Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Population grid-based assessment of the impact of broadband expansion on population development in rural areas

Olli Lehtonen\textsuperscript{a,}\textsuperscript{b}

\textsuperscript{a} Natural Resources Institute Finland, Latokartanonkaari 9, 00790, Helsinki, Finland
\textsuperscript{b} University of Eastern Finland, Department of Historical and Geographical Studies, P.O. Box 111, 80101, Joensuu, Finland

ARTICLE INFO

Keywords:
Broadband
Population development
Rural areas
Regression modelling
GIS
Finland

ABSTRACT

This paper estimates the effect of the expansion of broadband infrastructure, which enables high-speed Internet, on population development in panel of Finnish areas in the period 2010–2018. The study differs from previous studies in that it uses accurate statistics on the availability of broadband in 1 km \(^2\) population grids. Therefore, the impact of broadband availability on rural development is evaluated more accurately than previously. The results of the Difference-in-Difference (DiD) regression analysis show that the availability of broadband reduces depopulation of remote and sparsely populated rural areas. In this respect, the telecommunication policy in Finland has been successful, and the findings encourage the expansion of broadband infrastructure in rural areas.

1. Introduction

The expansion of broadband and the development of high-speed networks are for the most part focused on urban areas where the potential profits are higher for private-sector firms. The uneven development of the availability of broadband creates a digital divide where rural areas run the risk of being left behind in the use of information and communication technologies (Dickes et al., 2010). With broadband, it is possible to secure and increase housing opportunities and services in rural areas. The high-speed broadband connections will, for instance, improve opportunities for rural residents and entrepreneurs to use e-services, telemedicine and telecommuting (Townsend et al., 2013). The broadband connections can also potentially reduce the locational disadvantage of the distant rural areas (Lehtonen & Tykkyläinen, 2010) by improving their relational connections to growth areas and offering market access. Thus, the expansion of broadband for rural areas is a key element ensuring the regional competitiveness of the rural areas.

From a regional equality perspective, there are justifiable reasons for public support of expansion projects in broadband construction as regional differences in the availability of broadband have remained wide. Research has shown that broadband is most often missing in remote and sparsely populated rural areas, which are not attractive areas for commercial operators (Koutsouris, 2010; Townsend et al., 2013; Pyykönen & Lehtonen, 2016). The digital divide is a result of the higher construction cost of broadband in rural areas per customer than in urban areas, which is why the potential customer base is lower in rural areas and makes it even more difficult to anticipate rural broadband expansion. In the rural areas, the lack of broadband exacerbates the digital divide because it is often in these areas that the benefits of broadband could be the greatest since they are located far away from the physical services. Therefore, digital services could increase the opportunities to use services regardless of time and place in these otherwise service-poor environments (Pyykönen & Lehtonen, 2016).

E-mail address: olli.lehtonen@luke.fi.

https://doi.org/10.1016/j.telpol.2020.102028
Received 16 March 2020; Received in revised form 29 June 2020; Accepted 5 August 2020
Available online 25 August 2020
0308-5961/© 2020 Elsevier Ltd. All rights reserved.
In this study, GIS database is used to examine how the availability of broadband in Finland is related to the population development of the rural areas in the period 2010–2018. The aim is to increase the understanding of the impact of broadband availability on rural depopulation and thereby generally evaluate the use of public funding regarding the expansion of broadband in sparsely populated areas and overall telecommunication policy in Finland. The assumed chain of the impact of broadband expansion stems from public funding, which makes it possible to expand broadband and improve the availability of broadband in rural areas, thus boosting local economies and labour markets and finally leading to a more positive population development in the rural areas. However, the theory and empirical evidence on the effect of broadband expansion in rural areas are far less certain.

The study differs from previous studies in that it uses an accurate GIS database on the availability of broadband, and therefore, the impact of broadband availability on rural population development can be evaluated much more accurately than previously. In the past, for example, the availability of broadband has been studied at the municipal or county level (e.g., Czernich, 2014; Briglauer ym. 2018; Hasbi, 2020; Whitacre et al., 2014), but it is likely that local differences in broadband availability within municipalities or counties distort the research results. Despite the abundant research, there is no consensus about the impacts of broadband expansion on the development of rural areas. For instance, some studies have not found any evidence that broadband construction would have positive effects on regional development (e.g., Czernich et al., 2014, 2011; Whitacre et al., 2014) while other studies underlined positive effects on protecting depopulation (e.g. Briglauer et al., 2018; Townsend et al., 2013) or enhancing attractiveness for the creation of new businesses (e.g. Hasbi, 2020; McCoy et al., 2018). The varying results can partly be explained by the lack of accurate and comprehensive GIS databases and the diversity of the research areas. Due to varying research results, understanding of the importance of broadband availability for regional development is incomplete (Whitacre et al., 2014).

2. Importance of broadband for areas

The fast broadband networks were first available in the most densely populated areas, and only later did they spread into the more sparsely populated rural areas. The more distant areas can be tied to older technology, slower data and sometimes also to higher prices if local operators are small and struggling with costs. In the rural areas, people tend to use lower speeds and higher cost connections for technical reasons (Townsend et al., 2013). Koutsouris (2010) notices that the investments in rural areas worldwide remain behind those made in urban areas, and the last mile connectivity stays poor, which has been one of the arguments for public support. In addition to the availability of broadband, the benefits of connecting to households and businesses when the connection is available in the area are important. The deployment of broadband is not taking place at the same pace as its availability. The use of broadband appears to be influenced by the educational level of the population, household income, age, previous experience with the use of information technology and the construction cost of broadband (Carare et al., 2015; Whitacre et al., 2014; LaRose, Strover, Gregg, & Straubhaar, 2011).

From a housing perspective, the rural municipalities with broadband connections may be more pleasant places to live and work than those without them (Townsend et al., 2013). Enhanced broadband provision might also simply shift labour from metropolitan areas to more rural areas because broadband availability is assumed to increase telecommuting (Autor, 2001). In this way broadband expansion protects rural areas from depopulation but based on empirical findings it does not contribute to a further closing of the economic divide in the form of creating new jobs (Briglauer et al., 2018). In rural areas, broadband connections are essential for housing because digitalization improves the availability and accessibility of services in a diminishing rural physical service network. Therefore, broadband connections are a crucial part of the basic infrastructure comparable to roads and electricity, without which it is difficult to cope (Salemink et al., 2017; Skerratt et al., 2012). In the policy literature, there is a consensus that extensive broadband provision, applications and use are crucial factors for the future of employment and wealth creation in all nations and regions of the (especially developed) world (Preston et al., 2007).

The advancement of digitalization means that many everyday activities, such as bank services or shopping, have gone online. In addition, some social and health services, education, business and information services are also increasingly moving toward online digital services, highlighting the importance of broadband for a sense of community, telework opportunities, as well as possibilities for distance learning and telemedicine (Stenberg et al., 2009). The full use of digital services in rural areas requires functional, secure and fast broadband connections, which wireless connections cannot always guarantee (Townsend et al., 2013). The improving communication leads to, among other things, increased contacts with wider social networks, which are particularly important in rural areas for generating social capital and improving the development of the communities (Stern & Adams, 2010). Connections are also an important part of the multi-local living telework is projected to increase. In the simplest definition, the multi-local living means that a person or family have more than one residence or place to stay.

Several studies have demonstrated that fast broadband has a positive impact on employment as well as other issues of life quality (health, education, social relations) (Stenberg et al., 2009; Dickes et al., 2010). Rural studies have also shown a positive link between broadband availability and regional economic development (Holt & Jamieson, 2009; Koutroumpis, 2009). Thus, the construction of broadband connections can have positive effects later on, for example, population development, employment, income and job creation (Priege, 2013; Forzati et al., 2012; Atasoy, 2013; Mölleryd, 2015). The construction of broadband, particularly in rural areas, is seen as important as it creates opportunities for economic growth and business development in the area. For example, a study from Sweden found that the expansion of broadband had a positive impact on employment and population development of the regions (Forzati et al., 2012). According to the analysis of Stenberg et al., (2009), employment has grown faster in the areas with better Internet broadband. However, Galloway (2007) concludes that broadband access alone will not change the trend of rural economic development, and funding and promotion of it in isolation will be wasted if there is no development of firm motivations and skills for business. Similarly, a study relates advanced Internet technology to significant employment growth, but only in areas that already had a large and highly
skilled population, high income, and IT usage before the expansion of broadband (Forman et al., 2012).

The literature demonstrates that fast broadband enables business continuity and business development, but it also has an impact on business start-ups and locations. Municipalities with a very high-speed broadband network tend to be more attractive for companies, with a positive effect on establishment creation within the tertiary sector and the construction sector (Hasbi, 2020). Another study in Finland showed that one out of four companies reported that broadband connections had an impact on the location decisions of the companies (Kurvinen et al., 2018). However, other studies suggest that in general, broadband has a positive impact on productivity only in locales with high levels of human capital and/or highly skilled occupations (Mack & Faggian, 2013; McCoy et al., 2018). Companies need high-speed communications to access services and communicate with their customers and stakeholders (Barkley et al., 2007; Mack, 2014). In general, the effects of high-speed Internet connections are based on the cost savings and increasing business efficiency (Dickes et al., 2010; Stenberg et al., 2009). For instance, the entrepreneurs operating in rural areas can reduce marketing costs, increase their sales and reach new market areas with high-speed Internet connections (Prieger, 2013). In this sense, broadband technology reduces the importance of agglomeration advantages to firms and citizens and leads to more dispersion in the spatial distribution of economic activity (Ioannides et al., 2008).

3. Broadband policy and availability in Finland

The gap between urban and rural areas in terms of broadband access has increased as the expansion has progressed. The Finnish government recognized the deepening digital divide, with rural areas running the risk of being left behind in the information society. Thus, the Fast Broadband project was launched in December 2008 by the Finnish government aiming to ensure, with public support, the construction of high-speed broadband networks in areas where their commercial availability would be unlikely. The original goal of the Fast Broadband project was that by the end of 2015, more than 99% of the users would have access to a 100 Mbp/s broadband connection within 2 km of their permanent place of residence or place of business. In this paper, the definition for broadband refers to a connection at least 100 Mbp/s.

In order to achieve the ambitious goal, in 2009, the regional councils planned the regional programmes of projects for building

![Fig. 1. Availability of broadband in population grids in 2018 and location of Fast Broadband project areas in Finland.](image-url)
broadband infrastructure. In total, the programmes included some 800 projects and their combined costs were estimated to be nearly EUR 500 million. The projects included plans for expanding the broadband network system by 40,000 km, enabling the provision of advanced communications services to approximately 130,000 users in sparsely populated areas. A total of EUR 130 million of public support was available for broadband projects. Of the total funding, EUR 66 million is government support, approximately EUR 25 million is from the EU Rural Development Programme for Mainland Finland and approximately EUR 40 million from Finnish municipalities. The support is targeted at improving the availability of connections in the geographically remotest areas.

Geographically, the availability of broadband in the population grids in Finland was dispersed in 2018 (Fig. 1). In total, broadband was available in 27,191 population grids, which corresponds to 27.2% of inhabitant population grids. The areas with broadband availability appear to be vague and dispersed, located in different parts of the country without a distinct “spatial trend”. The availability seems to be most pronounced in urban areas and settlements that follow the road or river. The fragmented map of broadband availability is explained by the fact that broadband construction has progressed in a project-like manner and is driven by local initiative and activity. Thus, it seems that based on the availability map, local activity is sufficient to cover only part of the country with broadband but it also indicated that fragmentation may be influenced by the inactivity of some municipalities in supporting local broadband development and the institutional and cultural features of these types of areas, which do not favor broadband construction based on local project activities such as village activities.

The urban-rural typology is used to describe in more detail the expansion of broadband in Finland. The typology divides Finnish territory into 7 categories: 3 urban and 4 rural categories. The typology is based on the population, labour, commute, building and land use datasets which are used to calculate variables describing the amount, density, efficiency, accessibility, intensity, versatility and orientation of the areas (Helminen et al., 2014; see Data and Methods for the description of the YKR georeferenced statistical database). Since 2010, urban and exurban areas have been characterized by population increases (Table 1). Parallel to the population growth in urban and exurban areas, the four categories of rural areas in YKR typology have been mostly characterized by a loss of population. Rural areas close to urban areas witnessed the most favourable population development with only small depopulation during the years 2010–2018. Local centres in rural areas outside of the reach of urban influence have also had relatively stable population numbers (Table 1). The remaining two rural categories – core rural areas and sparsely populated rural areas, which together cover 84.2 percent of the Finnish land area – have experienced significant losses of population. Sparsely populated rural areas with the lowest population density lost 10.7% of their population during the years 2010–2018 (Table 1).

Broadband availability differs between urban and rural areas. The objective of the broadband project in Finland is realized only in the inner urban areas which belong to commercial areas (Table 1). In the outer urban areas, the availability of broadband is close to the target level, but other urban and rural areas remain clearly below the 99% target level. In the local centres of the rural areas, the availability is high in terms of population, but in the surrounding rural areas, the availability is much worse. The lowest availability of broadband is found in rural areas close to urban areas and in sparsely populated rural areas (Table 1) because less than 50% of the populations in these areas has broadband available. This structure will be a challenge for broadband construction in the future because it seems that most profitable and active rural areas, such as local centres and areas around them, have already been built and might decrease the probability of investments for the least profitable rural areas.

4. Data and methods

4.1. Construction of GIS database for broadband availability

The GIS database used for this study is compiled from Statistics Finland’s 1 km × 1 km population grids from the years 2010–2018 (Statistics Finland, 2019). This database contains total population and age and gender distributions of each inhabited grid, and it was used to calculate the annual population changes for the areas with and without broadband available inside the municipalities. The YKR

Table 1
Availability of broadband in urban-rural typology in year 2018 (%). Availability refers to the percentage of the population that lives in a population grid that contains broadband.

| Urban-rural typology | Population grids | Population | Population density in inhabited grids 2018 pers/km² | Population growth 2010–2018 (%) |
|----------------------|-----------------|------------|-----------------------------------------------|-------------------------------|
|                      | No broadband available | No broadband available | No broadband available | No broadband available |                       |                          |                          |
| Inner urban area     | 18              | 2.5        | 703                             | 97.5                          | 8360                    | 0.5                     | 18,32,874               | 99.5                      | 2553                     | 8.2                       |
| Outer urban area     | 284             | 13.5       | 1812                            | 86.5                          | 33,845                  | 2.4                     | 13,89,569               | 97.6                      | 679                       | 4.6                       |
| Exurban fringe       | 7535            | 23.6       | 2329                            | 76.4                          | 2,30,070                | 37.1                    | 3,89,530                | 62.9                      | 62                        | 6.3                       |
| Local centres of rural areas | 233       | 30.8       | 523                             | 69.2                          | 50,059                  | 17.4                    | 2,38,298                | 82.6                      | 381                       | –3.2                      |
| Rural areas close to urban areas | 16,632 | 84.0       | 3159                            | 16.0                          | 2,33,746                | 60.0                    | 1,56,096                | 40.0                      | 20                        | –2.4                      |
| Core rural areas     | 19,745          | 65.9       | 10,199                           | 34.1                          | 2,4,3707                | 40.1                    | 3,63,631                | 59.9                      | 20                        | –7.1                      |
| Sparsely populated rural areas | 26,205 | 73.1       | 9623                            | 26.9                          | 1,50,136                | 54.1                    | 1,27,486                | 45.9                      | 8                         | –10.7                     |
| Total                | 70,652          | 71.4       | 28,348                           | 28.6                          | 9,49,923                | 17.4                    | 44,97,492               | 82.6                      | 55                        | 0.3                       |
behind the project areas is most often a small company or a cooperative founded by local activity and operating outside the commercial payment decisions was combined with the geographical boundaries of broadband project areas (Traficom, 2019c). The operator instance, in planning public services (Kotavaara et al., 2011). Performing spatial population analysis independent from administrative analyses independent of administrative boundaries, which is a reason why their use has become more popular in recent years, for broadband availability can be very accurately subdivided into treated and non-treated population grids according to broadband availability within broadband project areas.

The GIS database was supplemented with the data of broadband availability published by Traficom at the same resolution of 1 km * 1 km rather than population grids (Traficom, 2019b). This data was supplemented by information on the year of construction of the broadband network in the areas, which was derived from the payment decisions of Traficom (2019c). The year of the construction of broadband is an essential element in assessing its impact on population development because the effects of broadband construction will be delayed by a few years (Stenberg et al., 2009). The data included 177 project areas, of which information from 119 project area payment decisions was combined with the geographical boundaries of broadband project areas (Traficom, 2019c). The operator behind the project areas is most often a small company or a cooperative founded by local activity and operating outside the commercial telecommunication areas. The funding can be only granted to the operator which has the financial capacity to carry out the construction project. The project areas are mapped on Fig. 1.

According to the GIS database, broadband availability improved for a total of 33,717 inhabitants in the period 2012–2019. The construction project activities are mainly located in sparsely populated rural areas, as 65% of the population for whom broadband has been made available live in sparsely populated rural areas (Table 2). The second largest class in the typology is the core rural areas, which accounts for 25% of the improvement in broadband availability (Table 1). Due to the removal of commercial areas and densely populated areas from the database, there has been no change in the availability of broadband in the local centres of rural areas (Table 2). According to the GIS database, a total of EUR 43,126,460 in state aid has been paid for these project areas.

The statistical unit of the analysis was compiled by aggregating the data on the population grids by project area and subdividing the data based on broadband availability within the municipalities by the project areas because a municipality is an administrative unit of broadband funding. The aggregation was done because the smallest population grid contained only 1 inhabitant in the year 2018, and thus, it would be impossible to analyze the impact of broadband availability on the population development of a single 1 km * 1 km population grid. The commercial population grids where the construction of broadband is expected to take place by commercial broadband companies were excluded from the analysis.

4.2. Regression modelling in examining the effects of broadband availability

Previous studies have demonstrated that the impacts of the expansion of broadband on regional development are difficult to assess (Whitacre et al., 2014) because the impact of technological investments on regional development is due to the simultaneous influence of several factors at the same time (Ramírez & Richardson, 2005; Stenberg et al., 2009). Therefore, this study utilizes matching methods and advanced statistical modelling in an attempt to come as close as possible to a randomized trial design given that we are using empirical data and, thus, to improve the generalizability of research results (Austin, 2011). The purpose of matching is to find one (or more) non-treated unit(s) (e.g. areas without broadband availability) for every treated unit (e.g., areas where broadband is available) with similar observable characteristics against which the effect of the treatment can be assessed. The use of matching enables the impacts of the expansion of broadband to be interpreted more reliably within these matched groups because matching aims to reduce selection bias (Rosenbaum & Rubin, 1983).

In this study, the matching is used in combination with Difference-in-difference (DiD) regression modelling because the standard DiD estimation can be combined with matching techniques if covariates are not influenced by the treatment (Lechner, 2011). A significant shortcoming of common matching methods such as Mahalanobis distance and propensity score matching is that they can make covariate balance worse across measured potential confounders (Sekhon, 2011). This can happen if covariates do not have ellipsoidal distributions, e.g., distributions such as the normal or t. One potential matching method proposed for the situations when the propensity score is not known is genetic matching, which is also utilized in this paper. Diamond and Sekhon (2005) and Sekhon and

| Table 2 | Change in the broadband availability in the period 2012–2019 in urban-rural typology. The figure represents inhabitants for whom broadband has been made available in broadband projects. |
|---------|----------------------------------------------------------------------------------|
| Rural typology | Year | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 |
| Local centres of rural areas | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Rural areas close to urban areas | 127 | 229 | 89 | 1115 | 141 | 17 | 1037 | 203 | 2958 |
| Core rural areas | 867 | 165 | 368 | 1198 | 1498 | 483 | 2133 | 1832 | 8544 |
| Sparsely populated rural areas | 2003 | 2181 | 2396 | 1112 | 3614 | 2655 | 6584 | 1470 | 22215 |
| Total | 2997 | 2575 | 2853 | 3425 | 5453 | 3155 | 9754 | 3505 | 33717 |
Grieve (2011) present a matching algorithm – genetic matching (GenMatch) – that maximizes the balance of observed covariates between treated and control groups.

A wide array of literature has already assessed the factors that can be related to the availability of broadband (e.g. Koutsouris, 2010; Dikes et al., 2010; Hypponen et al., 2014; Whitacre et al., 2014). For the purposes of this study, the matching is based on the educational level (share of highly educated people), average age, median income, total number of jobs and shares of primary and service industrial jobs in the area. These variables are used to ensure that the comparison of the population development in the treatment and non-treatment areas is made between similar areas and their selection is based on the literature but also available datasets. However, it should be noted that there might be some unobserved variables that are not included in matching that are related on the population development of the areas. All the variables used in matching had a base year of 2016. The summary plot of covariance balance before and after conditioning confirms the balance of the matched database (Fig. 2).

DiD regression is used with annual panel data from the years 2010–2018 to evaluate the impact of broadband expansion on the population development of areas by comparing differences in the population development between built (so-called intervention) and non-built (so-called control) groups (Heckman et al., 1999). Panel regression is used in the modelling because the timing of the broadband availability has been a major factor for the effects (Stenberg et al., 2009). To analyze the impact of the broadband expansion on the population development of areas, the following mixed-effect model with a DiD coefficient and a random intercept is used:

$$
\Delta \text{Pop}_{it} = \beta_1 + \beta_2 \text{time}_{it} + \beta_3 \text{availability}_{it} + \delta (\text{time}_{it} \times \text{availability}_{it}) + b_i + e_{it},
$$

Where:

- $$\Delta \text{Pop}_{it}$$ denotes the annual population change of the area $$i$$ at time $$t$$ as a percentage of the previous year’s population;
- $$\text{time}_{it}$$ denotes broadband availability in the area $$i$$ at time $$t$$ (the variable is coded as a binary: 1 means the time after broadband became available in the area, and 0 means the time before broadband was available in the area), indicating the average population change after broadband construction referring to the effect of time without actual intervention (expansion of broadband);
- $$\text{availability}_{it}$$ denotes the average population difference between the intervention and control areas prior to the expansion of broadband (the variable is coded as binary: 1 means that the broadband is available in the area unit and 0 means that the broadband is not available in the area);
- $$(\text{time}_{it} \times \text{availability}_{it})$$ denotes an interaction term which reflects whether the population development before and after the construction of broadband was similar in the intervention and control groups (the variable is binary and number 1 means that broadband is available in the area at time $$t$$, and number 0 means that broadband is not available in the area at time $$t$$);
- $$b_i$$ denotes for the random effect of the project and non-project areas, and
- $$e_{it}$$ denotes the error term.

The greatest interest of the regression model is directed towards the regression coefficient $$\delta$$ and its sign. This coefficient is called the DiD coefficient, and it measures the interaction effect of the time from the construction of broadband to the availability of broadband, that is, the difference between areas in population development over time. If the sign of the parameter $$\delta$$ is negative, the broadband construction has exacerbated the negative population development of the area. In contrast, a positive sign of the parameter $$\delta$$ denotes that population development has been more positive in the area where the broadband was built compared with the area where it had not been built. In the model, the random effects and residuals are assumed to be independent and normally distributed.

![Covariate Balance](image-url)  
**Fig. 2.** Covariance balance plot before and after conditioning.
with zero means and variances $\tau^2$ and $\sigma$. More specifically, parameter $\tau^2$ denotes the unexplained variability in the project and non-project areas means of $\Delta \text{Pop}$ and the residual variance $\sigma$ describes the unexplained variability across project and non-project areas in the fixed effects on the model. (Pinheiro & Bates 2000) These parameters are needed to calculate R2 and ICC (intraclass correlation coefficient) for the estimated mixed model. The marginal R2 is the fixed effects variance indicating how much of the model variance is explained by the fixed effects part only. The conditional R2 describes how much of the "model variance" is explained by the fixed and random effects. The ICC describes how much of the proportion of variance can be explained by the random effects only.

The regression model is fitted as a mixed model, because in the panel data, the population development in consecutive years is often not independent of previous years, and therefore the traditional least squares method, OLS, mean errors and thus interpretations of the statistical significance of the explanatory variables can be misleading. The model is fitted as a random mixed model by using restricted maximum likelihood estimates with random intercepts for areas because the population development in the project areas can be heterogeneous as the net migration and differences in natural population change, for instance, can have an impact on population development. In addition, the use of the random model can be justified because the selection of the broadband project areas can be considered random due to the availability of the information from the project areas varying a lot. Results of the Hausman Test supported the use of the random model instead of the fixed model (chisq 0.017, p-value 0.895). During the modeling it was considered using more advanced mixed-effects models with random slopes, but the model with a random intercept was deemed to be sufficient. The mixed model was estimated by using the lmer function from the lme4 package of the R program (Bates et al., 2014). The mixed model regression technique is explained in more detail in, for example, Venables and Ripley (2002). The models were fitted to the dataset with and without matching.

5. Results

5.1. Impact of the availability of broadband on population development

The population trends in the urban and rural areas seem to be linked to the availability of broadband. Table 3 summarizes the population trends in the years 2011–2018 by urban-rural typology and broadband availability. The findings show that in areas where broadband had become available by 2018, the population trends in the years 2011–2018 were much more positive than in areas where broadband was not available. In statistical significance, the only exceptions are the inner urban area and the sparsely populated rural area where no statistical difference can be observed between the categories of broadband availability (Table 3). However, in both cases, the population development in non-broadband areas has been negative. Another interesting finding in Table 3 is that the population growth has occurred in those areas where broadband was built and available. These growth areas are located in the inner and outer urban areas and in the exurban fringe, indicating ongoing urbanization processes. These findings suggest that the availability of broadband seems to have effects on population development of the areas, but the impacts of broadband expansion cannot be demonstrated by these figures.

The impact of the availability of broadband on population development is analysed with DiD regression. The main interest in the results of the regression modelling is directed towards the DiD coefficient which measures the combined effect of the broadband construction year and availability. The DiD coefficient shows whether the expected mean change in population development from before to after broadband expansion was different in the broadband and non-broadband areas. Before the construction of broadband, the annual population development trends in the broadband and non-broadband areas were similar (broadband areas $\Delta \text{Pop} = 0.081$, t-value 0.032, p-value 0.974). The results of the DiD regression models differ in terms of the estimation technique and used data (Table 4). The most reliable interpretations of the impact of the expansion of broadband on the annual population development of the areas are given in Table 4 from the results of the mixed model estimated for matched data.

In the mixed model estimated for matched data, the DiD coefficient is statistically significant and, in addition, has a positive sign (Table 4). This result indicates that the population development has been more positive in the areas with broadband availability after the construction compared with areas where broadband has not been built. In other words, without broadband construction, the depopulation would have been stronger, and the areas would have lost more population as their local competitiveness would remained lower than those with broadband availability. This interpretation is confirmed from the regression coefficient of the time variable

| Urban-rural typology | Population grids | Population grids | t-test |
|----------------------|------------------|------------------|-------|
|                      | No broadband available | Broadband available |       |
|                      | n | % | n | % | t-value | p-value |
| Inner urban area     | -84 | -1.0 | 125001 | 6.8 | -1.709 | 0.105 |
| Outer urban area     | -1331 | -3.9 | 54003 | 3.9 | -4.259 | <0.001 |
| Exurban fringe       | -5555 | -2.4 | 36572 | 9.4 | -9.513 | <0.001 |
| Local centres of rural areas | -3580 | -7.2 | -5792 | -2.4 | -4.068 | <0.001 |
| Rural areas close to urban areas | -9538 | -4.1 | -586 | -0.4 | -4.558 | <0.001 |
| Core rural areas     | -20787 | -8.5 | -21685 | -6.0 | -6.818 | <0.001 |
| Sparsely populated rural areas | -15831 | -10.5 | -13505 | -10.6 | 1.761 | 0.078 |
| Total                | -56706 | -2.2 | 174008 | 3.9 | -8.805 | <0.001 |
Table 4
The results of the DiD regression modelling fitted to data that excluded densely populated areas and commercial areas. Statistically significant regression coefficients are shown in bold in the table.

| Predictors | OLS Estimates | CI | p | Mixed model Estimates | CI | p | OLS, matched data Estimates | CI | p | Mixed model, matched data Estimates | CI | p |
|------------|---------------|----|---|------------------------|----|---|---------------------------|----|---|-------------------------------|----|---|
| (Intercept) | -0.40 | -0.83-0.03 | 0.066 | 0.34 | -0.36-1.03 | 0.342 | -0.42 | -0.85-0.02 | 0.060 | 0.26 | -0.49-1.01 | 0.497 |
| Treated     | 0.21  | -0.74-1.16 | 0.667 | -0.53 | -1.60-0.55 | 0.338 | 0.23 | -0.74-1.20 | 0.644 | -0.44 | -1.61-0.72 | 0.456 |
| Time        | -1.05 | -2.26-0.16 | 0.088 | -3.01 | -5.39-0.62 | 0.014 | -1.10 | -2.32-0.11 | 0.076 | -3.58 | -6.01-1.15 | 0.004 |
| DiD         | 0.50  | -1.36-2.37 | 0.597 | 2.46 | -0.30-5.22 | 0.081 | 0.54 | -1.34-2.41 | 0.573 | 2.96 | 0.15-5.76 | 0.039 |
| Random Effects |       |        |    |                       |    |    |                           |    |    |                               |    |    |
| $\sigma^2$   |        |        |    |                       |    |    |                           |    |    |                               |    |    |
| ICC          | 0.00  | 0.00   | 0.02 |                       |    |    |                           |    |    |                               |    |    |
| N            | 312   | 312    | 164 |                       |    |    |                           |    |    |                               |    |    |
| Observations | 3634  | 1854   |     |                       |    |    |                           |    |    |                               |    |    |
| Marginal R²/Conditional R² | 0.001/0.000 | 0.004/0.003 | 0.001/0.005 | 0.005/0.023 |
Table 5
Estimated effects of broadband construction on the population development of the project areas. Calculation is based on the mixed model fitted on matched data. T-test analyses the difference in the average population development of the project areas.

| Variable                     | Observed population development | Results of counterfactual analyses |                      |                      |
|------------------------------|---------------------------------|-----------------------------------|----------------------|----------------------|
|                              | Areas without broadband available | Areas with broadband available     | If the broadband had not been built | If broadband had already been built in 2010 |
|                              | Estimated population development | Diffe-rence in observed development | t-test (p-value)    | Estimated population development | Diffe-rence in observed development | t-test (p-value) |
| Population in 2018 (n)       | 293760                          | 71994                             | 66015                | 74549                |
| Change in the population 2010–2018 | −18301                         | −3313                             | −5979                | 2555                |
| %                            | −6.3                            | −4.6                              | −2666                | 5868                |
| annual change %              | −0.8                            | −0.6                              | −3.7                 | 5.6                 |
|                              | −0.4                            | −0.4                              | (<0.001)             | 0.7                 |

O. Lehtonen
which is significant and negative, indicating that after the expansion of broadband, the areas without broadband in the control group lost their populations (Table 4). The coefficient of the treatment variable is not significant, which means that there are no significant differences in the estimated mean difference in population development between the treatment and control groups prior to the broadband expansion (Table 4).

5.2. Counterfactual analyses based on DiD regression

The positive effects of broadband construction on the development of populations in the areas are further analysed and illustrated in Table 5 which depicts the population trends in rural areas by categories of broadband availability. There is a clear demographic difference between categories because the areas with broadband availability have lost 5.7% of their population in the period 2010–2018, and the areas without broadband availability have lost 7.9% of their population (Table 5).

In addition, Table 5 contains two counterfactual analyses which are based on a DiD-regression model estimated with the random effect model on the matched database. The analyses concretize the results of the DiD regression as broadband construction reduced population losses in the areas with broadband availability. Without the broadband construction projects, the population of these areas would have been reduced by 2666 inhabitants more in relation to the observed population development between 2010 and 2018 which would have increased annual depopulation by 0.4% per year (Table 5). The result from T-test also demonstrates that on average the construction of the broadband had an impact on the population development of the project areas (Table 5). Without construction the annual depopulation of the project areas would have been on average 2.6% while the observed depopulation was on average 0.1% in project areas.

The positive effects of broadband expansion on rural development will also be concretized in the second counterfactual analysis. If broadband had already been built in 2010 for the broadband project areas, the population development in the years 2010–2018 would have turned positive by 5868 inhabitants in the rural areas where broadband was constructed later. Annually, the counterfactual analysis shows that the population of the broadband project areas would have increased by 0.1% per year (Table 5). In addition to the positive population effects of broadband expansion, this finding emphasizes that the effects of broadband construction are delayed in the rural areas, meaning that the effects of construction will be more evident in the population trends of the areas after two to three years. This is indicated by T-test result which is not showing a significant change in the average population development even though the average change in population in project areas turns to be positive by 0.3%.

6. Conclusions

The findings of the study suggest that the availability of broadband is positively linked to the population development of the sparsely populated rural areas. According to what-if analysis based on the regression modelling, the depopulation in the broadband areas has been 0.4% more favourable annually in comparison with the alternative scenario where the expansion of broadband did not occur. The expansion of broadband has not eliminated rural depopulation, but investments have reduced depopulation in the rural areas where broadband has been built. Without the expansion of broadband, the depopulation of the rural areas would have been much higher. Therefore, broadband availability is an increasingly important part of the critical infrastructure that impacts business and household location decisions.

An important note on the expansion of broadband is that the effects of the construction appear in a few years after construction. This was evident in counterfactual analysis, which estimated population development in a situation where the broadband would already have been built in 2010. This would have resulted in an annual 0.7% more favourable population development in the constructed broadband areas compared with the observed population development in the rural areas. The effect was enhanced by the fact that the rural population could even have grown if broadband had already been built in the project areas in 2010. This increases the incentives for broadband expansion but on the other hand underlines that broadband cannot alone solve structural problems of the rural areas.

The results support the views that good broadband connections increase economic opportunities and improve the conditions for economic activity and development, as demonstrated in previous studies (Forzati et al., 2012; Kolko, 2012; Prieger, 2013; Atasoy, 2013; Mölleryd, 2015), thus contributing positively to population development of the rural areas. Indeed, the results provide an incentive to improve the availability of broadband in the locally active communities and to continue the expansion of the broadband with public support if such a local activity occurs in the rural areas. However, the connection to the local activity means that the implementation of broadband construction has remained the responsibility of small operators and local initiatives, and therefore not all project areas have been able to find operators for the construction. The geographically fragmented picture of broadband availability demonstrates that construction based on local activity excludes large areas from the broadband network. Thus, extending the geographical coverage of broadband availability and avoiding the deepening of the digital divide requires top-down coordination in broadband construction and more regionally tailored public funding. However, it should be noted that simply investing in broadband infrastructure in a structurally weak region may be insufficient to stimulate development if this region has a low level of human capital (Mack & Faggian, 2013; McCoy et al., 2018). For such a region, a more appropriate approach would be to complement the investment in broadband with measures to improve human capital (Hasbi, 2020; McCoy et al., 2018).

In interpreting the results of this study, one should consider that it mainly focused on the expansion of broadband in the core and sparsely populated rural areas, which mean that the results cannot be generalized to all rural areas. However, in the core and sparsely populated rural areas, the results support the view that in the 2010s, broadband had become an infrastructure comparable to roads and electricity (Salemink et al., 2017; Skerratt et al., 2012) which regulate the economic opportunities for the utilization of resources
located in these rural areas. The results reinforce the views of dispersion in the spatial distribution of economic activity related to broadband expansion, and this seems to be relevant to the rural areas (Ioannides et al., 2008). Broadband expansion is of particular importance in regional development because it is an element of structural competitiveness that can be directly influenced by policy action. Broadband construction has been able to support the conditions for the development of rural areas, especially in areas that have been spontaneously developed. Local activity, which can occur rapidly, develop informally, and focus on one-off development measures such as broadband construction, plays a key role in regional development and should be supported by policy measures.

In the future, regions’ disconnection from high-speed telecommunications networks will predict not only deepening population losses but also a decline in their economic opportunities. By connecting to broadband, the rural areas improve their relative position in relation to other regions, which in turn determines their chances of developing as part of networks on different levels. Without access to broadband networks, the digital divide in the rural areas will most probably widen further in the future. In this case, the rural areas outside the high-speed telecommunication networks may be cut off from many economic activities due to the underdevelopment of the infrastructure and high locational disadvantage. The COVID-19 pandemic has demonstrated the potential need for digital networks to support home working and service utilization, and therefore, it should be noted that broadband connections are also a threshold investment for the sustainable multi-local life and development of the place independent society.

A limitation of this paper is that the relationship between broadband availability and population development, estimated with DID regression and matching methods in an attempt to get closer to establishing causality, may still be subject to endogeneity. The matching of the broadband and non-broadband areas was based on the limited number of variables because of the data limitations of the population grid statistics which expose results for the problems of endogeneity. For instance, matching did not recognise the local activism on the broadband construction or differences between the project and non-project areas in human capital which can lead in estimation to the upward bias of the results. Nevertheless, the results give evidence that locally the broadband availability can reduce depopulation of the rural areas. However, in the future research, depopulation processes, such as negative net migration and natural population change, should also be integrated in analysis in order to understand more deeply the mechanism of the depopulation in rural areas after broadband expansion.

References

Amcoff, J. (2006). Rural population growth in Sweden in the 1990s: Unexpected reality or spatial–statistical chimera? Population, Space and Place, 12(3), 171–185.

Atasoy, H. (2013). The effects of broadband Internet expansion on labor market outcomes. ILR Re-view, 66(2), 315–345.

Austin, P. (2011). An introduction to propensity score methods for reducing the effect of confounding in observational studies. Multivariate Behavioral Research, 46(3), 399–424.

Autor, D. (2001). Wiring the labor market. Journal of Economic Perspectives, 15(1), 25–40.

Barley, D., Markley, D., & Lamie, R. (2007). E-commerce as a business strategy: Lessons learned from case studies of rural and small-town businesses. Clemson University: University Center for Economic Development.

Bates, D., Maechler, R., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models UsingEigenandS4.Rpackage version 1.1-7. http://CRAN.R-project.org/package=lme. (Accessed 3 January 2020).

Briglauer, W., Diirr, N., Falck, O., & Hüscherlath, R. (2018). Does state aid for broadband deployment in rural areas close the digital and economic divide? CESifo working paper series No. 6947. https://ssrn.com/abstract=3179949. (Accessed 21 June 2020).

Carrare, O., McGovern, C., Niorteg, R., & Schwartz, J. (2015). The willingness to pay for broadband of non-adopters in the U.S.: Estimates from a multi-state survey. Information Economics and Policy, 30, 19–35.

Czernich, N. (2014). Does broadband Internet reduce the unemployment rate? Evidence for Germany. Information Economics and Policy, 29, 32–45.

Dickes, L., Lamie, R., & Whitacre, B. (2010). The struggle for broadband in rural America. Choice, 25(4), 1–8.

Forman, C., Goldfarb, A., & Greenstein, S. (2012). The Internet and local spaces: A puzzle. American Economic Review, 102(1), 556–575.

Forzati, M., Mattson, C., & Aal-E-Raza, S. (2012). Early effect of FTTH/FTTx on employment and popu-lation evolution. An analysis of the 2007–2010 time period in Sweden. https://www.swedisch_centre.se/publications/earl-ef-ftth-f-ttx-employment-and-population-evolution-population-evolution.. (Accessed 3 January 2020).

Galloway, I. (2007). Can broadband access rescue the rural economy? Journal of Small Business and Enterprise Development, 14(4), 641–655.

Galloway, I. (2007). Can broadband access rescue the rural economy? Journal of Small Business and Enterprise Development, 14(4), 641–655.

Hashi, M. (2020). Impact of very high-speed broadband on company creation and entrepreneurship: Empirical Evidence. Telecommunications Policy, 44(3), 1–21.

Heckman, J., LaLonde, R., & Smith, J. (1999). The economics and econometrics of active labor market programs. In O. Ashenfelter, & D. Card (Eds.), Handbook of labor economics (pp. 1865–2097). Amsterdam: North-Holland.

Holst, L., & Jamieson, M. (2009). Broadband and contributions to economic growth: Lessons from the US experience. Telecommunications Policy, 33(10–11), 575–591.

Ioannides, Y., Overman, H., Rossi-Hansberg, E., & Schmidheiny, K. (2008). The effect of information and communication technologies on urban structure.

Koutsouris, A. (2010). The emergence of the intra-rural digital divide: A critical review of the adoption of ICTs in rural areas and the farming community. In I. Darnhofer, & M. Grözer (Eds.), Building sustainable rural futures: The added value of systems approaches in times of change and uncertainty (pp. 23–32). Vienna: 9th European IFSA Symposium.

Kown, R., Jokkonen, A., & Lemponen, V. (2018). Jokkonen, A., Lemponen, V. (2018). Verkosta vauhtia: Valokuvituskertoja ja digitalisaatiotyön, yritystoiminnan ja opiskelun mahdollistajina maaseudulla. In Ilmiönmääräisissä Aineissa: Alue- ja kuntatutkimuksenä Spatia.

LaRose, R., Strover, S., Gregg, J. L., & Straubhaar, J. (2007). The impact of rural broadband development: Lessons from a natural field experiment. Government Information Quarterly, 24(1), 91–100.

Lehtonen, O. (2010). Self-reinforcing spatial clusters of migration and socio-economic conditions in Finland in 1998–2006. Journal of Rural Studies, 26(4), 361–373.
Mack, E. (2014). Businesses and the need for speed: The impact of broadband speed on business presence. Telematics and Informatics, 31, 617–627.

Mack, E., & Faggian, A. (2013). Productivity and broadband: The human factor. International Regional Science Review, 36(3), 392–423.

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.

Möller, B. (2015). Development of high-speed networks and the role of municipal networks. OECD science, technology and industry policy papers No. 26. Paris: OECD Publishing.

Preston, P., Cawleya, A., & Metykova, M. (2007). Broadband and rural areas in the EU: From technology to applications and use. Telecommunications Policy, 40(7), 389–400.

Priejer, J. (2013). The broadband digital divide and economic benefits of mobile broadband for rural areas. Telecommunications Policy, 37, 483–502.

Pyykön, M., & Lehtonen, O. (2016). Importance of telecommunication networks for farms and municipalities. In Research, 56/2016 Helsinki: Natural Resources Institute Finland.

Ramirez, R., & Richardson, D. (2005). Measuring the impact of telecommunication services on rural and remote communities. Telecommunications Policy, 29(2), 297–319.

Rosenbaum, P., & Rubin, D. (1983). The central role of the propensity score in observational studies for causal effects. Biometrika, 70(1), 41–55.

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, 1–12.

Sekhon, J. (2011). Multivariate and propensity score matching software with automated balance optimization: The matching package for R. Journal of Statistical Software, 42(7), 1–52.

Stenberg, P., Morehart, M., Vogel, S., Cromartie, J., Breneman, V., & Brown, D. (2009). Broadband Internet’s value for rural America. United States Department of Agriculture: Economic research service.

Stem, M., & Adams, A. (2010). Do rural residents really use the Internet to build social capital? An empirical investigation. American Behavioral Scientist, 53(9), 1389–1422.

Townsend, I., Sathiaeseelan, 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(6), 580–595.

Traficom. (2019b). Availability of broadband with accuracy of 1km * 1km. https://eservices.traficom.fi/monitori/area?map=info-1249. (Accessed 20 April 2019).

Traficom. (2019c). Traficom’s state aid decisions for the Fast broadband project. https://www.traficom.fi/fi/viestintaverkot/laajakaista-hankkeen-tukias-maksupaatokset. (Accessed 6 May 2019).

Venables, W., & Ripley, B. (2002). Modern applied statistics with S. New York: Springer.

Whitacre, B., Gallardo, R., & Strover, S. (2014). Broadband’s contribution to economic growth in rural areas: Moving towards a causal relationship. Telecommunications Policy, 38(11), 1011–1023.