitive disadvantage. The observed patterns reflect some of the well-known regional determinants influencing digitization. High levels of population density, professional qualification and internal immigration as well as a (relatively) young population are urban rather than rural characteristics. Consequently, by showing the relevance of these determinants for digitization, our findings suggest a deep-rooted second-level digital divide favoring urban regions. The insignificance of GDP is a notable exception and opposes previous findings. Perhaps this can be interpreted as an encouraging finding because it shows that digitization measures do not depend on wealth or affordability issues. However, it must be taken into account that webpages are a very basic digital application that has already widely diffused. More research should be conducted on the diffusion of more advanced applications like cloud computing or big data and the influence of regional GDP on the diffusion.

While this general pattern of increasing regional disparities and technological development is not new, we suggest that the disruptive effects of digitization with substantial changes to market structures change the picture. They are not limited to specific sectors that undergo technological change, but rather affect all sectors and firms alike, which means that their impact on regional economic structures is likely to be substantially more grave than previous technological changes. Thus, regional disparities between rural and (semi-)urban regions might become substantially more pronounced if rural firms fail to be included into the technological dynamics of digitization. Since market pressure from the respective demographics only facilitates digitization in specific regions, we argue that there is a growing demand for mechanisms of technology transfer to firms in rural regions.

The findings regarding both levels of the digital divide raise several policy implications. Once again, our results highlight the need to promote the expansion of broadband internet. Some rural regions are still underserved, whereby firms in these regions could implement more digitization measures if they had a better access to the internet. However, it needs to be emphasized that the rolling out of broadband would hardly eliminate the second-level digital divide. In line with previous findings, our results show that other regional characteristics - notably professional qualification, population density, age structure and migration - play a major role in regional digitization. Socio-demographically 'weak' regions would still tend to benefit less from broadband internet than socio-demographically 'strong' regions. Differences in these characteristics are persistent and can hardly be changed. Given that these regional characteristics seem to determine the adoption of the technologically fairly simple webpage usage, an even stronger influence of these characteristics must be expected for more advanced technologies like big data analysis or cloud computing. In order to spread positive effects of digital technologies and prevent the rise in inter-regional disparities, policy measures must aim to diffuse more advanced technologies beyond the socio-demographically strong regions.

Digital training will presumably be indispensable for regions that lack some of the characteristics that 'automatically' drive regional digitization, especially since some studies have showed that (informal) life-long learning is even more important for the usage of digital technologies than professional qualification (Billon et al., 2016; Evangelista et al., 2014). Policy measures could thus promote the advancement of digital skills in socio-demographically 'weak' regions. Digital skills should be more easily obtainable than professional qualification and in many regions they should additionally be essential for technological diversification, potentially helping to develop new technological specialization patterns (Castellacci et al., 2019). Given the fact that we have used a novel indicator for digitization, the support found for previous studies demonstrates the strong robustness of the findings. We confirm that both levels of the digital divide are at work at the same time. Furthermore, we support the finding that there is a set of strong regional predictors for the second-level digital divide. In particular, socio-demographic variables are (again) found to be important determinants of regional digitization. Since differences in these variables are persistent, the digital divide will hardly vanish soon. However, contrary to previous studies, GDP is not significantly associated with our digitization indicator, possibly due to the wide diffusion of webpage usage. It would be interesting to ascertain whether GDP is a more reliable predictor of the usage of less-diffused technologies. Moreover, future studies could integrate some potentially-relevant regional characteristics that we have not included but that have been used in previous research, namely regional industry and sector structures, average firm sizes and institutional factors (Billon et al., 2016; Dengler et al. 2018; Moriset et al., 2012).

6. References

Akerman, A., Gaarder, I. & Mogstad, M. (2015). The skill complementarity of broadband internet. The Quarterly Journal of Economics, 130, 1781-1824. https://doi.org/10.1093/qje/qjv028.
Atasoy, H. (2013). The Effects of Broadband Internet Expansion on Labor Market Outcomes. Industrial & Labor Relations Review, 66(2), 315-345.
Autio, E., Nambisan, S., Thomas, L.D.W. & Wright, M. (2018). Digital affordances, spatial affordances, and the genesis of entrepreneurial ecosystems. Strategic Entrepreneurship Journal 12(1),72-95.
Berger, T. & Frey, C. B. (2016a). Structural Transformation in the OECD: Digitalisation, Deindustrialisation and the Future of Work. OECD Social, Employment and Migration Working Papers, No. 193. OECD Publishing: Paris.
Berger, T. & Frey, C. B. (2016b). Did the Computer Revolution shift the fortunes of U.S. cities? Technology shocks and the geography of new jobs. Regional Science and Urban Economics, 57, 38-45.
Bertschek, I., Briglauer, W., Hüschelrath, K., Kauf, B. & Niebel, T. (2015). The Economic Impacts of Broadband Internet: A survey. Review of Network Economics, 14(4), 201-227.
Billon, M., Lera-Lopez, F. & Marco, R. (2016). ICT use by households and firms in the EU: links and determinants from a multivariate perspective. Review of World Economics, 152, 629-654. https://doi.org/10.1007/s10290-016-0259-8.
Blank, G., Graham, M. & Calvino, C. (2018). Local Geographies of Digital Inequality. Social Science Computer Review, 36(1), 82-102.
BMI (Bundesministerium des Innern, für Bau und Heimat - Federal Ministry of the Interior, Building and Community; Ed.) (2019). Unser Plan für Deutschland - Gleichwertige Lebensverhältnisse überall. Retrieved from: https://www.bmi.bund.de/SharedDocs/ downloads/DE/publikationen/themen/heimat-integration/schlussfolgerungen-kom-gl.html [06.05.2020].
Bosma N. & Sternberg, R. (2014) Entrepreneurship as an urban event? Empirical evidence from European cities. Regional Studies 48(6),1016-1033.
Briglauer, W., Dürr, N. S., Flack, O. & Hüschelrath, K. (2019a). Does state aid for broadband deployment in rural areas close the digital and economic divide? Information Economics and Policy, 46, 68-85. https://doi.org/10.1016/j.infoecopol.2019.01.001.
Briglauer, W., Dürr, N. S., & Gugler, K. (2019b). A Retrospective Study on the Regional Benefits and Spillover Effects of High Speed Broadband Networks: Evidence from German Counties (ZEW - Centre for European Economic Research Discussion Paper No. 19-026). Retrieved from SSRN website: http://dx.doi.org/10.2139/ssrn.3423864.
Brydges, T. & Hracs, B. J. (2019). The locational choices and interregional mobilities of creative entrepreneurs within Canada's fashion system. Regional Studies 53(4), 517-527, doi:10.1080/00343404.2018.1478410.
Büchi, M., Just, N., & Latzer, M. (2016). Modeling the second-level digital divide: A five-country study of social differences in Internet use. New Media & Society, 18(11), 2703-2722. https://doi.org/10.1177/1461444815604154.
Cairncross, F. (1997). The Death of Distance. How the Communications Revolution Will Change Our Lives. Harvard Business School Press, Boston.
Cairncross, F. (2018). The Death of Distance - Has Distance Died? RSAI Newsletter. New Series 17, 8-9.
Camagni, R., & Capello, R. (2005). ICTS and territorial competitiveness in the era of internet. The Annals of Regional Science, 39, 421-438. https://doi.org/10.1007/s00168-005-0244-y.
Castellacci, F., Consoli, D. & Santoalha, A. (2019). The role of e-skills in technological diversification in European regions. Regional Studies, online first. https://doi.org/10.1080/00343404.2019.1681585.
Celbis, M. G. & de Crombrugghe, D. (2018). Internet infrastructure and regional convergence: Evidence from Turkey. Papers in Regional Science, 97(2), 387-409. doi:10.1111/pirs.12244.
Ciffolilli, A. & Muscio, A. (2018). Industry 4.0: national and regional comparative advantages in key enabling technologies. European Planning Studies, 26(12), 2323-2343, DOI:10.1080/09654313.2018.1529145.
Cumming, D. & Johan, S. (2010). The Differential Impact of the Internet on Spurring Regional Entrepreneurship. Entrepreneurship Theory and Practice, 3(5), 857-884.
Dahl, M.S., Sorenson, O. (2012). Home sweet home: Entrepreneurs' location choices and the performance of their ventures. Management Science 58(6),1059-1071.
De Falco, S. (2019): Digital and urban spaces: Oxymoron or binomial? Urban transformations in the digital era. Geography Compass.e12467. https://doi.org/10.1111/gec3.12467.
Dengler, K., Matthes, B. & Wydra-Sommagio, G. (2018). Regionale Branchen- und Berufsstrukturen prägen die Substituierbarkeitspotenziale. IAB-Kurzberichte 22/2018.
Dijkstra, L., Poelman, H., & Rodríguez-Pose, A. (2019). The geography of EU discontent. Regional Studies, forthcoming. https://doi.org/10.1080/00343404.2019.1654603.
EC (European Commission) (2020). Digital Scoreboard - Data & Indicators. Retrieved from https://digital-agenda-data.eu/.
Evangelista, R., Guerrieri, P. & Meliciani, V. (2014). The economic impact of digital technologies in Europe. Economics of Innovation and New Technology, 23(8), 802-824. https://doi.org/10.1080/10438599.2014.918438.
Fabritz, N. (2013). The Impact of Broadband on Economic Activity in Rural Areas: Evidence from German Municipalities. Ifo Working Paper No. 166.
Feldman, M.P. (2001). The Entrepreneurial Event Revisited: Firm Formation in a Regional Context. Industrial and Corporate Change 10(4), 861-891.
Feldman, M.P., Siegel, D.S. & Wright, M. (2019). New developments in innovation and entrepreneurial ecosystems. Industrial and Corporate Change 28(4), 817-826.
Galloway, L., Sanders, J. & Deakins, D. (2011). Rural small firms' use of the internet: From global to local. Journal of Rural Studies, 27, 254-262. https://doi.org/10.1016/j.jrurstud.2011.05.005.
Graham, M. (ed.) (2019). Digital economies at global margins. Cambridge, MA : MIT Press.
Grimes, S. (2003). The digital economy challenge facing peripheral rural areas. Progress in Human Geography, 27(2), 174-193. https://doi.org/10.1191/0309132503ph421oa.
Grimes, S. (2005). How well are Europe's rural businesses connected to the digital economy? European Planning Studies, 13(7), 1063-1081, DOI: 10.1080/09654310500242105.
Hasbi, M. (2020). Impact of very high-speed broadband on company creation and entrepreneurship: Empirical Evidence. Telecommunications Policy, 44(3). https://doi.org/10.1016/j.telpol.2019.101873.

Ivus, O. & Boland, M. (2015). The Employment and Wage Impact of Broadband Deployment in Canada. Canadian Journal of Economics, 48(5), 1803-1830.
Jung, J. & López-Bazo, E. (2020). On the regional impact of broadband on productivity: The case of Brazil. Telecommunications Policy, 44(1). https://doi.org/10.1016/j.telpol.2019.05.002.
Katz, R. L. & Koutroumpis, P. (2013). Measuring digitization: A growth and welfare multiplier. Technovation, 33(10-11), 314-319. https://doi.org/10.1016/j.technovation.2013.06.004.
Krone, M. & Dannenberg, P. (2018). Analysing the effects of information and communication technologies (ICT) on the integration of East African farmers in a value chain context. Zeitschrift für Wirtschaftsgeographie, 62(1), 65-81. https://doi.org/10.1515/zfw-2017-0029.
Li, R. & Shiu, A. (2012). Internet diffusion in China: A dynamic panel data analysis. Telecommunications Policy, 36(10-11), 872-887. https://doi.org/10.1016/j.telpol.2012.06.004
Mack, E. & Faggian, A. (2013). Productivity and Broadband: The Human Factor. International Regional Science Review, 36(3), 392-423.
Mack, E., Anselin, L. & Grubesic, T. H. (2011). The importance of broadband provision to knowledge intensive firm location. Regional Science Policy & Practice. 3(1), 17-35. https://doi.org/10.1111/j.1757-7802.2011.01026.x.
Malecki, E. (2002): The Economic Geography of the Internet's Infrastructure. Economic Geography, 399-424.
Malecki, E. J. (2003). Digital development in rural areas: potentials and pitfalls. Journal of Rural studies, 19, 201-214.
McCoy, D., Lyons, S., Morgenroth, E., Palcic, D. & Allen, L. (2018). The impact of broadband and other infrastructure on the location of new business establishments. Journal of Regional Science, 58(3), 509-534.
Mills, B.F. & Whitacre, B.E. (2003). Understanding the non-metropolitan-metropolitan digital divide. Growth and Change 34 (2), 219-243. https://doi.org/10.1111/1468-2257.00215.
Moriset, B., Brette, O., Chareire, O., Grasland, L., Richaud, A. & Thivant, E. (2012). The 'death of distance' fifteen years on. Information technology and knowledge-based service fims in rural areas. HAL archives-ouvertes, halshs-00728338.
Moriset, B. & Malecki, E.J. (2009). Organization versus Space: The Paradoxical Geographies of the Digital Economy. Geography Compass 3(1), 256-274, DOI 10.1111/j.1749-8198.2008.00203.x.
Müller, S. & Korsgaard, S. (2018). Resources and bridging: the role of spatial context in rural entrepreneurship. Entrepreneurship and Regional Development, 30(1 -2), 224-255.
Nambisan, S., Wright, M. & Feldman, M. (2019). The digital transformation of innovation and entrepreneurship: Progress, challenges and key themes. Research Policy 48(8), art.no.103773.
Prieger, J. E. (2013). The broadband digit al divide and the economic benefits of mobile broadband for rural areas. Telecommunications Policy, 37(6-7), 483-502.
Rodríguez-Pose, A. (2018). The revenge of the places that don't matter (and what to do about it). Cambridge Journal of Regions, Economy and Society, 11, 189-209. https://doi.org/10.1093/cjres/rsx024.
Ruiz-Rodríguez, F., Lucendo-Monedero, A. L., & González-Relaño, R. (2017). Measurement and characterisation of the Digital Divide of Spanish regions at enterprise level. A comparative analysis with the European context. Telecommunications Policy, 42, 187-211. https://doi.org/10.1016/j.telpol.2017.11.007.
Salemink, K., Strijker, D. & Bosworth, G. (2017). Rural development in the digital age: A systematic literature review on unequal ICT availability, adoption, and use in rural areas. Journal of Rural Studies, 54, 360-371. https://doi.org/10.1016/j.jrurstud.2015.09.001.
Scheerder, A., von Deursen, A. & van Dijk, J. (2017). Determinants of Internet skills, uses and outcomes. A systematic review of the second- and third-level digital divide. Telematics and Informatics, 34(8), 1607-1624. https://doi.org/10.1016/j.tele.2017.07.007.
Schleife, K. (2010). What really matters: Regional versus individual determinants of the digital divide in Germany. Research Policy, 39, 173-185.
Stam, E. (2015). Entrepreneurial Ecosystems and Regional Policy: A Sympathetic Critique. European Planning Studies 23(9), 1759-1769.
Sternberg, R., von Bloh, J. & Coduras, A. (2019). A new framework to measure entrepreneurial ecosystems at the regional level. Zeitschrift für Wirtschaftsgeographie (=German Journal of Economic Geography). https://doi.org/10.1515/zfw-2018-0014.
The Federal Government of Germany (2018). Shaping Digitalization. Implementation startegy of the Federal Government. Berlin.
Townsend, L., Sathiaseelan, A., Fairhurst, G., & Wallace, C. (2013). Enhanced broadband access as a solution to the social and economic problems of the rural digital divide. Local Economy, 28, 580-595. https://doi.org/10.1177%2F0269094213496974.
Townsend, L., Wallace, C., Fairhurst, G. & Anderson, A. (2017). Broadband and the creative industries in rural Scotland. Journal of Rural Studies, 54, 451-458.
Tranos, E. (2012). The Causal Effect of the Internet Infrastructure on the Economic Development of European City Regions. Spatial Economic Analysis, 7(3), 319-337.
Tranos, E. (2016). The internet: its geography, growth and the creation of (digital) social capital. In: Shearmur, R., Carrincazeaux, C. , Doloreux, D. (Hrsg.), Handbook on the Geographies of Innovation, 356-369. Cheltenham ,UK: Edward Elgar.
Vicente, M. R., & López, A. J. (2011). Assessing the regional digital divide across the European Union-27. Telecommunications Policy, 35(3), 220-237. doi:10.1016/j.telpol.2010.12.013.
What works Centre for local economic growth (WWG) (2015). Evidence Review 6 - Broadband.
Whitacre, B., Gallardo, R., & Strover, S. (2014a). Does rural broadband impact jobs and income? Evidence from spatial and first-differenced regressions. The Annals of Regional Science, 53, 649-670. https://doi.org/10.1007/s00168-014-0637-x.
Whitacre, B., Gallardo, R., & Strover, S. (2014b) Broadband's Contribution to Economic Growth in Rural Areas: Moving Towards a Causal Relationship. Telecommunications Policy, 38, 1011 -1023. https://doi.org/10.1016/j.telpol.2014.05.005.
Zhang, F. & Li, D. (2018). Regional ICT access and entrepreneurship: Evidence from China. Information & Management, 55(2), 188-198.
Zhu, K., Kraemer, K. & Xu, S. (2003). Electronic business adoption by European firms: a cross-country assessment of the facilitators and inhibitors. European Journal of Information Systems, 12(3), 251-268. https://doi.org/10.1057/palgrave.ejis.3000475.

7. Appendix

7.1 Professions and number of observations

Table A1. List of professions and number of observations

Professions	Number of firms	Number of firms with homepages	Share of firms with homepages
Apparatus construction	1,383	465	.336
Optician	11,792	5,451	.46.2
Concrete building	6,086	2,724	.44.8
Well construction	857	376	.43.9
Bookbinder	972	349	.35.9
Baker	23,219	4,486	.19.3
Roofer	11,957	4,819	.40.3
Electrician	19,621	5,532	.28.2
Butcher	14,806	2,827	.19.1
Floor tiler	9,629	2,384	.24.8
Hairdresser	25,486	4,703	.18.5
Building cleaner	8,771	3,520	.40.1
Scaffolder	3,117	990	.31.8
Glazier	6,471	3,097	.47.9
Goldsmith	3,295	1,006	.30.5
Engraver	738	229	.31.0
Heating engineer	19,237	6,171	.32.1
Audiologist	5,485	2,549	.46.5
Installer	4,931	2,523	.51.2
Automotive mechatronics engineer	16,528	4,921	.29.8
Plumber	13,364	4,099	.30.7
Confectionist	1,669	453	.27.1
Furrier	261	65	.24.9
Painter	14,696	5,086	.34.6
Mason	2,227	877	.39.4
Custom tailor	159	82	.51.6
Organ maker	272	63	.23.2
Orthopedic technician	4,858	1,833	.37.7
Interior decorator	8,437	2,487	.29.5
Plumbing Electrician	1,586	793	.50.0
Heating engineering	20,138	5,701	.28.3
Plumbing and heating installer	13,019	4,299	.33.0
Smith	5,697	1,398	.24.5
Chimney sweeper	4,803	609	.12.7
Shoemaker	2,009	367	.18.3
Stone carver	4,473	1,548	.34.6
Roadmaking	6,948	1,714	.24.7
Plasterer	4,180	1,115	.26.7
Dry-cleaner	3,427	758	.22.1
Cabinetmaker	17,135	5,189	.30.3
Ceramist	437	101	.23.1
Clocksmith	1,557	349	.22.4
Carpenter	9,937	2,842	.28.6
Bicycle mechanic	10,691	3,510	.32.8
No. of observations / share	**346,361**	**104,460**	**.30.2**

7.2 List of variables

Table A2. List of and description of variables

Name	Description	Calculation	Survey year
Population density	Inhabitants per km²	Inhabitants / km²	2015
Share of employees without professional qualification	Share of employees without professional qualification of all employees at place of residence in %	Employees without professional qualification at place of residents / all employees at place of residence x 100	2015
Share of employees with academic education (university or applied college)	Share of employees with academic education of all employees at place of residence in %	Employees with academic education at place of residence / all employees at place of residence x 100	2015
Turnover in craft sector	Turnover of craft companies in 1,000 € per employee	Turnover in craft sector / number of employees	2014
GDP per inhabitant	GDP in 1,000 € per inhabitant	GDP / number of inhabitants / 1,000	2015
Rate of (internal) emigrants per region	Emigrants per 1,000 inhabitants in a given region	Emigration / number of inhabitants x 1,000	2015
Rate of internal immigration per region	Immigrants per 1,000 inhabitants in a given region	Internal immigrants / number of inhabitants x 1,000	2015
Development of inhabitants aged 65+	Development of number of inhabitants 65 years or older in %	(inhabitants >= 65 years <2015> - inhabitants >= 65 years <2011>) / inhabitants >= 65 years <2011> x 100	2011-2015
Employees in craft sector overall	Share of employees working in crafts companies of all employees in %	Employees in crafts sector / all employees x 100	2014

Source of data: INKAR Database (www.inkar.de; last accessed 03.02.2020)