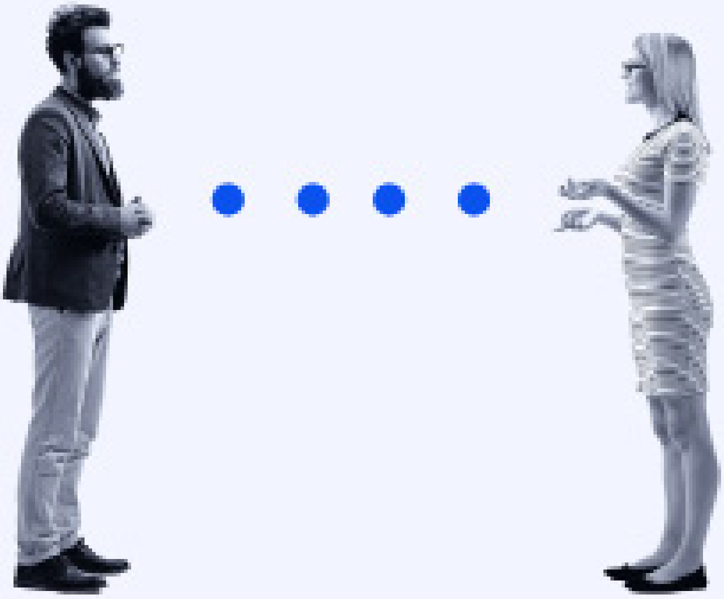
Technical Report

Measuring the Socio-economic Impact of **High-speed Broadband Deployment** in Rural Areas of Spain

Contents

Illustrations Index



Executive Summary

There is a reasonable consensus among economic agents, public authorities and the scientific community that the provision of fixed broadband connectivity is a key factor in boosting a country's socio-economic development. However, there aren't many specific studies on this impact in rural areas and particularly in Spain.

This study addresses the specific impact of fibre (FTTH) deployment on socio-economic aspects of rural municipalities in Spain between 2014 and 2018. Econometric techniques have been used to obtain favourable results regarding this impact. These include the impact on unemployment, the creation of businesses and household income.

A characterisation of the social, economic and demographic aspects of rural municipalities in Spain has also been carried out. The combination of this characterisation and the results obtained, enables conclusions to be drawn and recommendations to be made on the synergies between fibre deployment and other factors that contribute to the development of rural areas (the reduction of the digital divide, training, the generation of incentives for business development, the adoption of digitalisation and the promotion of new business models).

In addition, this study complements two other studies carried out within the framework of the project: 1) The analysis of the impact of fibre deployment on the Sustainable Development Goals as defined by the United Nations; and 2) The assessment citizens' perception through market research based on telephone surveys.

Context and Motivation

The Current Situation, Key Factors, Objective Definition and Initial Considerations

Telefónica

1.1 Introduction and Objectives

For at least the last 50 years and beyond, technological change has been considered the most important parameter for a country's economic growth [31]. In particular, a key element of this rule is investment in communication infrastructure. In fact, the United States Office of Economic Advisers estimated that for every dollar invested in Internet connectivity networks, an economic profit of almost three dollars can be generated [37]. Moreover, the Deputy Director of the Electronic Communications Market Analysis recently stated in the May 2019 economic bulletin that "encouraging investment in very high-capacity networks is increasingly important for competitiveness and economic development" [28].

In the late 1990s, the deployment of broadband networks, generally over DSL technology, transformed Internet connectivity by dramatically increasing access speeds [23]. Since then, broadband services have been considered as a way of providing benefits to all users (households, businesses and administrations), regardless of their size or geographical location [13]. For example, broadband services enable the growth and creation of businesses, thereby improving their capabilities, expanding their markets, improving their services and products and increasing employment by creating new business models. For this reason, it is widely agreed that the idea that the availability of broadband has a positive impact on social and economic aspects at a national and local level [1].

Additionally, the 2030 Agenda of the United Nations recognises the great potential of global connectivity to stimulate human development and as a cross-cutting tool that contributes to accelerating the achievement of each of the 17 Sustainable Development Goals (SDGs). Therefore, the goal of bringing universal and affordable connectivity is embedded within different targets and indicators (the SDG Targets: 9.c, 17.6 and 17.8).

However, as ITU also points out in its report "ICTs to achieve the United Nations Sustainable Development Goals", information and communication technologies (ICT), including the deployment of connectivity, are "a means to help to reduce poverty and hunger, boost health, create new jobs, mitigate climate change, improve energy efficiency and make cities and communities sustainable."

However, on top of the communication networks, it is necessary to build and develop all the services and capacities that allow society to benefit from these new technologies and access to the Internet. In recent years, there has been a growing line of studies seeking to assess the extent of this social and economic impact resulting from broadband network deployment [29, 30, 32]. This study looks into the effect of the deployment of connectivity services on a range of social and economic metrics, such as the GDP, employment and income, at international, national and regional levels. The study of the impact of broadband is a topic that is receiving a lot of dedication, with the effort applied by the body of authors being equivalent to that dedicated to studying a type of technology in itself [1].

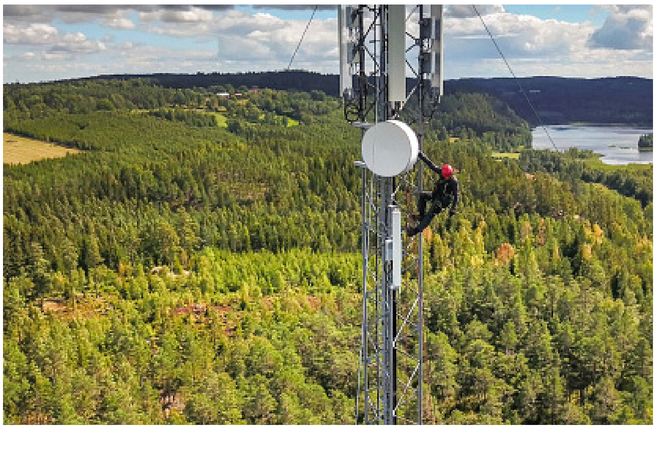
The authors intend to evaluate statements considered as fact by several organisations, for example that an increase in the speed of connection to the network implies social and economic benefits [3]. In this way, the studies analyse to what extent the deployment of broadband networks has a positive impact on economic and social variables and under what conditions. For example, one of the considerations presented in the literature is to evaluate in which cases there is a significant advantage in providing a broadband connection service over providing a non-broadband one, with the latter case being more cost-effective [1].

However, although a multitude of studies have been conducted globally focusing on regions with high population density, there are a lack of studies being carried out on evaluating the impact of broadband deployment on economic and social factors in rural regions [37]. Moreover, covering this lack of studies on the impact of broadband on economic and social variables in rural regions has been defined as essential by several authors [26, 29, 32]. Moreover, this deficiency is even greater in studies focused on regions in Spain [29]. The main reason for this need is that results obtained and conclusions generated by studies that cover urban regions may not be extrapolated to rural regions with lower population densities [35, 29], due to the influence on the impact level of some key factors such as the penetration level. Furthermore, some study variables highlighted by the authors as being key in the evaluation of the impact of broadband in rural regions, such as population density, the spread of population centres, ageing inhabitants or the reduction of the digital divide, are not considered in studies focused on urban deployments [29]. Taking all this into account, **this project aims to apply a**

systematic approach to provide a better understanding of the impact of a fixed broadband service on economic and social variables in rural Spain. To do this, this study seeks to generate new findings on the economic and social impact of broadband, as well as the main limitations, motivations and implications of the deployment. In addition, the study seeks to answer questions such as whether broadband deployment in certain municipalities and under certain conditions can generate a commensurate increase in social benefits and economic activity. In order to successfully complete the objective, an extensive compilation of high-quality data has been created from a variety of sources, as well as the use of a set of fully interpretable techniques that allow clear conclusions to be drawn for the design of strategies and action plans.

Taking all this into account, **this project aims to apply a systematic approach to provide a better understanding of the impact of a fixed broadband service on economic and social variables in rural Spain**. To do this, this study seeks to generate new findings on the economic and social impact of broadband, as well as the main limitations, motivations and implications of the deployment.

In addition, the study seeks to answer questions such as whether broadband deployment in certain municipalities and under certain conditions can generate a commensurate increase in social benefits and economic activity. In order to successfully complete the objective, an extensive compilation of high-quality data has been created from a variety of sources, as well as the use of a set of fully interpretable techniques that allow clear conclusions to be drawn for the design of strategies and action plans.



Beyond the impact on macroeconomic indicators, it is unquestionable that the availability of high-speed fixed broadband improves the quality of life for citizens. However, there is practically no data available to quantify this impact at the municipal level. The study is therefore complemented by field research based on direct surveys of citizens.

1.2 Broadband and Fibre to the X Technology (FTTx)

According to literature, the term broadband seems to have an ambiguous definition. This is probably due to the fact that the meaning of the term has evolved over the years and the consequent increase in network access capacity [52]. In fact, the first offer in Spain (1997) included broadband which provided speeds of 256 Kb/s [28]. For this reason, literature opts for a definition without numerical implications: "Broadband refers to a signal or a circuit that supports a relatively wide range of frequencies and can provide a substantially higher transmission speed than that provided by a conventional telephone line" [26, 54].

In this way, each author tends to specify the numerical value that they consider to be 'a substantially higher transmission speed than that provided by a conventional telephone line', thus establishing their own qualitative definition in their studies. In this study, we consider the specification of the International Telecommunication Union [43], which establishes two minimum speed values to consider as an access such as broadband: speeds of 10 Mbps worldwide, and 30 Mbps in certain countries, such as the member countries of the European Union. Therefore, **we consider 30 Mbps as the minimum high-speed broadband speed in this study**.

Moreover, the broadband service can be provided through multiple technologies, which present different challenges and costs for deployment and maintenance. Traditionally, broadband service has been provided over the xDSL technology family[1]. However, these technologies, with some exceptions, do not provide speeds of above 30 Mbps [26]. This is one of the reasons why national and international targets, such as the European Digital Agenda [53], consider fibre-based FTTx technology as key to meeting connectivity goals [43, 44]. In particular, this study considers FTTH, Fibre to the Home, technology.

[1] All DSL (Digital Subscriber Line) technologies, usually over copper cable.

> **The European Digital Agenda considers FTTx as the key technology to provide speeds above 100 Mbps, and thus facilitate socio-economic development.**

In addition, one of the main political and strategic objectives of the European Commission to boost the economy and foster job creation is the deployment of Next-Generation Access Networks [29]. These networks are highlighted in literature as an important challenge for the economic and social development of a country [13] and a means to greatly increase social and territorial cohesion [33]. Among the set of technologies that form the Next Generation Access Networks, FTTx technology is considered to be key for being the best option to provide high speeds, higher than 100 Mbps [28, 14]. FTTx is a fibre-based Internet access technology [39]. In particular, Spain is a success story in the deployment of networks based on this technology, with FTTP (Fibre to the Premises) coverage for 77.4% of households in 2018, being the country with the highest number of connected households in Europe [26].

For these reasons, this study addresses the impact of the high-speed broadband service provided by networks based on FTTx technology. However, there are different technological options to provide fixed network coverage above 30 Mbps, including, in addition to FTTH, wireless networks, VDSL (Very-high-rate Digital Subscriber Lines) and HFC (Hybrid Fibre Coaxial). Some similar studies [1, 31, 35] include information on broadband availability covering one or more technologies without clarifying the degree of deployment of the previous ones. This can sometimes produce distorted results. To combat this problem, this study limits its scope to the existence of broadband networks above 30 Mbps based on FTTH technology.

1.3 Key Social and Economic Impacts

The impact of the broadband service on social and economic variables has been addressed by several researchers [24, 25, 31]. The authors of these studies concluded that **broadband networks produce significant positive effects on various social and economic factors**, such as encouraging the creation of new businesses and increasing household income. In fact, studies have claimed that there is a positive impact of these networks on the GDP, with increases of 1.1% [23] and 1% [26]. This section describes the three indicators affected by broadband deployment that are considered most relevant in literature [15, 16, 38]. These three indicators are **companies, employment and income**.

1.3.1 Companies

The deployment of broadband networks creates multiple business opportunities and helps companies to increase their economic activity [29]. This is because the broadband connectivity service is a fundamental enabler for the digital transformation of companies, encouraging the development of new business models as well as improving existing ones. In this way, companies can expand their relationship channels with customers and suppliers, provide new and better services, reduce operating costs and expand their potential markets, among other benefits. Moreover, the authors claim that the deployment of broadband networks can provide an environment which is conducive to entrepreneurship, especially for small businesses [22, 37].

Moreover, broadband services have demonstrated the potential to change how business activities are organised, innovating processes and generally having a positive impact on the economic situation by increasing competitiveness and flexibility [23]. In [26, 37] it is concluded that broadband enables communication to be improved and costs to be reduced, using methods such as Voice over IP, and provides the possibility of electronic integration with suppliers and customers, software improvements, the possibility of teleworking, outsourcing, and ease of access to information and services.

In spite of all this, it is important to consider that a relevant materialisation of broadband deployment in companies is part of the digitalisation process of these companies and may require an adaptation time that allows for the preparation of personnel, technological resources, organisation and processes. Moreover, although there is abundant empirical evidence of the economic impact of broadband and its positive externalities on innovation, productivity and business restructuring, research shows that these effects vary depending on the environment in which broadband is deployed [8]. For example, there is an important difference on the impact if regions with greater or lesser prior technological development are considered, as it may be necessary for the company to first invest heavily in training employees in technological skills.

In particular, broadband deployment can provide rural areas with added advantages in comparison with urban ones. Some studies claim that broadband provides an attractive value to rural regions that encourages the creation or establishment of businesses [21, 22, 5]. Moreover, there can be additional benefits for rural businesses. For example, broadband deployment can improve communication between geographically isolated areas, supporting processes within an organisation and between multiple organisations. It also provides a way of staff accessing free or low-cost training resources. In fact, authors show that rural areas with broadband access see an increase in the number of medium and high skilled workers, which therefore, tends to reduce the level of poverty in the region [5].

As mentioned, in any given region the impact of broadband deployment on businesses depends greatly on the characteristics of the businesses and on the type of industry and the environments in which they operate. Fortunately, studies suggest that businesses in the service sector can benefit from broadband even in the most remote locations [37], this being the sector that benefits the most from the broadband service as it makes more intensive use of ICTs [14], and the service to which tourism belongs. However, a large number of companies in the service sector, such as entertainment, banking and education, usually need fewer employees than traditional sectors.

In addition, some authors [21, 14] emphasise that the availability of broadband service can encourage strong growth and the creation of rural businesses, such as livestock and agricultural enterprises. Moreover, the deployment of broadband in rural regions may have a positive effect on the number of companies operating in the construction sector, as the deployment of networks increases the workload and may require the creation of direct jobs, albeit of a temporary nature.

1.3.2 Employment and Income

It is important to extend the analysis to the micro and enterprise level to obtain results that can be representative of the magnitude of the social value generated by broadband [1]. For this reason, various studies [29, 30, 32] have chosen to study the impact of broadband on socio-economic variables, such as employment and income.

The impact of broadband deployment on employment levels has been the subject of study in recent years in economic and technological literature. Some authors have found that broadband could have a positive impact on employment generation [29, 30, 32]. However, there are difficulties when it comes to accurately measuring causality as there are four different impacts or repercussions [29], which are described below.

First, increased broadband penetration can increase productivity and process efficiency; this can lead to a reduction in jobs. Second, broadband service contributes to increased innovation and the creation of new business models, which can generate new jobs. Third, there may be an 'outsourcing' effect of certain business functions, which may result in job creation from the creation of outsourcing companies, or a loss of local employment opportunities if functions are transferred to other geographical areas. Fourth, the broadband infrastructure deployment work can have a positive impact on employment in the construction and manufacturing sectors, particularly if local recruitment occurs.

The four impacts, their consequences and their implications on unemployment are presented in Figure 1. The first three are part of the indirect impact, corresponding to medium and long-term results, while the last one is a direct, short-term impact. This paper studies the impact of broadband on job creation and on the employment rate, considering both the new jobs created as well as jobs that become obsolete.

Similarly, the effects of broadband deployment on employment are conditioned by the characteristics of different regions [1]. In fact, some authors claim that the potential of broadband to produce employment growth has materialised in technologically and economically developed regions, with middle/high income population and ICT skills, but not in the rest [4, 5]. In particular, these authors agree in distinguishing the level of impact on urban areas in comparison with rural regions. Moreover, while some papers highlight the importance of initial broadband deployment, others claim that improving the access speed can lead to job creation three times greater than that produced by the previous broadband service [13, 1].

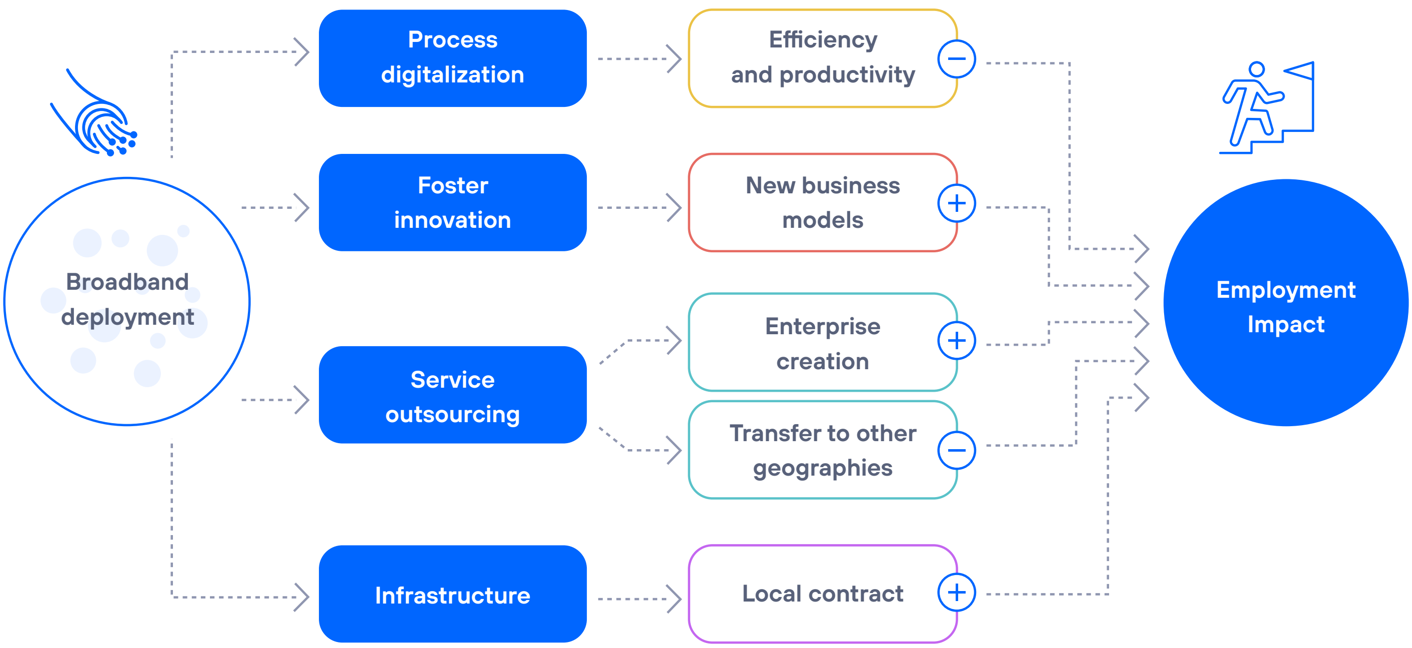


Figure 1: The impact of broadband deployment on employment.

In addition, other studies have obtained different levels of influence for broadband on employment depending on the type of industry being considered [6, 10, 19, 27]. For example, the authors of [6] state that the positive influence on job creation has a significant effect on the service sector, mainly in finance, insurance, education and health care, with the effect on the rest of the sectors being slight or non-existent. Similarly, another study [11] found that the deployment of broadband services promoted employment growth in ICT-intensive industries but did not positively affect salaries or average employment rates. By contrast, other authors have found no evidence that providing a broadband service reduces the unemployment rate in any sector [15, 20].

However, the income growth of broadband consumers is another relevant factor that can be positively influenced by broadband deployment [13, 29]. In fact, some studies [51, 52] have focused on studying the impact of a broadband service on household income. In particular, these studies investigate two aspects: the impact of increased broadband penetration and the impact of the increased access speed. According to the authors, the results obtained vary greatly depending on the characteristics of the region being studied and the quantitative values of access speed being considered.

Regarding the characteristics, the level of the impact of broadband on the income value is affected by the situation of each region in certain variables [51, 52]. Some of these variables include education, skills, occupation type, age as well as the previous income level of the population; type of housing; the proportional weight of each economic sector, in particular the size of the service sector; the technological adoption capacity of organisations and institutions; the quality of public services and the level of public spending; and its technological, energy and transport infrastructure, among others. In general, these variables are associated with the digitalisation capacity of the population and organisations of a given region. As a result of this, there is a risk of increasing inequality between regions within a country or between rural and urban areas.

Regarding rural regions, some authors who have chosen to study the impact of broadband in rural regions have obtained results that enable us to state that broadband deployment can indeed generate a positive impact on economic growth, highlighting a reduction in unemployment and an increase in average household income [22, 5]. This is mainly due to an increase in the number of new companies and a reduction in the number of companies that disappear.

In this regard, another study concluded that the deployment of broadband coverage increases employment in the service sector in rural regions, but does not produce the same effect in urban regions [12]. This is because broadband helps to overcome geographical difficulties, as discussed in the previous point. However, in rural regions it may be that improvements in the efficiency of a company's processes generate a reduction in employment in the area, leading to an erosion of the local market and local services [37].

In this regard, another study concluded that the deployment of broadband coverage increases employment in the service sector in rural regions, but does not produce the same effect in urban regions [12]. This is because broadband helps to overcome geographical difficulties, as discussed in the previous point. However, in rural regions it may be that improvements in the efficiency of a company's processes generate a reduction in employment in the area, leading to an erosion of the local market and local services [37].

Moreover, as discussed above, providing broadband services to rural regions may lead to a relocation of employment to urban municipalities. This fact is concluded in some studies such as [55], in which the authors state that broadband does increase employment in municipalities located far away from regional centres, but does not do so in municipalities which are close to regional centres. In any case, broadband can help to maintain the activity of companies which could disappear without this service, as a consequence with an increase in unemployment [5, 55]. In fact, the disappearance of businesses in rural areas is a serious problem [37].

> **Literature in this field has found that there is a relationship between the impact of broadband network deployment on different socio-economic indicators; particularly on employment, income and business creation.**

In terms of population density, broadband deployment can serve as a way of retaining and attracting people to rural communities [5, 37]. In addition to having a positive impact on economic aspects, such as those mentioned above, the broadband service can also generate indirect benefits for citizens. For example, broadband makes it possible to provide better social, education and health services, entertainment, and to enable teleworking and an electronic government. These benefits can increase citizens' quality of life [26]. In this way, a region can become a more attractive place in which to reside, thus preventing depopulation [37, 14].

In line with all this, various authors have obtained results that support both a significant positive and a significant negative impact, as well as a non-significant impact of broadband on employment and income [1, 30, 32, 52, 55]. According to several authors [16, 30, 32], this may be due to the dependence of the results obtained through a study with the situation of certain complementary variables, such as geographical conditions, the competitive climate and the general macroeconomic environment of each region being studied during the time period considered.

To visualise this fact, a comparison of the results from various studies is presented in Table 1. However, as can be seen in this table, there are a greater number of studies that have obtained positive impact results, of a greater or lesser significance. Therefore, there is a positive trend in the results for the different types of areas analysed. In particular, this study aims to address the impact of broadband in the specific scenario of municipalities in rural regions of Spain.

	Time Range	Geographical Scope	Municipality Type	Object of the Study	Impact Results
[1]	2011-2014	496 municipalities in the USA	They make up 80.7% of the U.S. population	Broadband availability and connection speed	Increased employment regardless of the connection speed provided
[4]	1995-2000	163 municipalities in the USA	Municipalities with more than 150,000 inhabitants	Availability of the broadband service	Increase in income and employment in 42% of urban regions and 6% of rural regions (danger of imbalance)
[5]	2008-2011	1,046 urban municipalities and 2,037 rural US municipalities	Municipalities with more than 50,000 inhabitants, the rest are rural municipalities	Broadband availability and adoption (200 Kbps)	Reduction of unemployment and increase of income in cases of high adoption (availability alone is not sufficient
[9]	2009-2014	351 municipalities in 39 regions of Canada	Municipalities with populations of between 64,000 to 5,336,000 inhabitants	Deployment of fibre networks in regions without it	100% penetration in a region increases employment by 2,9%
[12]	1997-2011	4,500 municipalities in 38 urban and 31 rural regions of Canada	Municipalities with more than 50,000 inhabitants, the rest are rural municipalities	Broadband availability	Increase in employment in rural regions (1.1%) at the expense of urban regions and an increase in income by 1%, although a minimum penetration rate is required
[13]	2010-2015	4,933 municipalities in France excluding Paris, Lyon and Marseille	Municipalities with more than 2,000 inhabitants	High-speed broadband (fibre) availability	A positive impact on business creation (3.9%) and non-significant impact on employment with increased availability and increased speed
[15]	2002-2006	6,919 municipalities in Germany	Municipalities with less than 22,803 inhabitants	Increased broadband availability (over xDSL)	Lack of impact on income and employment
[20]	2008-2011	415 municipalities (174 rural) in 8 US states	Not specified	Increased bandwidth availability	No impact on employment in rural municipalities and 0.49% growth in urban municipalities
[23]	2011-2012	45 metropolitan areas in 9 US states	Municipalities with less than 1 million inhabitants	Increased the availability of gigabit networks over FTTH	Increase in Employment and GDP growth (1.1% ~ 1 Billion) in high penetration regions
[25]	1998-2002	22,390 communities (postcode) in the USA	Not specified	Broadband availability (200 Kbps)	Increase in employment (1%), if there is connectivity versus no connectivity, with no impact or negative impact on income
[31]	2002-2007	22 OECD countries (including Spain)	Country level	GDP impact of broadband broadband penetration (DSL, cable, fibre and others)	A 0.025% increase in economic growth (GDP) for every 1% increase in broadband penetration.
[55]	2005-2009	8,460 rural municipalities in Germany	Municipalities with less than 50,000 inhabitants	Broadband availability (384 Kbps)	0.03% increase in service sector employment for every 10 percentage points increased through broadband.
[57]	2010-2012	290 rural and urban municipalities in Sweden		FTTH penetration level	1.1% increase in employment rate with a minimum penetration rate of 10%.

Table 1: Summary of the literature reviewed that addresses studies on the impact of connectivity deployments on socio-economic factors.

In conclusion, there is also a set of studies which include the study of the impact of an increase in broadband connection speed in comparison with the increase in the availability or its penetration [1, 13, 23, 58, 59, 60]. These studies show that an increase in the broadband speed can have positive impacts on socio-economic variables [58], such as increasing GDP by 0.3% if the access speed doubles (from 8.3 Mbps to 16.6 Mbps) [60]. However, these impacts can only occur if a high network penetration is achieved [23], and their level of impact is highly dependent on a region's economic growth and its use of technology, with the impact being higher in economically and technologically developed economies [9].

Moreover, high-speed network deployments, which increase both the availability and connection speed, achieve higher impacts than deployments aimed at increasing speed: up to a 2.5% increase in GDP [31] in comparison with up to a 1.1% increase [23], and an increase in income and employment [9, 25, 57, 55] compared to no increase occurring [1, 13]. Furthermore, according to what has been previously discussed (summarised in Table 1) and in [1], it can be stated that broadband availability has the potential to provide positive impacts on socio-economic variables, regardless of the speed provided (positive results with speeds ranging from 200 Kbps xDSL to high speeds through FTTH). Finally, it is interesting to note that the articles dealing with rural areas have considered the availability of broadband as a key aspect rather than the increase in access speed.

In this study, we analyse the impact of broadband availability without considering the increase in access speed. Having said this, this increase in speed may occur indirectly, due to the existence of prior technologies other than FTTH not being considered.

Moreover, although data on broadband availability can provide representative information on the impact of broadband on socio-economic variables, information on adoption or uptake provides a better representation of the impact [61]. To the best of our knowledge, there are a lack of studies being carried out to look into the difference of the impact of broadband adoption versus the impact of availability. This is probably due to the fact that there are great difficulties in obtaining uptake level data [61]. Moreover, the importance of knowing adoption information is even greater in rural regions, which are more difficult to obtain accurate data about [63].

In particular, the only paper found that addresses this issue is [62], whose authors claim that reaching high levels of broadband adoption (above 60%) has a positive impact on income, the number of companies and employment variables, while low adoption (below 40%) is associated with a decrease in employment and number of firms. Thus, the authors of [62] conclude by stating that deployment strategies and policies should be on demand and adoption oriented. In fact, there are multiple articles that address those factors that influence adoption and propose measures that can be applied to encourage uptake, such as [64, 65].

This study has not specifically analysed the casuistry of adoption, but it has found, for example, that household spending on "communications" is a factor that has a simultaneous impact with fibre coverage. Of course, this factor is not a direct measure of adoption, but it points to a trend that would certainly be interesting to explore further.

1.4 Governmental Action Policies

Broadband networks are considered to be a key factor for socio-economic development. For this reason, their deployment has become a priority worldwide and is considered an essential investment for the future [28]. In recent years, countries such as the United States, Australia, Japan, Mexico and numerous African countries have developed national intervention and action policies to ensure that all citizens have access to broadband services [13].

Currently, policy makers are shifting their focus from broadband services to high or super broadband connectivity, with gigabit speeds [1, 18]. This is due to the fact that certain regions, the vast majority of which are urban, have already had large broadband network deployments available for several years. However, there are still regions, generally rural ones, which do not have access to this service. Moreover, recent studies [35] state that the vision of entire populations or geographical areas having access to broadband infrastructure may not be achievable because there are issues with the return on investment. In these regions, the broadband service can be provided by private initiatives through other methods rather than fixed networks, such as satellite or mobile ones. However, this is not always possible [14].

For this reason, broadband availability has lagged behind in certain regions, especially in non-urban ones, creating a divide between individuals, businesses and geographical areas. This is known as the broadband divide, which is included in the digital divide. To combat this problem, national and international policies seek to address this need and support initiatives to ensure a broadband service to all households. In this way, the aim is to ensure that citizens living in regions where network deployment is difficult to accomplish have the same opportunities to increase their household economy[52] as well as access to information, culture and public services [29].

In our closest international area, the European Commission has defined a set of digital tasks, reflected in the European Digital Agenda [45]. One of these tasks is broadband universalisation, which includes providing 30 Mbps coverage in all homes and 100 Mbps in 50% of homes by 2020. Thus, the aim is to increase broadband penetration and reduce the digital divide. A set of studies assesses the economic benefits that could come from fulfilling the tasks defined in the European Digital Agenda [17, 18]. The authors obtain positive results and in line with the objectives of the European Commission, affirming economic benefits that affect companies, users and the national economy.

With regards to Spain, in October 2019 the Ministry of Industry, Energy and Tourism stated that there is a relationship between the deployment of new broadband networks and the increase in GDP and the creation of new jobs in the short and long term [2]. In addition, it also stated that Spain has significant room for improvement compared to other European countries to provide a contribution to the national and local economy through broadband deployment [2].

However, according to T. Krestshmer [16], Director of Science, Technology and Industry at the OECD, both public and private initiatives may encounter difficulties in defining broadband network deployment strategies because, as previously discussed, there are a wide variety of studies that reach diverse and inconclusive conclusions. In fact, some authors [35, 17] have devoted their efforts to analysing multiple action policies and have concluded that there are a wide range of public measures that do not provide enough detail or do not include certain relevant aspects, making their application somewhat difficult. As a solution, this study has a defined scenario, rural municipalities[2] of Spain, and recent information, the last five years (2014 to 2018). Thus, the objective is to provide representative knowledge that validates the existence of the impact of broadband on socio-economic variables in rural regions of Spain. This information may be considered by public and private entities and thus encourage the possibility of the deployment of these networks.

1.5 Broadband in Rural Areas

Some authors have defined the study of broadband availability in rural regions as essential [26], since the lack of equal availability in all parts of a region limits the possibilities for growth and territorial rebalancing, thereby increasing territorial, social and economic imbalances. This problem is known as the digital divide. In terms of fixed connectivity, Spain reached 92.9% rural coverage of standard fixed networks (VDSL) in 2018, higher than the European average. In terms of new generation broadband networks, coverage in rural areas increased from 23.13% in 2013 to 47.64% by 2018. This data shows very significant growth, but it is below the European average and far from the percentage of standard fixed network coverage [53]. This difference is also notable between rural areas and the national average, as the data shows that, while broadband coverage of more than 30 Mbps in 2018 reached 85.07% of the total population, in rural areas such coverage only reached 54% [51].

Among the factors that explain this difference are the cost of deployment associated with geographical dispersion, topography, etc. In addition, the profitability criteria of the operator themselves must be taken into account, given that the deployment of new generation networks must have a profitable and sustainable economic model that motivates the necessary investments by all the agents involved. The Spanish State provides aid to encourage such investments. In the State aid rules, it defines areas that do not have coverage of new generation broadband networks, nor forecasts their provision by any operator within 3 years, based on credible investment plans as white NGA areas.

[2] *According to Law 45/2007 on the Sustainable Development of the Rural Environment, rural municipalities are those with less than 30,000 inhabitants and with a population density of less than 100 inhabitants per km^2.*

> Due to the benefits it brings to businesses and society in general, the European Digital Agenda proposes that 100% of households should reach 30 Mbps, and 50% should reach 100 Mbps by 2020.
>
> Spain has made a pioneering effort in fibre deployment in recent years. However, in rural areas it is necessary to continue doing so.

In 2016, the resident population in the set of 6,664 Spanish municipalities that we have defined as rural reached 7.76 million, of which 4.7 million resided in white areas, i.e., more than 60% of the population lived in areas with little prospect of receiving fibre deployment. Moreover, this distribution is not uniform, with residents in small towns being especially penalised (Figure 2). For example, in municipalities with less than 2,000 inhabitants, the percentage of the population living in white areas reached 83%.

State incentives for investment in fibre deployment in white areas and the efficiency of deployment plans have had a very positive impact on reducing this gap. In Figure 3 it can be observed how the volume of rural population residing in white areas has reduced significantly between 2016 and 2018, going from 4.7 million people to 1.9 million in 2018.

This very positive fact includes some nuances, since the reduction has been unequal when we consider the population range of the municipalities (Figure 4). While a sharp reduction of the population in white areas is observed for municipalities between 2,000 and 5,000 inhabitants, the reduction in municipalities with less than 1,000 inhabitants has been rather more moderate, with a population of approximately 1 million people residing in white areas of municipalities with less than 1,000 inhabitants being recorded in 2018.

% Resident population in white areas and population per geotype (2016)

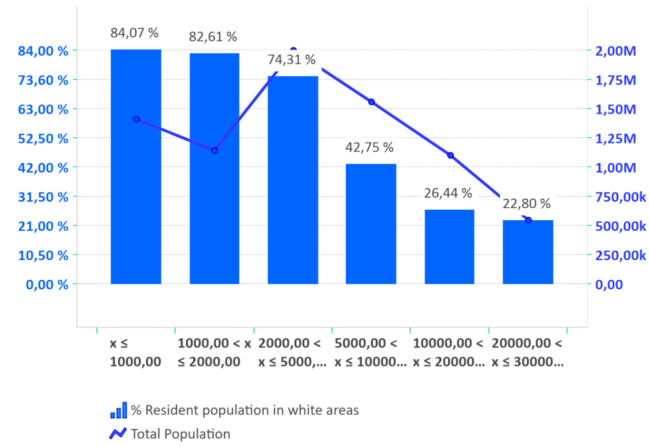


Figure 2: Percentage of resident population in white areas and population volume for rural municipalities in 2016 according to their population range. It can be observed that the ratio of people living in white areas is highly dependent on the number of inhabitants of the municipality, clearly unfavourable for municipalities with less than 5000 inhabitants. (Source: Our own creation).

Resident population in white areas / Year

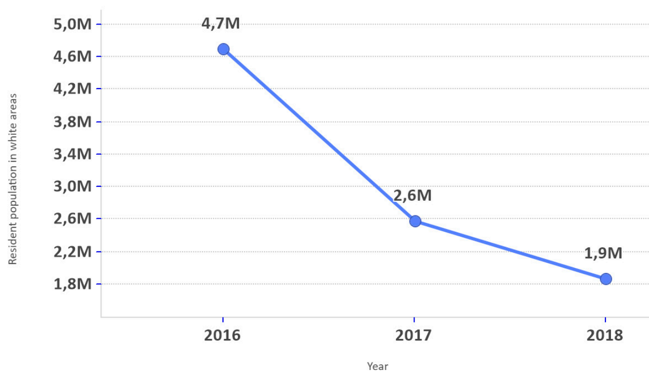


Figure 3: Evolution of the rural population residing in white areas for the period ranging from 2016-2018 (Source: Our own creation).

Resident population in white areas / Year

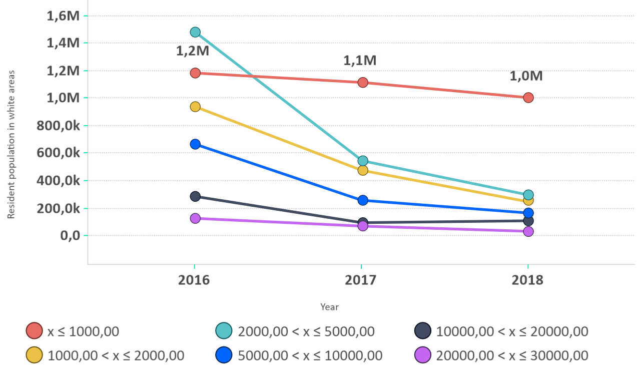


Figure 4: Evolution of the resident population in White Areas according to the population ranges of the municipalities.

Study Methodologies

Studies that aim to analyse the impact of ICTs on social and economic variables and the circumstances of their use must consider and establish three different aspects [16]. First, the type of ICT product or service. Second, the level of aggregation: international, national, regional, industry and the company. Third, the methodology used and the variables considered. The first two points have already been discussed in the previous section in the study of broadband service in municipalities in rural regions in Spain. The third point, related to the analysis methodology, is discussed in this section. This is a key aspect when it comes to drawing conclusions from the study.

2.1 Study Methods

The study of broadband network deployment policies and strategies and their impact on social and economic factors has been addressed in literature considering an analysis of the available historical information [30, 32]. For this purpose, historical information can be analysed by applying two complementary methods [35, 36]: case study analysis and econometric or empirical analysis.

The first method is based on comparing the situation and policies and strategies in large regions, usually different countries, and finding patterns that can explain the causes of broadband penetration; whereas the second method is based on using empirical aspects, such as data, models and statistical techniques, to discover the key factors that explain broadband penetration and its impact on economic, social and political variables. Both methods are shown in Figure 5, based on [36].

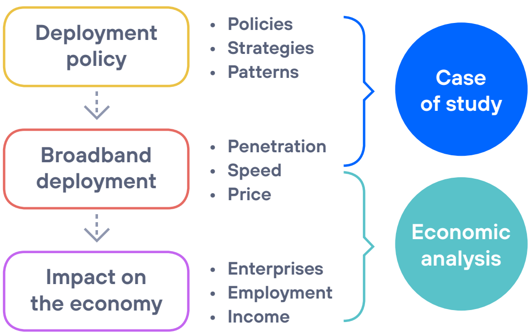


Figure 5: Methods for studying the deployment of broadband networks.

The objective of this project is to study the impact of broadband on economic and social variables, considering the quantitative variables. Therefore, the econometric method based on empirical data is used in this study.

By contrast, empirical studies can be divided into two approaches according to the information they consider to assess the impact [23, 29, 30]. First, those that perform an analysis of broadband network deployment with the objective of assessing the **direct benefits** of infrastructure investment and of direct consumer spending on services, known as surpluses. Second, studies that aim to analyse the indirect benefits, which encompass multiple aspects that are affected to a greater or lesser extent and in the medium and long term by the deployment of broadband networks. Some of the **indirect benefits** considered by this second approach include cost savings, increases in average household income, productivity gains in enterprises and the creation of new companies and business models.

In this study, the indirect approach is used in order to provide information of greater value than the direct approach. This is because the indirect approach allows the study of key economic and social factors that have a greater and longer-lasting impact over time. Both approaches have been applied simultaneously and individually in this study, considering the increase in employment in the construction sector in the direct approach. However, for this sector, no significant impacts have been obtained in the study.

2.2 Econometric Method

Similar studies which apply an econometric or empirical approach vary in the method used depending on issues such as the type and amount of data and the relationship between variables [15, 16, 31]. As discussed in section 2.1, all these studies consider the analysis of the historical data available to generate knowledge about the impact of the broadband service and the deployment strategies. This analysis can be carried out considering three methods. Each of them applies a different approach to the study of the relationships and patterns of the variables. The information presented below on each of the methods has been obtained from [40, 41, 42].

i) Application of **time series**. This method analyses the information of the variables considering a temporal dimension divided into instances in time. For this, each instance of time constitutes the population or sample element.

ii) **Cross section**. This method applies an analysis of the information considering multiple population elements in a single instant in time. In other words, this method is characterised by having a structural dimension rather than a temporal one. Thus, each unit of analysis forms a population element.

iii) **Econometric method using panel data**. This considers a complete sample of the analysis units or various instances in time. Therefore, this method considers a structural dimension over a temporal one.

According to the literature on panel data techniques [46], the main characteristics are that (i) they allow studying the patterns of each element, in this case, municipalities, over time, (ii) simultaneously preventing the problems of masking effects when aggregations are made, and (iii) they allow estimating models which consider permanent differences between individuals even if they are not observable.

The first two characteristics refer to the fact that it is possible to study individual effects and time effects. Individual effects are those that affect each of the individuals that make up a sample (whether households, companies, municipalities, etc.) or even groups of them in a different way, with these effects being invariable over time. For example, this effect makes it possible to study aspects such as the type of sector, the technological base or the qualification of the population. Similarly, temporal effects are those that affect all the units under analysis within the study equally and vary over time. This type of effect usually occurs when there are macroeconomic factors, which can affect all the units of study equally, such as the effect of an economic crisis on the economic variables of companies and households.

The third characteristic, which is a very relevant contribution of the panel data method, refers to the fact that it is possible to capture unobservable heterogeneity, in other words, obtaining a reflection of the pattern of variables even when it is impossible to find out all the information about them or about other related variables.

In addition, the panel data method has other advantages over the time series and cross-sectional methods. Some of these advantages are listed below. First, it allows a better representation of high volatility processes, which are abundant in studies related to macroeconomics. Second, this method allows us to consider a larger number of observations, increasing the degrees of freedom[3] and reducing the effect of correlation between explanatory variables. Finally, the panel data method allows us to consider the possible existence of unobservable heterogeneity[4], thus reducing the bias of the results.

> **Most studies in this area use econometric analysis, an in particular panel data.**
>
> **In applying these techniques it is necessary to address the endogenous effects that exist between the variables involved.**

In particular, the panel data method is advantageous for this type of study because it solves a common problem related to data precision [1, 40]. This problem occurs because the available datasets often provide information that is approximate and timely. However, panel data uses these data sets as representations of time periods and generates the models on them. As a result of all this, most of the similar studies considering an empirical approach use a panel data model [1, 29, 15]. For all these reasons, the panel data model has been chosen to study the relationship between broadband deployment and the socio-economic variables.

[3] *The degrees of freedom express the amount of information from the data that can be used to calculate the variability of the estimates. The greater the degrees of freedom, the more reliable the estimates obtained will be.*

[4] *Unobservable heterogeneity is the error that occurs in a model due to variables which, by their nature, are unobservable, and therefore cannot be explicitly included in the model (e.g., population attitudes, such as a predisposition to technological change or of an entrepreneurial nature).*

As a result of all this, most of the similar studies considering an empirical approach use a panel data model [1, 29, 15]. For all these reasons, the panel data model has been chosen to study the relationship between broadband deployment and the socio-economic variables.

However, the panel data method can also present some complications. One of the biggest problems that can arise when a panel data study on a highly complex aspect is required is endogeneity[5] [12, 9, 29]. Endogenous effects occur as in the same way that there can be an impact of fibre deployment on economic conditions, there can also be the opposite effect, i.e., that the socio-economic conditions of the municipalities trigger the fibre deployment. In this case it is said that **broadband service deployment can be endogenous to economic conditions**.

For example, broadband deployment may have an impact on population density and income levels, which in turn may also be related to the corresponding regional economic activity, the latter being a factor that may directly or indirectly influence operators' broadband deployment selection decisions. To counter this problem, some studies have considered using instrumental variables[6] [38, 31, 16]. This is the solution used in this study.

Another of the most relevant problems in similar studies is the absence of a benchmark to assess the differential effect of a broadband network deployment and thus prevents bias [9]. In other words, a reliable benchmark is needed, such as regions which do not have a broadband service and which have similar characteristics to regions that do have a service. To combat this problem, this study has considered a wide range of municipalities which are representative of different levels of fibre coverage penetration. This, added to the fact that the panel data technique also takes into account this problem, allows us to control the bias of the results obtained.

[5] *In a model that seeks to explain Y with Xk variables, endogeneity occurs when either of the X and Y variables are bidirectionally related. For example, in a supply-demand model, if one tries to predict price (Y) from demand (X), endogeneity will arise, since producers adjust their price based on the demand (X causes changes in Y) but consumers also change their demand in response to price (Y causes changes in X).*

[6] *In a model that seeks to predict the Y variable with the Xk variables, an X instrument is a new variable to the model which, being related to X, does not directly affect Y except through X.*

2.3 Studies Applied to Spain

There are some studies that look into the deployment of broadband networks in Spain. These studies can be categorised into two groups depending on their focus. One group of studies address the reasons motivating as well as the consequences of broadband deployment at an international level and includes Spain as part of their analysis [26, 35, 31, 32]. This study focuses on studying national and international policies that encourage the deployment of broadband networks through government incentives or other mechanisms. To do this, they evaluate the impact of these policies on Spain at a national level or, occasionally, at an autonomous community level.

Similarly, the second group is made up of studies which focus on Spain [28, 30, 33, 29]. The studies included in this group has used two approaches, similar to the ones presented in section 2.1. The first approach studies the government policies and business strategies that have been followed to carry out broadband deployment in certain regions, both nationally and at municipality level [28, 30, 33]. To do this, the studies carry out research into the historical description of deployments, operators and government interventions to enable a technical and market assessment to be made. The studies then conclude with a description of possible future scenarios, providing recommendations, relevant investment factors, deployment strategies and intervention policies.

In particular, the author of [30] studies the importance of a regulatory framework and public policies that encourage intense competitiveness to develop the industry, increase investment incentives and facilitate the diffusion and adoption of access networks, thus having an impact on society and the economy. With a view to the future, the study proposes two alternatives: competition in services, which would reduce entry obstacles, but would also reduce investment in infrastructure, and inter-modal competition (vertically integrated operators), thus resulting in the opposite effect. In any case, the author reaffirms the need for regulatory intervention to combat adverse effects, such as the reduction of the profitability of network deployment in the case of competition in services, and to encourage the deployment of networks that boost economic growth through initiatives.

Furthermore, in [28] and [33] studies are carried out on beginning of the deployment of FTTH networks and the regulatory interventions that have boosted investment in the infrastructures of these networks since 2008. In particular, the authors consider the decisions made by the Telecommunications Market Commission (CMT in its Spanish initials) in 2009 as key in promoting the definition of a wholesale telephone service over its infrastructure between 2010 and 2013. In addition, they highlight the review of this strategy carried out by the National Commission on Markets and Competition (CNMC in its Spanish initials) in 2016, whose objective was to achieve the right balance between encouraging efficient investment and ensuring the maintenance of sustainable competition by studying the profitability of the regions. Thus, solutions were regulated to provide access to competitive municipalities (traditional deployments) and certain non-competitive municipalities (deployments through virtual and disaggregated access). Moreover, these works also address the reasons that encouraged the decision to invest in FTTH networks, mainly by Telefónica between 2010 and 2013, such as the density and average age of the population, the unemployment rate in the region and the infrastructure already in place. As recommendations for the future, the authors of these papers [28, 33] address the relevance of specifying regulatory frameworks at a European and national level as well as public aid plans to increase the level of coverage penetration, especially in areas in which deployment with private investment might not be possible and taking into account the aspects mentioned above. In fact, the authors affirm the value of the CNMC in their consideration of non-competitive municipalities, and conclude that it would be interesting to increase the number of municipalities included.

The second approach is an empirical study on the impact of broadband deployment on social and economic variables and it reveals the deployment strategies at a local or regional level. According to our study of the available literature, and in agreement with other authors [26], there are a lack of studies that apply this method in Spain. Moreover, this deficiency is greater in studies focused on rural areas.

According to our bibliographic research, the only study that considers this method is [29]. In the said study, the authors assess the potential reduction of the digital divide through the deployment of broadband networks in rural regions. This digital divide is studied considering key social and economic variables. This project therefore fills a gap in the empirical studies focused on the rural regions of Spain.

> **For the specific case of Spain, this is a pioneer study in its consideration of the impact of fibre deployment in rural municipalities in a broad and recent time scope (2014-2018).**

Going into a greater level of detail, two of aforementioned articles [33, 29] are the ones that present a greater similarity with the considerations of our study. In [33], the authors analyse a dataset composed of 6,063 Spanish municipalities which have access to broadband services in order to analyse the fibre optic deployment strategy of incumbent operators in the period from 2010-2013. Thus, they study aspects related to the effect of the competition of telecommunication operators to deploy and provide the broadband service and the conditions of the region so that the operator can make the decision to deploy the network.

In particular, **the authors highlight the market size and population density as a positive effect on investment, and the level of unemployment and the percentage of an elderly population as a negative effect**. The methodology followed is based on the use of panel data models, in particular, logistic regressions considering fixed effects[7].

Moreover, the authors of [29] present a study on 63 municipalities considering a dataset from 2002 to 2009. In particular, the municipalities considered are small, with populations of between 500 and 10,500 inhabitants, and generally with a large number of white areas, where there are no deployment initiatives in place.

[7] In regressions with fixed effects, we estimate a different ordinate at the origin (α in a regression)
$Y=\alpha+\beta X$ for each observation, being able to consider different starting points for each municipality.

The authors state that **there is a positive relationship between the decrease in the digital divide, associated with the deployment of broadband, and a consequent increase in the level of education and income and a decrease in the size of the ageing population**.

However, the authors also conclude that this relationship is highly context-dependent, with very different results across regions, autonomous communities, and time periods. As can be seen in Table 2, this study completes this analysis in terms of updating the time periods and the number of years considered, using a sufficiently relevant sample size of municipalities.

Study	[29]	[33]	This Study
Time Range	2002-2009	2010-2013	2014-2018
Geographical Scope	Municipalities and the Autonomous Community	Municipality	Municipality
Quantity	63	6,063	847
Population	500-10,500 (30)	Not specified	1,000-30,000
Methodology	Difference in the differences	Panel Data	Panel Data
Subject	Impact of broadband deployment (>1 Mbps) on reducing the digital divide assessed by socio-economic indicators: education, population density, education level, income and age.	Factors to consider when making decisions on FTTH fibre deployment.	Impact of FTTH fibre deployment on socio-economic indicators. Factors to consider when making decisions on fibre deployment.
Initial Situation	Not specified	High penetration of broadband technologies over xDSL and/or cable (<30 Mbps) and very low penetration of FTTH (<100 Mbps).	High-speed broadband penetration (>30 Mbps) non-existent or below 30% coverage of individual properties.
Results	Positive impact of broadband deployment on reducing the digital divide. Positive, but inconclusive, impact of broadband deployment in income and education levels and a reduction in the number of ageing populations.	In favour of deployment: market size and population density. Against deployment: unemployment and an ageing population.	Impact of fibre deployment on the unemployment rate, business creation and household income[8]. In favour of deployment: population size. Against deployment: ageing population and distance to municipalities with more than 100,000 inhabitants[9].

Table 2: summary of studies applied to Spain (including this study) on the impact of broadband on socio-economic indicators and on deployment decisions.

[8] *For the results of this study about the impact of fibre deployment on socio-economic indicators, see section 3.5*

[9] *For the results of this study on fibre deployment decisions, see section 4 Conclusions*

Experimentation

The general objective of this study is to look into the impact of fibre deployment on econometric and social variables. For this reason, and in line with the context and the aforementioned methodologies, we have opted to work with panel data techniques. The application of these techniques allows the estimation of a regression model that relates to socio-economic variables (unemployment rate, volume of companies, etc.) with the level of fibre optic penetration. Moreover, panel data allows us to consider the effects relative to each region (fixed) and to each year (time). Thus, we can consider the factors common to a region and invariable over time as well as those that are common to several regions, but variable over time. In short, the model allows us to observe the effect produced by variations in high-speed broadband coverage in a region, for example, with regards income or unemployment, both globally and regionally.

The geographical area considered in this study is made up of rural municipalities in Spain. The definition of rural municipality used is that defined by Law 45/2007, of 13 December, for the Sustainable Development of the Rural Environment (see footnote 2.). Thus, **the upper limit of inhabitants of the municipalities included in this study is 30,000, and the population density is always less than 100 inhabitants per square kilometre**. Moreover, due to the limitations of public data availability for very small municipalities, the lower limit of inhabitants for this study has been set to 1,000.

The time range considered is 5 years, from 2014 to 2018 inclusive. The choice of this period is based on two criteria. On the one hand, we have chosen to carry out an analysis of the most recent socio-economic data available and on the other, the greatest increase in fibre optic deployment in Spain has taken place particularly around 2016. Thus, the period from 2014 to 2015 will contain information on the status and initial trend of municipalities without the impact of this deployment, as the vast majority of rural municipalities still didn't have relevant fibre deployment available at that time. The period from 2016 to 2018 provides impact information on the evolution of each municipality where fibre was deployed, while there will still be municipalities where the deployment will have not yet arrived. In this way we can study the panel data from a sample that helps us compare the effects.

It should be emphasised that this study only analyses the impact of FTTH fibre deployment, therefore, it is not intended to measure the differential impact between having fixed data connectivity or not having it, nor between having high-speed connectivity or low-speed xSDL connectivity[10]. In addition, many of the municipalities without high-speed fixed connectivity have high-speed mobile connectivity, which to some extent compensates for the lack of fixed connectivity. For all of these reasons, the positive impacts may be more moderate than expected.

3.1 Data Sources and Variables

The scope of this study includes the impact assessment on social and economic aspects. Because of this, the set of variables considered to analyse this impact includes commonly used indicators in socio-economic and demographic studies, such as employment, the presence of companies from different sectors, income levels, etc. The complete list of indicators used in the study can be found in Annex 5.1. This section presents a discussion of the availability of data associated with these indicators, the information they provide, and how the variables used are treated.

The data has been obtained mainly through four strategies. On the one hand, from open data sources of Official Bodies or which form part of the Spanish Government, such as the Spanish National Institute of Statistics, the Ministry of Territorial Policy and Public Function and the National Social Security Institute. On the other hand, by publicly processing and transforming the available data, e.g., with characteristics such as the distance from the municipal capitals to the nearest motorway accesses, the distance to other municipalities with larger populations or the length of coastline that each municipality has. The third way has been in the contribution of Telefónica's own data, for example, the number of domestic customers, SMEs or Large Companies that use fibre connectivity, the amount of traffic generated by each municipality and, of course, the progress of the fibre coverage deployment undertaken by Telefónica. The possible presence of fibre deployment by other operators, both globally and locally, has also been taken into account. Finally, the fourth source of information has been the provision of data by the AIS-Group.

[10] *In 2018 in Spain, the coverage of networks with speeds higher than 2 Mbps for municipalities of between 1,000 and 50,000 inhabitants was around 90% [51].*

This company combines data from the Census, the Household Budget Survey (EPF in its Spanish initials), the Living Conditions Survey (ECV in its Spanish initials), from the State Public Employment Service, the Employed Population Survey (EPA in its Spanish initials), the National Statistics Institute (hereinafter, INE in its Spanish initials) and from other private partners to provide additional variables that are not available directly or with the necessary geographical or time resolution.

Data from multiple sources have been used:

- **INE, the National Social Security Institute (INSS in its Spanish initials) and other official sources.**
- **Personal data processing.**
- **Telefónica's own data.**
- **Private partners' data.**

Annex 5.1 provides details on the characteristics of all the data used in this study. Each type of variable is described, the source is cited as well as the availability according to the time range and the number of municipalities for which data exists. Data availability is an extremely important consideration. In fact, the lack of information on socio-economic or demographic aspects has led a large number of municipalities to not be included. Despite this problem, the final set of municipalities, which have the necessary information for the study here proposed, is broad enough for the conclusions to be statistically significant. In general, the volume of data that can be discovered decreases in municipalities with smaller populations. This is mainly due to the fact that the economic and organisational effort involved in data collection is often not accessible for small municipalities. In addition, in order for the data to be reliable and comparable, the data collection must be done using a structured and homogeneous methodology throughout the country.

In addition, it is important to note that standardisation techniques have been applied to all variables. This allows the variables to be compared between municipalities and regions as well as with each other, as well as being used as an input for the different statistical processes to be carried out. More details on the standardisations applied is provided in the descriptions on the variables presented in the Annex. For example, applying rates on the working population in the case of the unemployment rate, rates per 1,000 inhabitants in the case of the number of companies or in population movements, standardisation, etc.

The most relevant considerations regarding the datasets considered in the study are included below.

3.1.1 Companies

Information on business activity is used in multiple similar studies [23, 26, 37] because they provide information on the impact on the innovation of processes and production and on competitiveness. Two complementary data sources have been used in this study.

The initial source is the INE, which provides company data for the entire selected time range (2014-2018), although not for all levels of disaggregation of the activity sectors. In particular, for municipalities with less than 1,000 inhabitants, the INE only provides the total number of companies, without disaggregation by sector and, for municipalities with between 1,000 and 5,000 inhabitants, it provides data according to the reduced classification that can be seen in Table 3. In this study, the INE's company data used is classified according to the activity sectors of the aforementioned table.

The alternative data source is the AIS-Group, whose data is used in this study for certain sectors that are particularly relevant and are not offered by the INE for all the municipalities analysed (for example, finance and insurance, agriculture, livestock and fishing).

Activity Sector	CNAE Sections	Description
A2 - Industry	B, C, D, E	Extractive and manufacturing industries; Supply of Electricity, gas, steam and air conditioning; Water supply, sewerage, waste management and remediation activities.
A3 - Construction	F	Construction
A4 - Retail, Transport and Accommodation	G, H, I	Wholesale and retail trade, repair of motor vehicles and motorcycles; Transport and storage; Accommodation.
A11 - Services	J, K, L, M, N, P, Q, R, S	All the companies in the sectors of: Information and Communication; Financial and insurance activities; Property activities; Professional, scientific and technical activities; Administrative and support service activities; Education; Health and social work activities; Arts, entertainment and recreational activities; Other services)

Table 3: Grouping of the activity sectors (according to the National Economic Activity Classification, CNAE in its Spanish initials) used for the creation of variables.

3.1.2 Income, Household Income and Spending, Poverty and Unemployment

Various studies have assessed the influence of broadband connectivity on income and unemployment [13, 1, 29, 30, 32]. With regard to the data relating to income, the data available from the INE at a municipal level is limited to 2015 and 2016. For this reason, the gross income data published by the Tax Agency has been used, since this data covers the greatest time availability, although it is not complete (2014 to 2017), for all the municipalities under study, except for those belonging to the Autonomous Communities of the Basque Country and Navarre, whose income data are not directly comparable with those of the rest of the country's territory.

Moreover, the dataset obtained from the AIS-Group covers four areas of interest. First, monthly household income, both general and for the specific cases of employed households, retired households, etc. This data is similar to the gross income data mentioned above, but provide the value of being more specific and available for the whole range of time and municipalities considered in this study. Second, annual household expenses for communications and education. Third, the level of education of the inhabitants of the municipalities and the number of working, employed and student members of the households.

Finally, the population at risk of poverty, both at a global level and disaggregated into children, adults and the elderly, in addition to the associated probability of defaulting on debts. All this data provides information on different facets of the potential impact on households of the availability of access to improved connectivity opportunities.

The information on unemployment has been included both as totals and broken down by gender and age ranges. This is especially interesting when looking for impact on the most vulnerable sectors of society in this area. In turn, information has been included on affiliations to the General, Self-Employed and Special systems, to complete the overall view of the impact on job creation.

3.1.2 Demographics, Population Variations and Mobility

Since this study focuses on rural municipalities, it is essential to consider the demographic characteristics [35, 29, 26] and population volume [37, 14] in order to assess both the impact of fibre deployment on these characteristics and the mutual influences with broadband deployment. Therefore, data that demographically characterises the municipalities has been used.

On the one hand, public data from the INE has been used to characterise the population size, density, average age, vegetative growth, percentage of foreigners and percentages of population sectors in comparison with the total according to age segments, gender, etc.

On the other hand, two strategies have been used to characterise population movements.

First, with public data from the INE, we have counted the number of new arrivals in each municipality due to moves within the same province, from the same Autonomous Community or from other Autonomous Communities; and in a similar way, with data relating to deregistrations.

Second, based on Telefónica's processed by LUCA[11], a set of indicators have been generated regarding the capacity of each municipality to attract visitors to it. Always using anonymised, aggregated and extrapolated data, three data samples were taken in February, August and November of each year being studied, and the number of non-resident people who travelled to a municipality each day was calculated. Adding this calculation according to working days, weekends and time of the year, several types of indicators are obtained, the most relevant being those that attract visitors for work or tourism purposes.

3.1.4 Telecommunications Infrastructure Data

The main variable in this dataset is that named the "Maximum FTTH coverage", which describes the progress of fibre optic deployment in each municipality. Specifically, this variable indicates the maximum percentage of property units (homes or business premises) with Telefónica fibre coverage that has been reached each year. Using this variable, the granularity achieved in the measurement of coverage is superior to that achieved in other studies that consider average coverage values in large regions, which can present important variations in the access speed in large regions.

In addition, if there have been other operators in the municipalities that have deployed fibre, regardless of whether they are local or global operators has also been taken into account. For this study, it has been decided to not include the municipalities in which this overlap has been detected, as there is no public data available on the deployment of each operator. In this way, the variable that will indicate the level of fibre optic penetration will always be the scope of coverage achieved by Telefónica in each municipality and in each year, this being the coverage that determines the total effective coverage of the municipality.

All variables in the telecommunication services area are also listed and described in Annex 5.1.

3.1.5 Data Analysed but not Used in the Models

Below, we discuss two sets of variables that have been explored and analysed but could not be used in the panel data models of this study for various reasons.

A priori, one of the most interesting sources for analysing the context of the municipalities is the Local Infrastructure and Equipment Survey[12], given that its scope is municipalities with less than 50,000 inhabitants. However, two limitations have prevented the use of this data. The first is that, at the time this study was conducted, data had only been published up to 2017, with 2018 being a significant year to be analysed. The second limitation, and more important than the first, is that there is a considerable number of Autonomous Communities and provinces for which this data is not available. Specifically, this group is made up of the Autonomous Communities of Cantabria, Catalonia, Madrid, Navarre and the Basque Country; the autonomous cities of Ceuta and Melilla; and the provinces of A Coruña, Avila, Cáceres, Huesca, Seville, Valencia, Valladolid, Zamora and Zaragoza).

[11] LUCA is Telefónica's business unit which offers Big Data and Artificial Intelligence solutions to other companies and public organisations (https://luca-d3.com/).

[12] Available on the website of the Ministry of Territorial Policy and Public Function of the Government of Spain (https://ssweb.seap.minhap.es/descargas-eiel/index.php)

Data on the number of patents registered in each municipality has also been explored, but it has been discovered that this does not provide significant information on the models as it is a variable with a very unbalanced distribution. For example, considering the whole group of municipalities selected for this study, in 2016, 100% of patents were generated by 5% of the municipalities.

Other data analysed is that which provides contextual information. This is the case of data related to the geographical location (altitude, distance to the nearest access to a main road or motorway and the length of coastline) or the availability of public services (presence of schools, hospitals and a train station). This type of data does not vary (e.g., altitude), or hardly varies (e.g., number of hospitals) over time, so no relationship with broadband penetration over the period considered (5 years) will be obtained either. Therefore, it has only been used to enrich the characterisation of the municipalities, rather than being incorporated into the impact models.

Finally, the budgetary data for all the municipalities and for all the years within the scope of the study has been downloaded from the Ministry of Finance website and subsequently analysed. The main difficulty arising from using this data is that comparisons between municipalities are complex due to the different ways of organising the competences and public investment policies both between municipalities and between Autonomous Communities. For example, if one municipality has a lower budget than another for a certain concept, it may be because the apparently missing investment is being covered in other ways not accounted for in these budgets. In short, it has not been possible to use deductions regarding relevant information on concepts such as public expenditure on Research, Development and Innovation, Knowledge Management, Employment Promotion and Education.

3.2 Characterisation of the Analysed Municipalities

Spain has a total of 8,126 municipalities. Of these, according to the applied definition of rural municipalities (see footnote 2.), we selected 6,691 municipalities for the study. The data availability constraints described above reduce this set to 1,127, all of which have more than 1,000 inhabitants. Finally, only municipalities that have either not reached 30% coverage of property units with fibre access in the whole period studied, or have reached that minimum of 30% in 2016 or subsequent years, have been considered. We therefore have a sufficient number of municipalities that represent both those that do not receive the benefit of fibre or receive it to a low degree, as well as those that receive it in a consolidated way. At the same time, municipalities where another operator (global or local) has deployed fibre are not included. Thus, the final size of the whole group of municipalities under study (N) is 847. For all of them, the number of periods (T) considered is 5 years (from 2014 to 2018), except when analysing the impact on the average gross income, which is 806 municipalities (without the Basque Country and Navarre) and over 4 years (2014 to 2017).

In general, for both static and dynamic panel data methods, the use of samples with a large number of individuals (N) and a short time period (T) is recommended. The published literature does not define an exact value for N and T. However, some authors have indicated that a suitable value for N should be greater than 100, while the value for T should not exceed 15 and, ideally, be less than 10 in the case of dynamic panels (Roodman, 2008) [41]. Therefore, it is considered that the size and representativeness of the Data Panel used in this study are adequate.

The following is a characterisation of the selected municipalities and a comparison of these municipalities with the set of all the rural municipalities in Spain. This will allow us to have a better knowledge of the different casuistry of the municipalities analysed and to be able to evaluate the scope of the conclusions of this study using better criteria.

3.2.1 Geographical Characterisation

In terms of the geographical distribution of the selected municipalities, Figure 6 and show the number of municipalities for each Autonomous Community, both for the whole set of rural municipalities without the restrictions mentioned above, and for the whole group of selected municipalities.

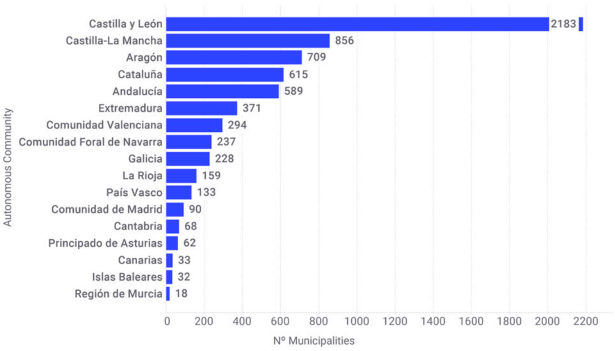


Figure 6: Geographical distribution by Autonomous Community for all rural municipalities in Spain according to the definition set forth in Law 45/2007 on the Sustainable Development of the Rural Environment.

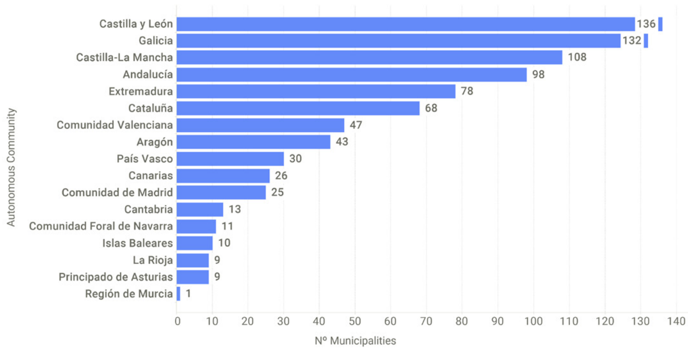


Figure 7: Geographical distribution by Autonomous Community for the municipalities contained in the selected group of municipalities

The most noteworthy aspect of these two graphs is that the whole group of municipalities selected for the study eliminates the over-representation that Castilla y León has in their total number of rural municipalities. On the one hand, the restrictions applied to the selection of municipalities make the criteria for fibre deployment more homogeneous, but, on the other hand, they partly penalise municipalities with less than 1,000 inhabitants, since in this case there was not enough data available for the analysis. Therefore, we are looking at a compromise between the representativity of the aspects of fibre deployment and that of the Autonomous Communities.

Figure 8 and Figure 9 show these same distributions throughout the geographical territory of Spain. In particular, this can be observed in Figure 9, with the areas which are less coloured than those in Figure 8 corresponding above all to areas where the municipalities have a smaller surface area and, consequently, a smaller population, which is coherent with not including municipalities with less than 1,000 inhabitants mentioned above.

Moreover, the map representations suggest other geographical aspects that characterise the municipalities and enable similar sets of municipalities to be created. For example, whether the municipality is coastal or inland, the proximity of the municipality to communication routes or to other municipalities, etc.

In particular, Table 4 summarises the main geographical characteristics which data has been collected for in this study and how the municipalities are distributed around the average values for each variable (for each variable, the percentage of municipalities that fall on either side of its average value is shown). Once again, it is noteworthy that in this table, the particular decrease in municipalities with a smaller surface area than the average, from 69.9% of representativity in the case of all rural municipalities to 36.1% in the whole group of selected municipalities. The increase in the representation of coastal municipalities is also significant, with these apparently being beneficiaries to the detriment of those with less than 1,000 inhabitants, since coastal municipalities tend to have a larger population, even if they are smaller.

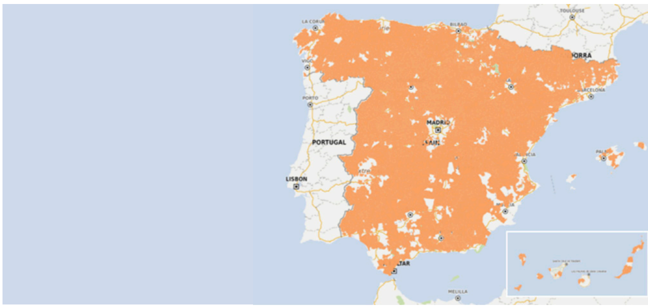


Figure 8: Geographical distribution of all of the rural municipalities in Spain.

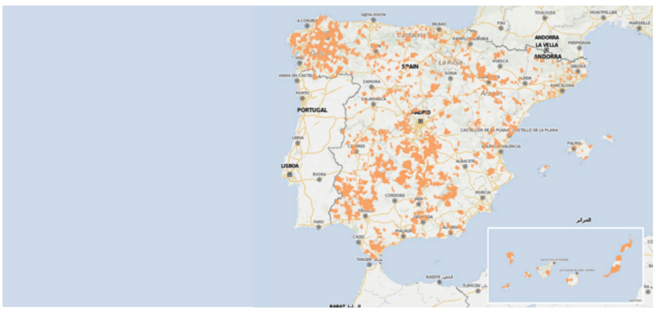


Figure 9: Geographical distribution of the whole group of municipalities selected for this study.

	Nº of municipalities with a coastline	Nº of municipalities with motorway access less than 16 km away	Nº of municipalities with another municipality of more than 100,000 inhabitants less than 58 km away	Nº of municipalities with a surface area smaller than 63 km²
All rural	127 (1.9%)	4,091 (61.3%)	3,522 (52.7%)	4,669 (69.9%)
Whole group of selected municipalities	54 (6.4%)	555 (65.5%)	505 (59.6%)	306 (36.1%)

Table 4: Other geophysical aspects of the rural municipalities compared to the whole group of municipalities selected for the study. The applied threshold values of 16 km, 58 km and 63 km² correspond, respectively, to the average distances of all the rural municipalities to the nearest motorway access, the average distances of the municipalities to the nearest municipality with more than 100,000 inhabitants, and the average surface areas of the municipalities.

In short, despite the limitations, in view of the data and graphs in this section, we can conclude that the whole group of municipalities selected for the study are a reasonable and sufficient representation of the total number of rural municipalities in Spain in terms of their geophysical characterisation and distribution.

3.2.2 Evolution of the FTTH Coverage Characterisation

In 2018, the coverage level of FTTH fibre technology stood at an average of 32.6% of households in rural areas of Spain [53]. This value is lower than in urban regions due to the high deployment costs. However, this value is higher than the average level of coverage in European households, which stood at 14.2% in 2018. Moreover, in recent years, an effort has been made in the FTTH deployment in rural regions of Spain, leading to a considerable increase with 2.5% of homes covered in 2013.

Figure 10 and Figure 11 show the percentage of fibre coverage achieved by Telefónica by year and by population range. The total set of municipalities in Spain has been separated again from the whole group of the 847 selected municipalities.

On the one hand, Figure 10 demonstrates that, in general, deployment starts in the municipalities with the largest populations. From a telecommunication operator's perspective this is natural given the economies of scale for deployment logistics and return on investment. However, deployment made in the smaller population ranges is also considerable, experiencing a remarkable increase for municipalities with less than 5,000 inhabitants throughout the entire time range and especially in 2018 for those with less than 2,000 inhabitants.

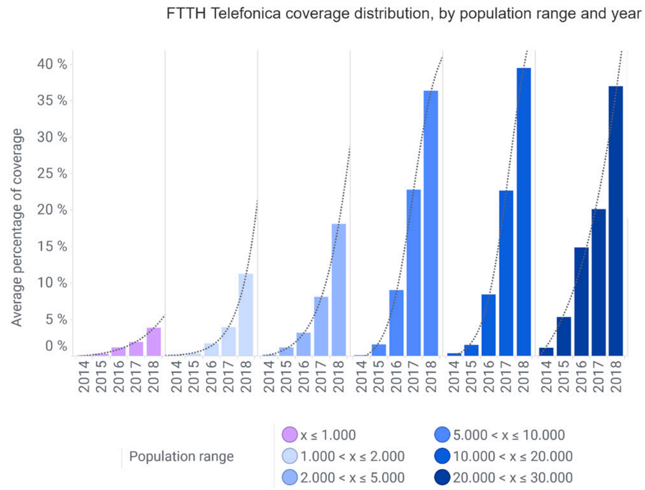


Figure 10: Fibre coverage (percentage of connected property units) by population range for all rural municipalities in Spain, including those with less than 1,000 inhabitants.

By contrast, the comparison with Figure 11 it once again shows that the whole group of municipalities selected for the analysis is a reasonable representation of the evolution of fibre coverage in rural municipalities in Spain when we include municipalities between 1,000 and 30,000 inhabitants. The lower presence of municipalities in 2014 and 2015 shown in Figure 11 is due to the fact that precisely one of the selection criteria has been that the municipalities should reach the minimum of 30% coverage between 2016 and 2018, or that didn't reach this level in the entire period from 2014 to 2018, thus, the analyses of the representativity of the municipalities that reach low levels of penetration or zero penetration is improved.

Specifically, for our complete study group, in 2018 (Table 5) most of the municipalities considered had a population of less than 5,000 inhabitants. Practically none of the municipalities had FTTH coverage in 2014. During the period under study, Telefónica's fibre deployment progressed reaching 345 of the 847 municipalities, although the coverage percentages are not equal, depending on the population range of the municipalities.

It is noteworthy that, once deployment begins in a municipality, on average 30% coverage is reached in 7.6 months, and 85% coverage is reached in 20.3 months. In the case of the whole group of municipalities selected, 30% coverage is reached in 6.4 months, and 85% in 12.5 months.

In terms of fibre deployment costs, it is worth noting that the smaller the municipality, the more representative the cost of bringing the first installation of a fibre node to its vicinity is. Especially if this cost is compared to the cost of subsequently covering the property units in that municipality from the said node. This effect is further accentuated if the dispersion of the municipalities is greater or if the orography is more complex.

3.2.3 Socio-economic Characterisation

Since this study investigates the possible impact of fibre deployment on indicators of unemployment, wealth and business creation, this section focuses on characterising these aspects of the municipalities.

Figure 12 shows the average unemployment rates of the municipalities in 2018 and their standard deviations. There is a general tendency for unemployment to be higher in larger population sizes. This trend is reversed in municipalities having between 20,000 and 30,000 inhabitants. Moreover, the dispersion of the unemployment values, represented by the standard deviation, tends to decrease as the population size increases, indicating that in larger municipalities the unemployment values are more concentrated around the average. The representativity of the values obtained for both variables is equivalent between both groups of municipalities (all rural municipalities in Spain compared to those of the whole group selected for this study).

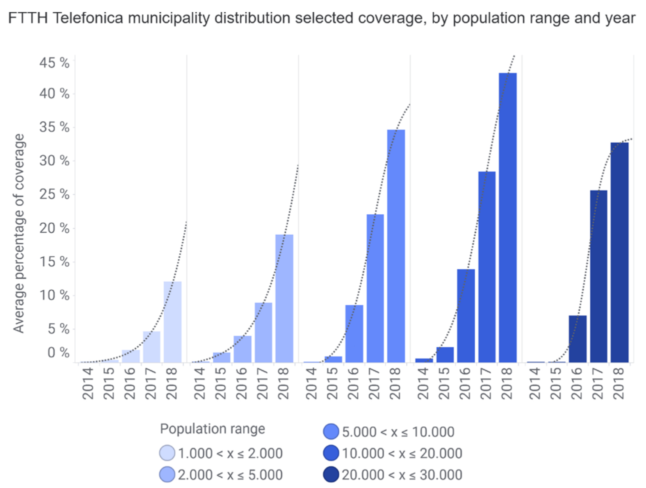


Figure 11: Fibre coverage (percentage of connected property units) by population range for the whole group of municipalities selected for the study. The whole group of municipalities selected does not include less than 1,000 inhabitants due to the unavailability of other datasets.

Population Range	Nº of municipalities	% municipalities having Telefónica's FTTH coverage in 2018
from 1,001 to 2,000	375	23.2%
from 2,001 to 5,000	356	39%
from 5,001 to 10,000	92	68%
from 10,001 to 20,000	20	95.0%
from 20,001 to 30,000	4	100%

Table 5: Population ranges by municipality for the sample considered in the study and average FTTH coverage in 2018. The municipality is considered to have coverage when it reaches at least 30% of its population.

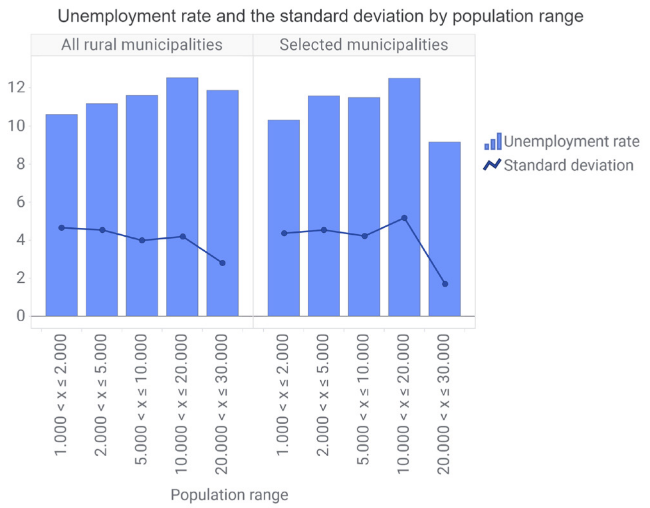


Figure 12: Average unemployment rates in 2018 and their standard deviations by population range. Left: all rural municipalities. Right: municipalities selected for the study. To facilitate the comparisons, municipalities with less than 1,000 inhabitants are not included.

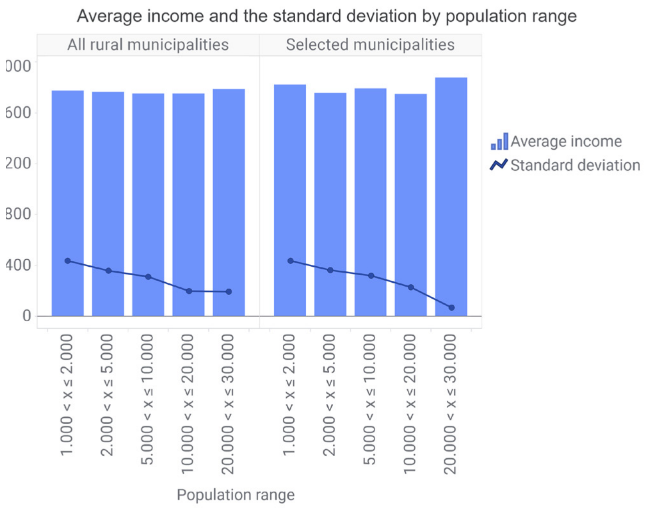


Figure 13: Average household income and its standard deviation by population range. Left: all rural municipalities. Right: municipalities selected for the study. To facilitate the comparisons, municipalities with less than 1,000 inhabitants are not included.

Figure 13 shows average household incomes by population range in 2018 and their standard deviations. It can be observed that, for all rural municipalities as well as for the selected ones, income levels are relatively similar in all population ranges. This may lead to the conclusion that there is a certain equality between the sizes of municipalities, except for the group of municipalities which have between 20,000 and 30,000 inhabitants, where the level is higher. However, the values of the standard deviations show that the greater dispersion of income values the smaller the size of the municipality, which points to the tendency of greater inequalities in smaller municipalities. Once again, the representativity of both groups (all rural municipalities compared with those selected for the study) is equivalent.

3.2.4 Mobility Characterisation

In order to characterise the mobility of the population of rural municipalities, two types of data have been used, as mentioned in section 3.1.3. Firstly, the mobility data provided by Telefónica and, secondly, the data on population balances published by the INE. In Figure 14, the ratio of the non-resident population in that municipality and the number of people who usually travels to it is represented, divided by the population that actually resides there. This is a mobility indicator measured from Telefónica's data and, for the purposes of this illustration, it is based on data for all working days during the month of February 2016. As an example, a ratio of 0.5 means that, for every 2 inhabitants who reside in that municipality, 1 person who does not reside in the said municipality usually goes to work in there (1÷2=0.5).

For clarity, the maps show only an enlarged part of the peninsula in a way that is representative of different typologies of municipalities in terms of their location and economic activity. The municipalities in purple have more than 30,000 inhabitants and are not therefore analysed in this study, but they are represented on the maps to show that, in general, the municipalities with the largest populations attract the working population from neighbouring or nearby municipalities. The shades of blue show different levels of attraction of workers to the municipalities, depending on the ratio of workers who go there per local inhabitant. In fact, the distribution of the blue shades on the maps reflects that this attraction of workers mentioned above works on several levels, since the municipalities in purple are usually surrounded by others which "provide" workers to them but, in turn, these municipalities also receive workers from other ones close to them. In addition, the municipalities in dark blue that do not have a purple municipality nearby are often also surrounded by other ones in a lighter blue, thus forming their own zones of influence that function like those of the municipalities in purple.

This structure with levels of zones of influence represents the well-known flows of a working population that habitually travel to nearby municipalities that have more economic activity. This is often negative for retaining the population or activities; however, it offers great potential if new centres of attraction can be generated. This is probably the case of groups of municipalities in light blue "led" by one or more municipalities in darker blue, without having to have a municipality with more than 30,000 inhabitants nearby.

In turn, Figure 14 shows the comparison with all the rural municipalities in Spain and the municipalities selected for this study. The difference lies mainly in the fact that some of the zones of influence seen in the map showing all of the rural municipalities do not appear among the selected municipalities or are not complete, since the data that has been obtained when making the selection constitutes a sample of the initial map.

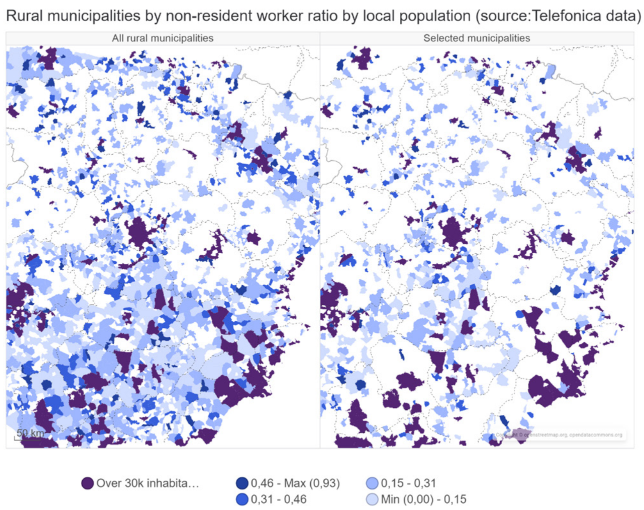


Figure 14: Heat maps that represent the ratio of non-resident population in each municipality that usually travels there for work, divided by the amount of population that actually resides there.
The municipalities with more than 30,000 inhabitants are in purple, and the municipalities according to the ratio are shown in different shades of blue. The regions of influence of some of the municipalities on the municipalities that surround them can be observed.

Figure 15 represents the net population balances of the municipalities according to INE data from 2016. Once again, the municipalities with more than 30,000 inhabitants are represented in dark purple in order to show their relationship with the rural municipalities, as well as a representative enlarged area being represented. Municipalities with a positive population balance are represented in two shades of blue (the darker the blue the greater the net population growth), and the municipalities with a negative population balance are represented in two shades of red (the darker the colour the greater the net population loss).

> **There are rural municipalities with their own socio-economic potential that also benefit the municipalities in their area of influence. However, this potential is fragile.**
>
> **In addition, other areas have specific potential in the tourism sector.**

The most significant aspect of this map is that the areas that are not located near municipalities with more than 30,000 inhabitants generally lose population, to the benefit of ones that have larger ones nearby (the municipalities in blue are normally in the areas of influence of the large municipalities). In addition, the zones of influence that were seen in Figure 14 (clusters of municipalities around a darker-coloured one), in Figure 15 the population loss can be observed, despite their potential.

This suggests that **the potential for rural areas to exert economic traction is limited. In short, an overall flow of population towards the larger municipalities can be observed which leads to the "emptying" of rural areas**.

At the same time, Figure 15 also compares the total number of rural municipalities with the group of municipalities selected for the study, representing a sample of the casuistry that has just been explained.

Rural municipalities by population balance (source Statistics National Institute Spain)

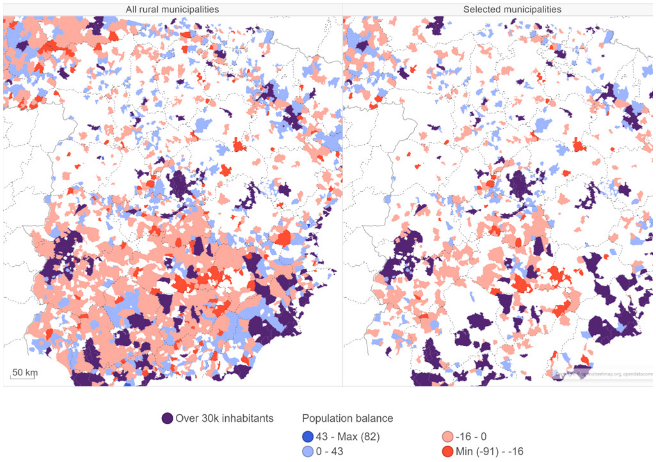


Figure 15: Heat maps representing the population balance of the municipalities according to the INE's data from 2016.
Municipalities with more than 30,000 inhabitants are shown in purple. The two shades of blue show the municipalities with a positive balance and the two shades of red show the municipalities with a negative balance.
Population loss is observed in areas located far away from municipalities and rises in population is shown in areas close to them.

Although it is not within the main objectives of this study, this conclusion about the traction potential of certain municipalities deserves a brief look at in order to provide a more cross-cutting characterisation. For this purpose, the distribution of the municipalities has been observed for the metric of workers who move to work in these municipalities each day represented in Figure 14, and the third quartile level (Q3) has been chosen to separate two groups: those with a high level of attraction of workers, and the rest[13].

In Figure 16 the distributions of different variables separating these two groups of municipalities are represented: those with a metric of non-resident workers greater than the Q3 level mentioned in red, and those with a lower metric in blue. Only the variables that have been found to best characterise these groups have been represented.

Similarly, it has been found that there are a number of key variables that do not differentiate the two groups. For example, both groups have similar characteristics in terms of population density, the proportion of the service sector in comparison with the other sectors, the number of companies per 1,000 inhabitants, the proportions of population by age range and the proportions of companies by size ranges. In these important ways they seem to be very similar, so what makes them different form one another? According to Figure 16, the group of municipalities with the highest level of non-resident workers presents distributions that tend to be more favourable for the following groups of variables:

- Migrations: they have more new inhabitants ("Registrations per 1,000 inhabitants") and, especially, those coming from municipalities in the same province ("Registrations same province"). However, they also have more people who leave ("Deregistrations per 1,000 inhabitants"), which, once again, shows that these municipalities are magnets for attraction but, at the same time, they have a loss of population. In fact, they suffer from greater losses than the other group.

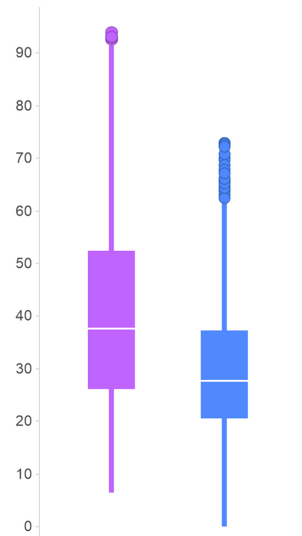
- Geopolitical: they tend to be located closer to municipalities with more than 100,000 inhabitants ("Distance 100 k"), closer to motorway accesses ("Distance from Motorway"), and have a smaller surface area ("Surface Area"). These characteristics point to factors that contribute to a better access to infrastructure, transport, trade and services.

- These aforementioned factors probably contribute to the fact that these municipalities have better metrics on income ("Income per person"), unemployment ("Unemployment rate"), the attraction of workers ("Attraction of workers") and tourists ("Overnight Tourist Stays").

- Similarly, household expenditure on communications ("Spending on communications") and education ("Spending on education") is higher, which can act as both a cause and a consequence of the previous factors.

[13] To find this threshold and for the representations shown in Ilustración 16, box-plot diagrams have been used to represent how the municipalities are distributed according to the values of a variable. For more information see "Box-plot diagrams" in the Glosario.

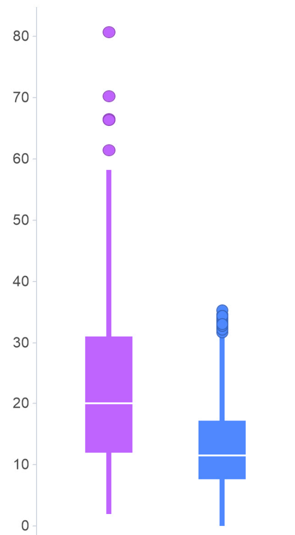
Registrations per 1.000 inhab



	Marked	Unmarked
Count	510	1068
Median	37,60	27,77
Q3	52,43	37,28

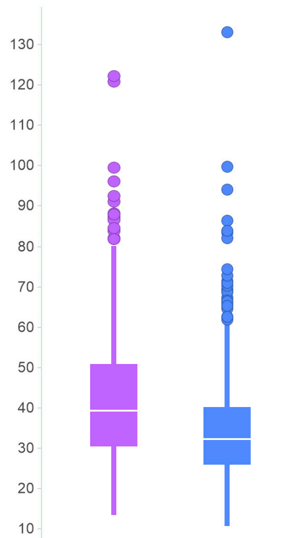
Marked 18:27:06

Registrations same provice



	Marked	Unmarked
Count	510	1068
Median	20,01	11,55
Q3	31,00	17,18

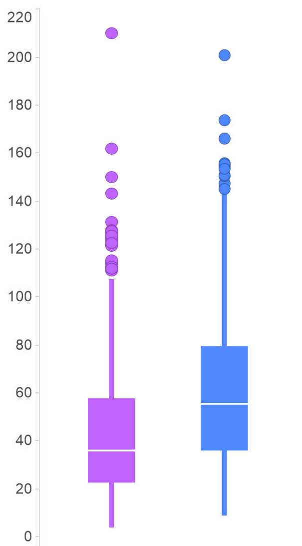
Cancellations per 1.000 inhab



	Marked	Unmarked
Count	510	1068
Median	39,38	32,20
Q3	50,97	40,27

Marked 18:27:06

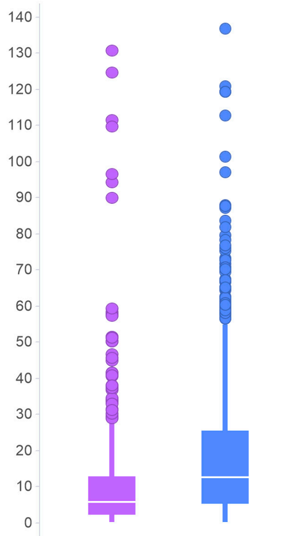
Distance 100k



	Marked	Unmarked
Count	510	1068
Median	35,81	55,43
Q3	57,56	79,64

Marked 18:27:06

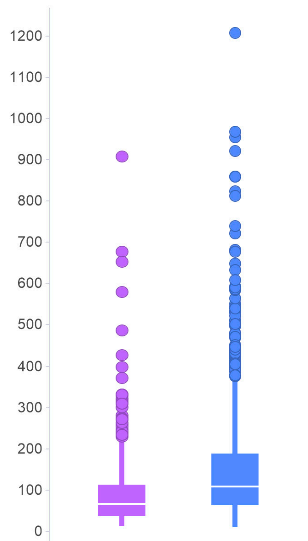
Distance highway



	Marked	Unmarked
Count	510	1068
Median	5,76	12,54
Q3	12,89	25,44

Marked 18:27:06

Area



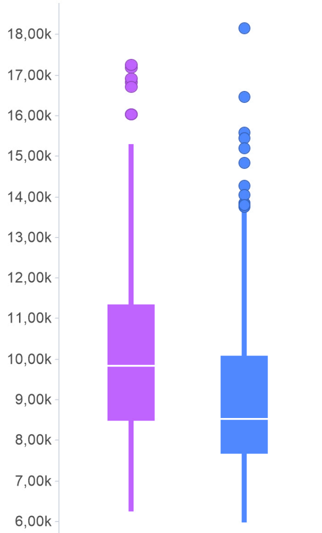
	Marked	Unmarked
Count	510	1068
Median	65,64	108,58
Q3	112,37	188,41

Marked 18:27:06

Figure 16: representation of the distributions of different variables separating the municipalities based on the third quartile (Q3) level of the worker attraction metric represented in Figure 14. The municipalities with a higher metric than that mentioned in Q3 level shown in red, and those with the said metric being lower in blue.

Telefónica

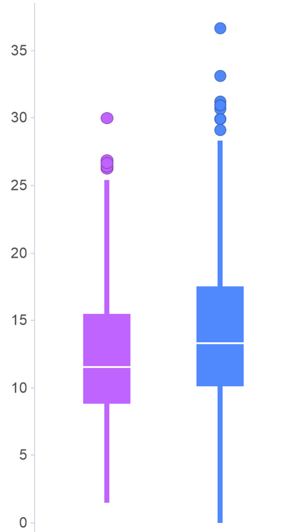
Income per capita



	Marked	Unmarked
Count	510	1068
Median	9840,00	8517,00
Q3	11340,50	10086,75

Marked 18:27:06

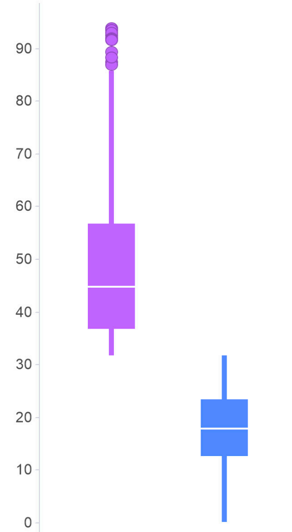
Unemployment rate



	Marked	Unmarked
Count	510	1068
Median	11,54	13,33
Q3	15,50	17,52

Marked 18:27:06

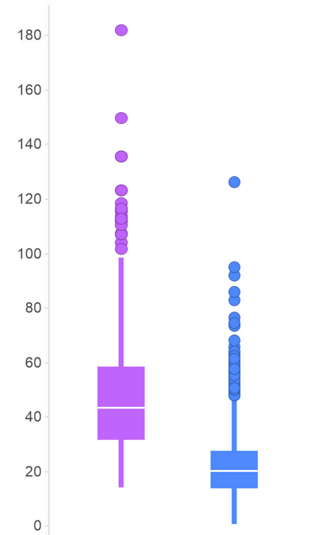
Worker attraction



	Marked	Unmarked
Count	510	1064
Median	44,78	17,95
Q3	56,73	23,48

Marked 18:27:06

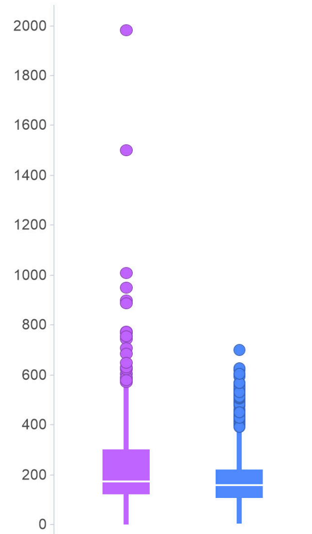
Tourist overnight stays



	Marked	Unmarked
Count	510	1064
Median	43,22	20,20
Q3	58,60	27,53

Marked 18:27:06

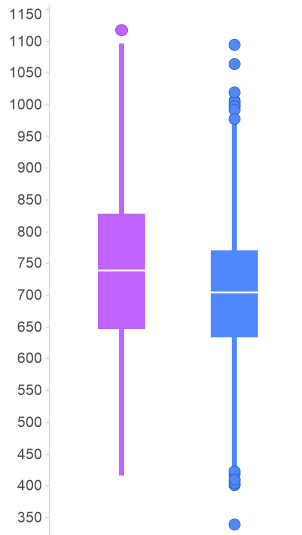
Education expenses



	Marked	Unmarked
Count	510	1068
Median	173,97	157,64
Q3	302,03	222,05

Marked 18:27:06

Comm. expenses



	Marked	Unmarked
Count	510	1068
Median	739,08	704,40
Q3	827,86	770,69

Marked 18:27:06

Finally, Figure 17 represents the ratio of overnight stays of tourists compared with the number of local inhabitants. The pattern of the formation of the groupings of municipalities is similar to that observed in the case of workers' travel habits, although in this case there are new groupings. This means that there are tourist attraction centres which are independent to the labour attraction ones, and that they constitute a specific socio-economic potential for certain rural areas.

3.3 Characterisation of the Relationships between Variables

Section 1.3 describes how impact chains between socio-economic variables work. This inspires the objective of this section, which is to statistically characterise the relationships of dependence and influence that exist between the main socio-economic variables used in this study.

The technique used to carry out this characterisation is that of Bayesian Networks, which enables probabilistic relationships to be found from among all the variables, through the use of conditional probabilities.To do this, the first step is to start from the same representativity of the high penetration group of fibre coverage and that of the low penetration group, since this characteristic is the one that most conditions this study. To achieve this, the two groups have been balanced[14].

The second step is to obtain a simplified set of relevant variables when it comes to deciding whether a municipality belongs to one group or the other. For this purpose, a logistic regression was performed using the variables that showed the best correlation with the "high fibre penetration" category. This step enables us to draw a first conclusion regarding the importance of the variables involved through the regression model. As can be seen in Figure 18, the variables that have shown the greatest importance are the percentage of the population over the age of 65, the total population and the distance to municipalities with more than 100,000 inhabitants. In other words, according to the model, socio-demographic and territorial aspects seem to be even more relevant than economic ones (the latter represented by the variables service companies and average gross income) when classifying whether a municipality will have a high or low penetration of FTTH coverage.

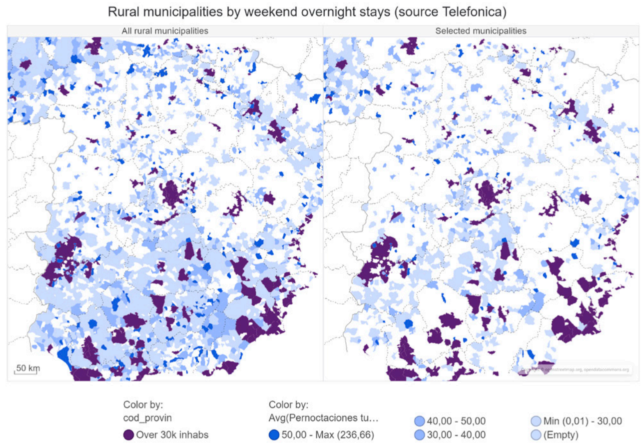


Figure 17: Heat maps representing tourist overnight stays per local inhabitant.
Municipalities with more than 30,000 inhabitants are shown in purple. Data from weekends in August 2016. The darker blue colour corresponds to more overnight stays per local inhabitant.
There are coincidences with the groupings shown in Figure 14 but there are also other areas with specific tourist potential.

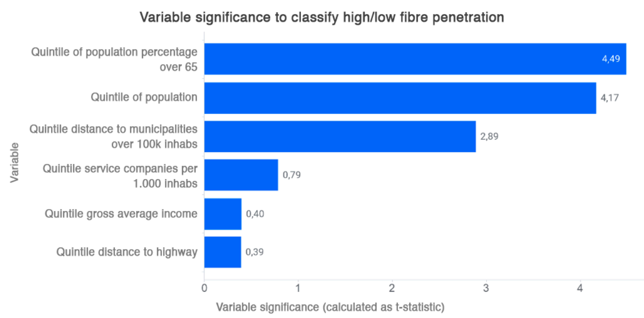


Figure 18: Top 6 variables, ordered by importance, to distinguish between "high penetration" and "low penetration" groups of municipalities in accordance with the logistic regression model. The significance is calculated as the absolute value of the t-statistic (the higher the value, the greater the significance). A greater importance of socio-demographic and territorial aspects rather than to economic ones can be observed.

[14] The purpose of group balancing is to have homogeneous groups in comparison with the number of samples contained in each one. This prevents certain statistical effects derived from one group having a greater representation than another. Therefore, some samples are usually randomly removed from the largest group until the number of samples is equal to the smallest sample group.

The third step before applying the Bayesian Network technique is to address the potential endogeneity of socio-economic variables, as this may distort the study. This is because such variables may condition deployment plans and at the same time reflect the impact of the deployment itself. This may contribute to perpetuating or even increasing the starting conditions of municipalities which are potentially at risk of exclusion in terms of fibre coverage.

Figure 19 shows the distributions of both groups of municipalities ("high penetration" and "low penetration") in quintiles of different variables. For example, in the first graph (distance to municipalities with more than 100,000 inhabitants) it can be observed that the greater the number of the high penetration group (orange), the greater the distance quintile (0 is the quintile with the greatest distance, 4 is the quintile with the least distance) and; in the graph relating to the population over the age of 65, it can be observed that the greater the quintile, the higher the penetration group (0 is the quintile with the highest number of its population over 65 years of age, 4 is the quintile with the lowest number of its population over 65 years of age). This is a more descriptive way of expressing what was already shown in the important points obtained in Figure 18.

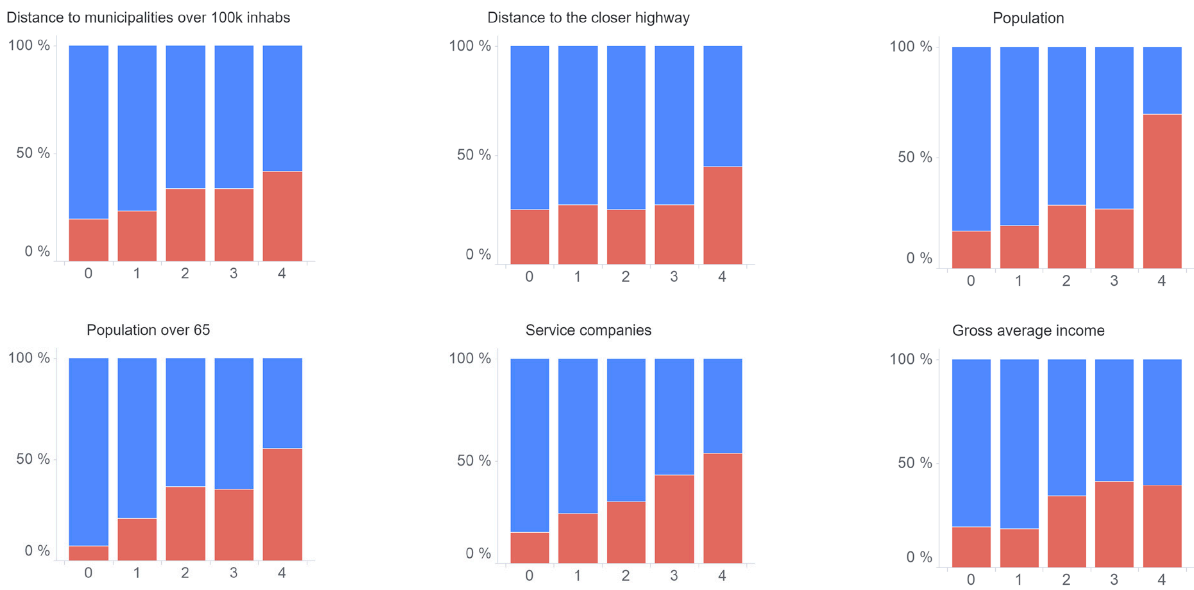


Figure 19: The distributions of the "high penetration" groups (in orange) and "low penetration" ones (in blue) in quintiles of different variables. 0 corresponds to the least favourable value of the variable (less population, further away from the motorway, lower average age, etc.) and 4 to the most favourable value.

It is therefore necessary to ensure that the set of municipalities analysed does not have an a priori bias in terms of fibre penetration. In other words, the a priori probability that a municipality belongs to the high penetration group of fibre coverage or to the low penetration group is the same regardless of the values that the municipality has for other variables.

For this purpose, the quintile scale seen in Figure 19 and each variable is weighted by the importance given to it according to the regression model. The weighted values for each variable are added together and assigned to each municipality. In this way, the municipalities with the highest score are, a priori, better positioned for fibre deployment. In other words, this score allows us to gauge whether or not a municipality has an a priori bias that makes it more or less interesting for a telecommunications operator to deploy fibre in it.

Figure 20 shows the distribution histogram of this score for both groups.

Figure 21 shows the corresponding box-plot diagram where it can be observed that both distributions are reasonably separated but there is an overlap zone between the 2 score averages [21 and 33]. These averages are the ones chosen as thresholds to determine that the score values between them are not conclusive enough to assign their municipalities to one group or another. Applying these thresholds results in the set of 73 "high penetration" and 71 "low penetration" municipalities that is represented Figure 22. This set manages to minimise the a priori bias and, therefore, the endogeneity, and it is the one that will be used next for the analysis with Bayesian Networks.Therefore, the last step consists of using this group of municipalities to calculate the Bayesian Networks, with their subsequent analysis enabling conclusions to be drawn regarding the relationships between the variables.

Bayesian Networks enable us to extract information about conditional probabilities. In the case of the study, since the municipalities are assigned the corresponding quintiles for each of the variables, for example, it is possible to calculate the probability that the municipalities belonging to the *M* quintile of the X_k variable, which in turn belong to the *N* quintile of the X_l variable. The relevance of this feature of Bayesian Networks when characterising relationships between the variables is that they enable dependence relationships to be established, and even probabilistic causality, between the different variables.

For example, calculating the Bayesian Network by setting the fibre coverage variable as a class, we obtain the graph shown in Figure 23. The existence of a link between two nodes represents the probabilistic causality relationship between the two variables. In other words, when an arc connects variable A with variable B (A→B), this means that the probability of B occurring is conditioned by the probability of A occurring. Specifically, in this graph it is interesting to see the following chains of impact:
Coverage → Unemployment → Companies → Population
Coverage → Unemployment → Income

These sequences show how coverage can influence wealth (Income) and business creation through its simultaneous influence on unemployment. According to the diagram, this influence can also be direct, without Unemployment having to be a mediating influence. As can be seen in section 3.5, these relationships between coverage and the rest of the variables are completely in line with the results obtained in the panel data experiment.

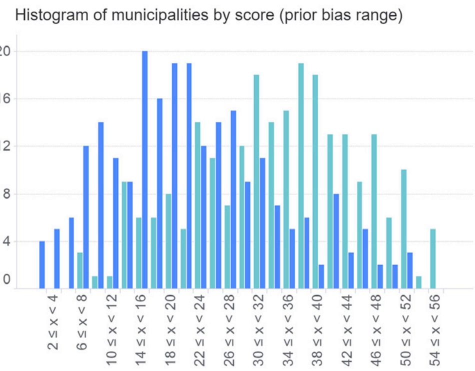
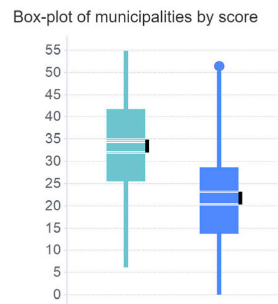


Figure 20: Histogram of the "high penetration" (orange) and "low penetration" (blue) groups by score range. The score allows us to gauge the a priori bias of a municipality which has greater coverage.



	Alta cobertura	Baja cobertura
Count	237	239
Avg	33,47	21,69
StdDev	11,16	11,49

Figure 21: Box-plot diagram of the "high penetration" and "low penetration" groups by score range.

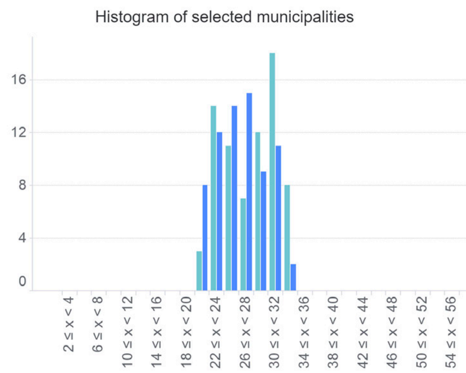


Figure 22: section of the Figure 20 histogram selected to minimise the bias.

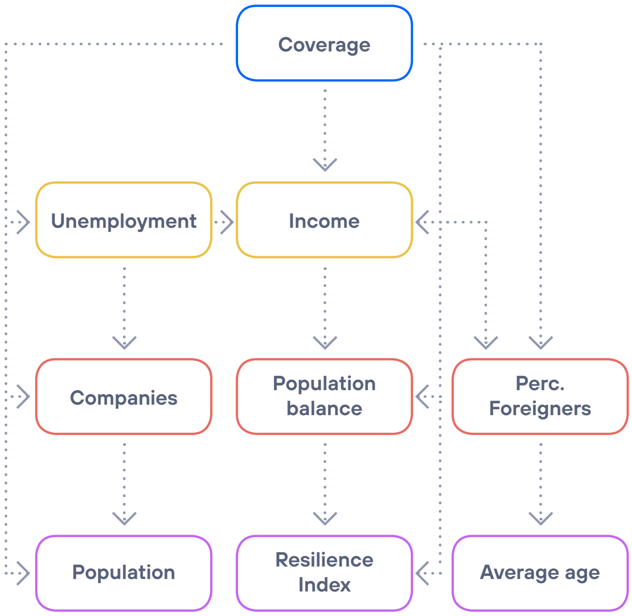


Figure 23: Bayesian Network diagram, establishing the fibre coverage variable as a class. Noteworthy is the Coverage → Unemployment → Companies → Population and the Coverage → Unemployment → Income sequences.
The presence of a link means that there is a relationship of dependence between the two variables. The network has been calculated using data from 2016, which is the central year of the entire time scope of this study.

Another case of interest is the one represented in the graph of the Figure 24 which shows the Bayesian Network setting the Gross Income variable as a class. In this case, the most relevant characteristics are as follows:

Income → Unemployment → Companies → Population
Income → Unemployment → Resilience Index / Average Age

These sequences show how the variation in levels of wealth (Income) can influence aspects such as stabilising the population, the average age of the population and even the diversification of employment sectors (Resilience Index[15]), both directly and through the improvement of Unemployment rates and the creation of companies.

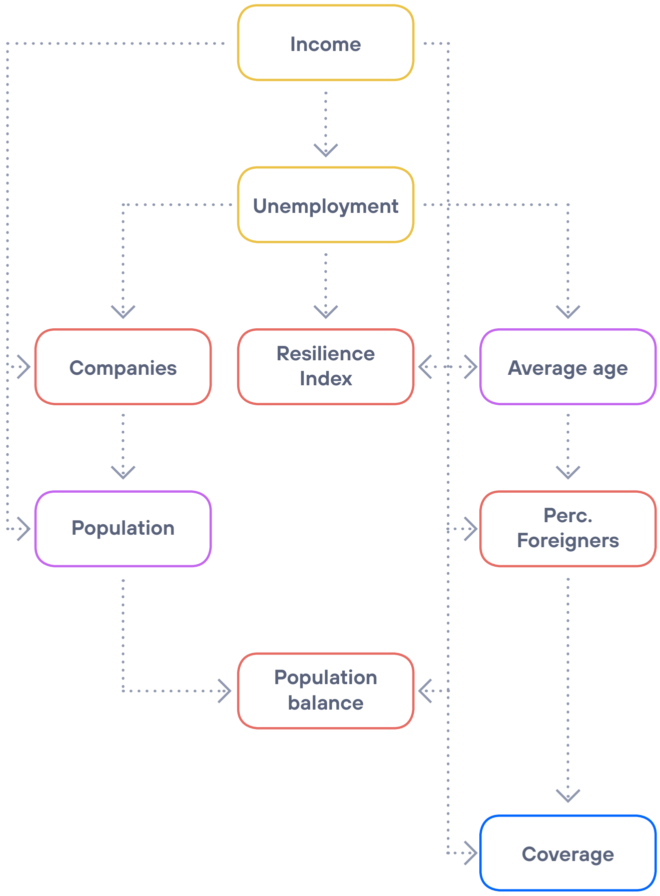


Figure 24: Bayesian Network diagram establishing the gross income variable as a class. Noteworthy is the Income → Unemployment → Companies → Population and Income → Unemployment → Resilience Index sequence.
The presence of a link means that there is a relationship of dependence between the two variables. The network has been calculated using data from 2016, which is the central year of the entire time scope of this study.

If the calculation of these networks is iterated by alternating the variable that is set as a class, multiple networks like the ones just discussed are obtained. In these networks the recurrent appearance of the relationships described above between Coverage, Companies, Unemployment and Income can be observed, showing that this is the most characteristic chain of impact that arises from the relationships between the variables considered in this study. This, in addition to being a conclusion in itself, further reinforces the relevance of assessing the impact of fibre deployment on the other three variables using panel data, in which case it would also be necessary to address the endogeneity problem.

[15] The Resilience Index is an indicator created by researchers at the Polytechnic University of Madrid Innovation Centre in Technology for Development. This indicator summarises the composition by employment sectors into a single value in the said municipality compared to the average of the other Spanish municipalities. This gives a measure of a municipality's vulnerability to a crisis in a particular sector resulting from the low diversification of employment.

3.4 Panel Data Methodology

Having presented the data in section 3.1 and characterised the whole group of municipalities to be used and the pattern of the relationships between variables, the next objective is to build a panel data model that extracts information on the impact of fibre optic deployment on indicators such as business creation, a decrease in unemployment and an increase in wealth. As has been mentioned, this model will consider 847 municipalities with annualised information between 2014 and 2018.

3.4.1 Optimisation of the Set of Variables

First of all, it is necessary to reduce the number of variables used by optimising the information they provide, given that there is a wide variety of indicators and the information provided by some of them is redundant. In turn, in order to select the variables to be used in the models, the forward selection technique has been used, consisting of adding variables iteratively by adding the variable that provides the greatest significant improvement to the model result based on the minimisation of the root-mean-square-error (RMSE) and the maximisation of the representativity of the panel data models used (R^2 adjusted) in each iteration. When running these iterations, a Pearson correlation analysis has also been carried out between the variables selected, so that, at the same time, multicollinearity[16] problems, as well as preventing the over-representation of groups of variables. Although there are studies that seek to optimise and complete techniques such as these [48, 49, 50], in this case the results of the models with the selected variables have been sufficient and relevant with a reasonable computational cost.

With this optimisation, the set of variables to be used in the models is reduced as indicated in Table 6. The process is carried out for each variable on which the impact will be estimated, which also simplifies the interpretation of the results obtained.

[16] *Multicollinearity occurs when, in a model whose objective is to explain the pattern of a Y variable from another set of explanatory X_k variables, there is correlation between several or all the explanatory variables.*

	Initial Variables	Variables after applying Forward Selection and the Correlation Threshold
Unemployment Rate	79	11
Service Companies	93	13
Household Income	91	7

Table 6: Reduction of the initial set of available variables using "Forward Selection" and the correlation threshold for each of the variables on which the impact of fibre deployment is estimated.

3.4.2 Panel Data Model Application

The next step is to address the specific problem of panel data based on its usual initial formula:

$$Y_{it} = \alpha_i + \beta_k X_{kit} + e_{it}$$

Y is the variable to be explained (e.g., the Unemployment Rate) from the regression of the explanatory variables X_k (e.g. own fibre coverage, household expenditure, etc.) through the panel data. In this case, k refers to each of the variables to be considered, i refers to each of the municipalities, and t refers to each year. With α_i being the intercept of the regression for each municipality, and e_{it} models the regression errors.

From this point different problems arise which have to be addressed. The most relevant ones that have been faced in similar studies [13] and that also apply to our case are the effects of time and endogeneity. In [46] a comparison of the theory of the various techniques that can be applied in different cases is approached in a very strict way, considering techniques such as the first differences, delays of independent and/or dependent variables ([13] proposes and applies some of these techniques whilst making some variations to them), or the use of instrumental variables as very useful options in this context.

In our case we have used three panel data techniques that address these issues in different ways: fixed effects, first differences and "Two Stage Least Squares" (TSLS).

For the possible choice of the fixed effects technique, it is necessary to evaluate whether the intercepts α_i, which represent the starting point of each municipality in the regression that will estimate the model, can be considered constant for each municipality (fixed effects) or whether it is necessary to consider α_i as a random variable (random effects)[17].

In statistical terms, estimation with fixed effects is more consistent[18], and random effect estimation is more efficient[19]. The Hausman Test allows us to check if the estimates of both types of effects are significantly different, which is the case in our study, and subsequently the fixed effects model should be followed because it is more consistent.

Moreover, the first differences technique is particularly tolerant to the existence of endogeneity among the variables. Starting from the initial formula, it consists of calculating the differences with the delayed variables, so that the terms which are considered constant are cancelled out and it is no longer necessary to estimate the α_i. The formula is as follows:

$$(Y_{it} - Y_{it-1}) = \beta_k (X_{kit} - X_{kit-1})' + (e_{it} - e_{it-1})$$

In our case the results obtained in the first difference models present very similar significance values and coefficients to those obtained by simply applying the fixed effects model.

Finally, the TSLS technique allows us to address the existence of endogeneity in a more specific way. Specifically, we have considered this possibility for the variable that describes the deployment of fibre optics ("Maximum FTTH Coverage"), because, as we have seen in the previous section, this variable usually has a bidirectional dependence with the socio-economic variables which we measure the impact on (Business Creation, Unemployment Rate and Income). The specific technique used is TSLS with fixed effects to estimate panel data using instrumental variables. Instrumental variables are ones that replace the variables that present this type of problem. In the first of the two TSLS phases, the fibre variable is estimated based on another variable ("instrument"):

$$\tilde{F}_{it} = \alpha'_i + \beta' Z_{it} + \beta'_k X_{kit} + e'_{it}$$

Where $\tilde{F}$ the estimate of the fibre coverage values as a function of the variable instrument Z, and the rest of the non-endogenous variables, X_{kit} (in our case, the rest of the variables of the model). In the second phase, these estimated values of $\tilde{F}$ are used to estimate the original dependent Y variable:

$$Y_{it} = \alpha_i + \beta_k X_{kit} + \beta_f \tilde{F}_{it} + e_{it}$$

Where X_k no longer includes the fibre coverage variable as it is represented by $\tilde{F}$.

In our case, we have used the population percentage in white areas in each municipality as an instrument for the fibre deployment variable. On the one hand, this helps to solve the possible existence of misspecification of the model due to endogeneity, and, on the other hand, it makes us accept the limitation that this variable is only available for the years 2016 to 2018, which, fortunately, still includes the time range of greater progress of fibre deployment.

The validity of using the white zone percentage as an instrument is based on two conditions. First, the exogeneity condition, whereby the instrument variable would not directly affect the dependent variables but rather it would through the instrumented variable. In this case it is reasonable to think that the number of companies, unemployment or income are not directly determined by the percentage of white area, but there is a relationship between this and fibre deployment. Second, the relevance condition, which can be verified using a regression of the percentage of coverage of the white area (modelled as $\tilde{F}_{it} = \alpha' + \beta' Z_{it} + e'_{it}$), so that the null hypothesis H_0: $\beta' = 0$ can be posed, and if this hypothesis is rejected, there is evidence of the relevance of the percentage of white area as an instrument.

[17] A more sophisticated analysis (see [46]) explains that the individual effects can always be considered as random without loss of generality, thus, what is relevant is whether or not the effects correlate with the Xk variables. In any case, the Hausman Test makes it possible to make a distinction between the models.

[18] In statistics, "consistency" refers to the fact that the estimated value converges towards the true value when the number of samples converge towards infinity.

[19] In statistics, "efficiency" refers to the fact that an estimator may have more variance (which would be less accurate) than another.

In our case, the significance is tested with a confidence level of 99.99%.

As in the case of the first differences technique, we have found that, in the models based on the conclusions from which this study are drawn, the significance values obtained by applying or not applying the TSLS technique are similar. Therefore, only the results of using the fixed effects models are presented below, since this model is statistically more consistent.

3.5 Results of the Panel Data Models

In the results tables below, the variables presented are placed in order, going from the most to the least significant and the p-value is included as an indication of the significance of the variables in the models obtained[20].

3.5.1 Overall Unemployment Rate

Table 7 shows that the estimated impact of fibre coverage on the Overall Unemployment Rate. According to the coefficient and significance obtained, there is a 0.008% decrease in the unemployment rate for each percentage point of the progress of fibre deployment with a 99.9% confidence level. This would mean that reaching 100% coverage in a municipality would reduce the unemployment rate by almost 1% (0.8%). This model has obtained an R2 of 0.63.

> **Reaching 100% coverage in a municipality would mean a reduction of almost 1% in the Overall Unemployment Rate.**

Variable	Coefficient	P-value
Annual spending on communications per household	- 0.004	<0.001 ***
Personal Income	- 0.004	<0.001 ***
Companies Services per thousand inhabitants	- 0.103	<0.001 ***
Maximum FTTH coverage (percentage of property units covered)	**- 0.008**	**<0.001 *** **
Companies >50 M Sales per thousand inhabitants	- 2.784	<0.001 ***
Affiliates to Social Security Special Schemes per the working population	- 0.265	<0.001 ***
Companies <10 Employees per thousand inhabitants	- 0.145	0.001 ***
Probability of being in default for debts	0.039	0.001 ***
Population density	- 0.088	0.023 **
AIS - Agriculture, Livestock and Fishing Companies per thousand inhabitants	- 0.169	0.101
Construction companies per thousand inhabitants	- 0.037	0.112

Table 7: Result of the fixed effects model on the Impact on the Overall Unemployment Rate.

[20] *See the meaning of the concept of the Level of significance in the Glossary (section 5.2).*

According to Table 1, enabling connectivity has an impact on unemployment in several of the results from the literature reviewed in this study ([1], [4], [5], [6], [9], [12], [13], [23], [25], [55]). Despite dealing with different typologies of the analysed areas, studies [9] and [25] are particularly aligned with the result obtained here, since [25] presents a 1% increase in employment by enabling connectivity and [9] presents a 2.9% increase in employment by enabling 100% coverage.

In turn, Table 7 also shows that this impact is favourably influenced by the emergence of companies, especially those in the Services sector, those with fewer than 10 employees and those with sales of more than €50 M. Evidently, the appearance of high turnover companies can have a very high impact on the Unemployment Rate (note the magnitude of the coefficient of the model in this case), although it is not something that happens frequently (in this case it has been observed in 10.5% of the whole group of municipalities analysed). In turn, although it is small, the impact of the creation of especially small SMEs (represented by the Companies <10 Employees per thousand inhabitants variable) on unemployment is also consistent.

The appearance of all these variables together suggests the existence of synergies between the traction that both SMEs and large companies generate in employment and the catalytic effect that fibre deployment can have on this process, as well as other variables such as personal income or spending on communications. The latter variable would confirm the use of deployed communication infrastructures such as fibre. Furthermore, it should be noted that the impact of fibre deployment can also be seen in the decrease in unemployment among women under 25 and men between 25 and 45, although with a lower level of confidence. This, together with the clear significance obtained for Overall Unemployment, suggests that the tendency of the impact on unemployment at lower ages, especially for women really exists.

3.5.2 Number of Service Companies

The effect on the number of companies in which fibre deployment has been found to be significant is specifically in the Services sector[21]. The specific results of the model can be seen in Table 8.

The impact estimated by the model is an increase of 0.006 companies per thousand inhabitants for each percentage point of progress in fibre deployment, with a confidence level of 99.6%. This would mean that reaching 100% coverage in a municipality of 5,000 inhabitants would facilitate the creation of 3 new service companies. This model has obtained an R^2 of 0.39.

> **Reaching 100% coverage in a municipality of 5,000 inhabitants would facilitate the creation of 3 new service companies.**

[20] *According to section 3.1, this sector includes all the companies from the Information and Communications, Financial and insurance activities, Property-related activities, Professional, scientific and technical activities, Administrative and support service activities, Education, Health and social work activities, Arts, entertainment and recreational activities, and other services.*

Variable	Coefficient	P-value
Employment Resilience Index	- 0.346	< 0.001 ***
Companies <10 Employees per thousand inhabitants	0.342	< 0.001 ***
Construction Unemployment Rate	- 0.269	<0.001 ***
Annual spending on communications per household	0.002	0.001 ***
Affiliates to the Social Security General Scheme per working population	0.033	0.002 ***
Maximum FTTH coverage (percentage of property units covered)	0.006	0.003 ***
AIS - Property Companies per thousand inhabitants	0.733	0.003 ***
Probability of being in default for debts	- 0.036	0.005 ***
AIS - Finance and Insurance Companies per thousand inhabitants	1.042	0.012 **
Percentage of children in poverty	- 0.010	0.028 **
Personal Income	0.001	0.029 **
Percentage of women	0.273	0.037 **
AIS - Technical companies per thousand inhabitants	0.440	0.126

Table 8: Result of the fixed effects model for the Impact on the number of Service Companies.

This result is aligned with the findings obtained in [13], which, according to Table 1, analyses municipalities in France with more than 2,000 inhabitants and shows an impact of 3% on the creation of companies.

In addition to the intuitive agreement of the impact on the creation of companies by variables such as worker affiliations and unemployment, it is interesting to highlight the favourable effects of the percentage of women in the municipality and the number of companies with less than 10 employees. Similar to the point regarding the effects on the Overall Unemployment Rate, the importance of SMEs as drivers of employment and their synergies with aspects such as fibre coverage and spending on communications arises here.

This point is complemented by the significance of the Resilience Index, whose contribution to the model means that the worse the resilience of a municipality is in terms of its diversification of employment in different activity sectors (which happens when the value of the variable is further from 0 in its absolute value), the fewer the companies that will be created. In this case, the model shows that when municipalities have a low diversification of employment sectors, they have greater difficulty in creating new companies. Additionally, by contrast, the better the resilience of a municipality, the more effective the contribution of other variables are (including fibre deployment) in the creation of service companies.

3.5.3 Income and Revenue

There are several factors relating to the income level of the population regarding which favourable results have been obtained in terms of the impact that fibre deployment can have on them: Average Gross Income, Personal Income, Household Income, and Retired Household Income.

Telefónica

Intuitively the concordance of the impact on these variables is reasonable, however two aspects should be noted. Firstly, at the time this study was conducted, the data on Average Gross Income was only available homogeneously for all municipalities for the years 2014 to 2017, therefore the time scope is different from that used throughout the rest of the analysis. Second, the Average Gross Income data is based on presented tax returns, it is therefore an objective and accurate measurement but it does not represent the entire population (not everyone who has income is required to present a tax return).

However, the Income data is based on the Household Budget Survey and the Living Conditions Survey and, although approximations are used to calculate values by municipality, it completes the picture of the population with the full time range going from 2014 to 2018.

In short, as the results of the panel data models are similar and consistent for these variables, we present the specific result for Household Income below as it has a greater time scope.

According to Table 9, the impact estimated by the model is an increase of 0.634 in Household Income for each percentage point of progress in fibre deployment, with the confidence level at 99.93%. This model has obtained an R^2 of 0.5.

> **Achieving 100% coverage in a municipality would mean that a typical household income of €1,600 per month would increase to €1,663.**

Variable	Coefficient	P-value
Percentage of students per household	9.916	< 0.001 ***
Unemployment Rate	- 14.281	< 0.001 ***
Average age	49.206	< 0.001 ***
Maximum FTTH coverage (percentage of property units covered)	0.634	< 0.001 ***
Probability of being in default for debts	- 3.789	< 0.001 ***
Industry Unemployment Rate	- 10.807	0.131
Registrations of foreigners per thousand inhabitants	1.469	0.231

Table 9: Results of the fixed effects model for the Impact on Household Income.

The impact estimated by this model means that a typical average household with a monthly income of around 1,600€, would increase to 1,663€ once their municipality reached 100% coverage, which would be an increase of 3.9%.

This impact is aligned with several of the findings in the literature reviewed ([4], [5], [12], [20], [25], [29] and [31]) summarised in Table 1. Specifically, [31] estimates a 0.025 increase in GDP for every 1% increase in broadband penetration. In the case of this study, Income and Revenue are indicators equivalent to the GDP, since this data does not exist at the municipality level.

3.5.4 Population Density

In the results of the panel models we have also found an impact on the population density variable. In this case the coefficient obtained for the fibre coverage variable is not a significant magnitude (0.003 inhabitants per Km^2 for each percentage point of fibre deployment), but the significance value is (99%), thus demonstrating the existence of a relationship between the two, although it is probably an intermediate and/or long-term effect.

> **Reaching 100% coverage in a municipality would mean the recovery of the population lost in one year.**

In any case, there are two interesting aspects of the effect on population density that should be noted.

On the one hand, rural municipalities in Spain lose on average 37 inhabitants each year, therefore, the impact of reaching 100% fibre coverage for a typical municipality of 5,000 inhabitants and with a 134 Km^2 of surface area, would be the recovery of about 40 inhabitants, which despite the low magnitude in relation to the coefficient, is equivalent to the population lost in one year.

On the other hand, other variables arise that impact population density together with fibre deployment. Although this study does not attempt to make an exhaustive analysis of the problem of depopulation, using the combination of these variables it can be deduced that population density is favoured by the confluence of all or some of the following factors:

- A greater presence of adults, young people and children.
- A better access to the services sector, including retail and accommodation, due to the availability of companies in these sectors. In turn, in view of the aforementioned results, this presence of companies influences employability.

> **Other factors which encourage population density:**
>
> - **Lower average age.**
> - **Greater presence of service, retail, transport and accommodation companies**
> - **Increased household spending on communications**
> - **Higher household incomes**
> - **Higher percentage of foreigners.**

- Increased household spending on communications, demonstrating a more intensive use of the associated media.
- Higher household incomes, which in turn has a positive effect on the other factors related to consumption.
- A higher percentage of population of foreign residents.

3.5.5 Diminishing Returns

Finally, it has also been observed that, in all the cases mentioned (Overall Unemployment Rate, Service Companies, Household Income, etc.), when the models are executed separating the municipalities in two groups of high and low fibre penetration (less than 30% and more than 30% coverage respectively), the coefficients obtained are higher for the low penetration group. This highlights a "diminishing returns" pattern of the impact of fibre deployment.

> **Diminishing Returns:**
>
> **The first steps of fibre deployment in a municipality have a greater impact than completing the coverage up to 100%.**

In other words, the impact of fibre deployment being constantly relevant, is more so in its initial deployment (for example, when the coverage goes from 0% to 30%) than it is once the coverage of the municipality has been completed (for example, when the coverage goes from 60% to 90%).

Conclusions and Recommendations

Telefónica

4.1 Conclusions

Broadband networks can be a key factor for socio-economic growth. In fact, countries around the world have adopted a national plan with ambitious targets to deploy networks that provide broadband access to every household. Among these countries are those belonging to the European Union, including Spain.

As a result of this, numerous studies have been published looking into the impact of broadband deployment on economic and social variables, such as employment, income and population density. Although some of these studies conclude that there is no impact or even a negative impact on these variables, most of them show positive effects, despite the fact that the typology of the areas analysed in each case are very different (see Table 1).

> **The results of most of the literature published are in line with the impacts detected in this study, in particular with regards to:**
>
> - **Unemployment Rate**
> - **Companies**
> - **Income**

Furthermore, there is a scarcity of studies that focus their efforts on looking into the impact of broadband in Spain, particularly studies focused on rural regions. In fact, for Spain there are no references of studies that attempt to quantify the impact of broadband deployment at a municipal level or that specifically consider rural municipalities.

This is something that has been addressed in this analysis and whose results are in line with the results of other studies in similar fields. For example, in [9] and [25] impacts on unemployment are presented, whereas in [13] impacts on business creation are presented, and [31] shows the impact on the GDP. Consequently, this study complements the current literature by looking into the impact on social and economic variables of broadband deployment in rural Spain.

In addition to the parallels explained between this study and the literature which has been reviewed, the analysis of the relationships between the variables carried out in section 3.3 confirms, from a probabilistic point of view, that the relationships between the variables analysed (Coverage, Unemployment, Companies and Wealth), thus providing further support for the conclusions presented in this section.

Moreover, this analysis may be of interest when formulating joint policies for the digitalisation of the rural environment. These policies must combine the perspective of deploying connectivity with that of taking advantage of the opportunities that this connectivity makes possible. In this way it will be possible to maximise the effectiveness of the different types of investment in rural development.

The methodology used is based on estimations using panel data, reinforced with the application of endogeneity minimisation techniques between the variables, so that the results obtained have statistical rigour and reasonable significance for the **scope of the rural municipalities analysed (847) throughout the years going from 2014 to 2018.**

Generally, the models in this study conclude the significance of the impact of fibre deployment on the **Overall Unemployment Rate, the Creation of Service Companies and Household Income.**

According to these models, **reaching 100% fibre coverage in a municipality would mean:**

- **A decrease of 0.8% in the Overall Unemployment Rate in the said municipality.**
- **An increase of three Service Companies (for the case of a municipality with 5,000 inhabitants).**
- **An increase of 63 euros in monthly Household Income (a 3.9% increase of the average monthly income going from €1,600 to €1,663).**

A significant impact on **population density** can also be observed.

Telefónica

Of course, these impacts are not generated purely due to the isolated fact of fibre being deployed. The results obtained suggest that there are several factors that combine synergies when it comes to producing a favourable impact on the aforementioned indicators. Two main groups of contributory factors should be highlighted:

- Factors related to the business fabric: especially the creation of companies in the Services Sector, SMEs (particularly those with less than 10 employees) and, of course, those with a high turnover (more than €50 M). SMEs appear as powerful development drivers which occurs at the same time that fibre coverage is being deployed. The diversification of the activity sectors also appears as a relevant factor.
- Income and expenditure factors: mainly personal income and household spending on communications.

> **Fibre deployment is an effective direct catalyst for the development of rural Spain.**

Based on these results, which have been obtained essentially from the panel data models, and added to the conclusions on the effects analysed in section 3.3, it is concluded that **this study presents fibre deployment as a very effective direct catalyst to encourage the fundamental indicators of development of rural municipalities in Spain (Unemployment Rate, Service Companies and Household Income).**

4.2 Recommendations

Based on the conclusions and results presented, it can be concluded that, in order to maximise the beneficial impact of fibre deployment and, in general, of broadband connectivity, it is advisable to combine it with plans that, through the digitalisation of rural regions, encourage the economic and social fabric which can benefit from this deployment. For example:

- Incentives for the development of the rural business fabric, especially for SMEs.
- Training plans for specific sectors of the population (young people, women, the elderly), especially in the use of ICTs, so that the connectivity made available can be fully utilised.
- Training plans for the reconversion of employment: connectivity provides access to training that would not otherwise be available and can also help to increase the chances of the success of businesses and services which are supported by connectivity (training in entrepreneurship, marketing of local production, access to third-party services, etc.).
- The attraction of high turnover companies can have a major impact in rural municipalities, but obviously this is uncommon. Having said this, good connectivity is a fundamental prerequisite when these types of companies are looking for a new location. Once again, when it comes to promoting this type of initiative, there must be a set of complementary policies to attract investment in which infrastructure is a key factor.

In addition, as we have seen in section 3.5.4, the benefit of fibre deployment starts from the first property units covered, with the favourable impact being greater in the early stages of coverage than in the stages that are able to reach the entire population of a municipality. However, bringing connectivity to a rural municipality for the first time is precisely the most expensive phase of coverage expansion, therefore generating incentives for these initial deployments is beneficial. In this way, more significant progress is made in breaking down the digital divide that affects rural municipalities.

> **It is necessary to combine the effort made in fibre deployment with other impact factors:**
>
> - **Training plans**
> - **Promoting business development**
> - **Investment attraction**
> - **New business models**
> - **Digitalisation**

These recommendations, which naturally arise from the results of this study, are fully aligned with the objectives and strategic lines of the Digitalisation Strategy for the Agri-Food and Forestry Sector and that of Rural Environments [56]. This document was approved in March 2019 within the Ministry of Agriculture, Fisheries and Food with the aim of encouraging the promotion of these sectors through digitalisation.

In Table 10 a summary of the objectives and strategic lines of the said document is presented. Specifically, there is alignment between them and the recommendations of this study in terms of **bridging the digital divide through connectivity, capacity building, the digital adoption of SMEs and promoting business development and new business models.**

Objective	Strategic lines
1. Bridging the digital divide	1. Connectivity 2. Training
2. Promote data usage	1. Interoperability 2. Open data 3. Value chain and environmental data
3. Promote business development and new business models	1. Strengthening the digital innovation ecosystem 2. Providing advice for digital adoption in Agri-food, Forestry and Rural Knowledge and Innovation Systems 3. Promoting new business models

Table 10: Objectives and Strategic Lines of the Digitalisation Strategy for the Agri-food and Forestry Sector and that of the Rural Environment. Approved in March 2019 by the Ministry of Agriculture, Fisheries and Food.

The results presented in sections 3.2.4 and 3.5.4 suggest the relevance of carrying out a more thorough analysis of the relationship between population movements between territories and migrations as well as, in general, a **specific analysis of the depopulation problem.**

In addition, section 1.3.2 also suggests the need for an analysis that takes into account both the availability of connectivity as well as its adoption by potential users (individuals, companies, public administrations, etc.).

Next steps:

- **Use of other variables.**
- **Further the analysis of the depopulation problem.**
- **Consider the impact of adoption.**

Finally, mainly due to three reasons, this study has not been able to take into account factors that could help to complete the findings and the interpretation of the conclusions. Therefore, for future analyses, it would be interesting to consider other sets of variables such as the following suggestions:

- Variables relating to different types of social assistance, whether of a public nature or that provided by the Third Sector.
- Variables related to customer services which require connectivity for their provision: card payments, other financial services, online shopping, etc.
- Variables related to services requiring connectivity and provided between businesses and between these businesses and customers (B2B, B2G, G2G, G2C)[22].
- Variables related to public investment: the section on data sources discusses the consideration of municipal budget variables in this study and the difficulties for them to represent the effort made through public policies in a relevant and homogeneous way throughout the country. A more in-depth extraction and analysis of this information (bringing together local, regional and national policies) would be advisable in order to gain a more accurate picture of its impact. Obviously, this would by no means be a trivial analysis.

[22] Acronyms referring to services involving Businesses (B), Governments (G) and Customers (C). B2B for services between companies (Business to Business), B2G for services from companies to governments (Business to Government), G2G for services between governments (Government to Government, for example, between City Councils and Autonomous Communities or the Central Government), G2C for services from Governments to Customers (e.g., online procedures that citizens can carry out in City Councils and Public Administrations).

Annexes

Annex I

5.1 Variables and Data Sources

Sociodemographic Data

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Population	Number of inhabitants registered in the municipal register of inhabitants	INE	OK
Variation in population	Evolution of the population during the years going from 2014 to 2018, calculated as the slope of a simple regression	INE	OK
Population density	Number of inhabitants per square kilometre of the municipality	INE	OK
Percentage of foreigners	Ratio between the number of foreigners registered in the municipal register and the total population Expressed as a percentage.	INE	OK
Percentage of women	Ratio between the number of women registered in the municipal register and the total population. Expressed as a percentage.	INE	OK
Percentage of men	Ratio between the number of men registered in the municipal register and the total population. Expressed as a percentage.	INE	OK
Percentage of young people under the age of 16	Ratio between the number of inhabitants aged between 0 and 16 registered in the municipal register and the total population. Expressed as a percentage.	INE	OK
Percentage of people aged between 16 and 64	Ratio between the number of inhabitants aged between 16 and 64 registered in the municipality register and the total population. Expressed as a percentage.	INE	OK
Percentage of population over the age of 65	Ratio between the number of inhabitants over the age of 65 registered in the municipal register and the total population. Expressed as a percentage.	INE	OK
Dependency rate	Ratio between the number of dependent inhabitants (number of inhabitants under the age of 16 + the number of inhabitants over the age of 65) and the number of inhabitants between the ages of 16 and 64.	INE	OK

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Average age	Average age of the population of the municipality calculated based on the municipal register by age	INE	OK
Birth rate	Ratio between the number of births and the total population. Expressed in per thousand	The Autonomous Regions' Statistical Institutes	Only available until 2017
Mortality rate	Ratio between the number of deaths and the total population. Expressed in per thousand.	The Autonomous Regions' Statistical Institutes	Only available until 2017
Percentage of low levels of education	Percentage of population with low levels	AIS-Group[23]	Available for municipalities with less than 30,000 inhabitants and for all the years involved
Percentage of intermediate levels of education	Percentage of population with intermediate levels of education	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved
Percentage of high levels of education	Percentage of population with high levels of education.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Working members of households	Average number of working people per household	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved
Employed members of households	Average number of employed people per household	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved

[23] *When the AIS-Group is mentioned as the source, it refers to this company's compilation of data coming from various sources: The Census, EPF, ECV, the Working Population Survey (EPA in its Spanish initials), the State Public Employment Service (SEPE in its Spanish initials), INE and Iberinform.*

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Percentage of students per household	Percentage of the members of the household who are students	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Economic capacity	Economic capacity rate anchored to the National figure	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Poverty rate	Poverty rate	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Percentage of children in poverty	Percentage of children at risk of poverty	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Percentage of adults in poverty	Percentage of adults (not including the elderly) at risk of poverty	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Percentage of the elderly in poverty	Percentage of the elderly at risk of poverty	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Probability of being in default for debts	Probability of being in default for debts	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.

Geographical Data

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Altitude	Altitude measured in metres above sea level of the central point of the municipality	Telefónica	OK
Distance 2 k	Distance measured in kilometres as the crow flies to the nearest municipality with >2,000 inhabitants	Open Street Maps + INE	OK
Distance 100 k	Distance measured in kilometres as the crow flies to the nearest municipality with >100,000 inhabitants	Open Street Maps + INE	OK
Distance from motorway	Distance measured in kilometres by road to the nearest motorway access	Open Street Maps	OK
Coast length	Length in kilometres of the municipality's coastline	Open Street Maps	OK
Coast indicator	Flag which indicates whether the municipality has a coastline (1) or not (0), calculated based on the length of the coastline	Open Street Maps	OK
Surface area	Area of the municipality measured in square kilometres	National Geographic Institute	OK

Employment Data

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Unemployment Rate	Ratio between the number of unemployed people registered in the municipality and the working age population (Spanish and foreign population between the ages of 16 and 65). Expressed as a percentage	State Public Employment Service (SEPE)	OK
Unemployment Rate <25	Ratio between the number of unemployed people under the age of 25 and the population between the ages of 16 and 25 in the municipality. Expressed as a percentage	SEPE	OK

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Unemployment Rate 25-45	Ratio between the number of unemployed between the ages of 25 and 45 and the population between the ages of 25 and 45 in the municipality. Expressed as a percentage	SEPE	OK
Unemployment Rate >=45	Ratio between the number of unemployed over the age of 45 and the population between the ages of 45 and 65 in the municipality Expressed as a percentage	SEPE	OK
Unemployment Rate men <25	Ratio between the number of unemployed men under the age of 25 and the population of men between the ages of 16 and 25 in the municipality. Expressed as a percentage	SEPE	OK
Unemployment Rate men 25-45	Ratio between the number of unemployed men between the ages of 25 and 45 and the population of men between the ages of 25 and 45 in the municipality Expressed as a percentage	SEPE	OK
Unemployment Rate men >=45	Ratio between the number of unemployed men over over the age of 45 and the population of men between the ages of 45 and 65 in the municipality. Expressed as a percentage	SEPE	OK
Unemployment Rate women <25	Ratio between the number of unemployed women under the age of 25 and the population of women between the ages of 16 and 25 in the municipality. Expressed as a percentage	SEPE	OK
Unemployment Rate women 25-45	Ratio between the number of unemployed women between the ages of 25 and 45 and the population of women between the ages of 25 and 45 in the municipality. Expressed as a percentage.	SEPE	OK
Unemployment Rate women 25-45	Ratio between the number of unemployed women over the age of 45 and the population of women between the ages of 45 and 65 in the municipality. Expressed as a percentage	SEPE	OK
Percentage of Unemployment in Construction	Ratio between the number of unemployed in the construction sector and the total number of unemployed people in the municipality. Expressed as a percentage	SEPE	OK

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Percentage of Unemployment in Industry	Ratio between the number of unemployed in the industry sector and the total number of unemployed people in the municipality. Expressed as a percentage	SEPE	OK
Percentage of Unemployment in Agriculture	Ratio between the number of unemployed in the agriculture sector and the total number of unemployed people in the municipality. Expressed as a percentage	SEPE	OK
Percentage of Unemployment in Services	Ratio between the number of unemployed in the services sector and the total number of unemployed people in the municipality. Expressed as a percentage	SEPE	OK
Percentage of Unemployed without previous employment	Ratio between the number of unemployed people without previous employment and the total number of unemployed people in the municipality. Expressed as a percentage	SEPE	OK
Construction Unemployment Rate	Ratio between the number of unemployed people in the construction sector and the population between the ages of 16 and 65. Expressed as a percentage.	SEPE	OK
Industry Unemployment Rate	Ratio between the number of unemployed people in the industry sector and the population between the ages of 16 and 65. Expressed as a percentage	SEPE	OK
Unemployment Rate in Agriculture	Ratio between the number of unemployed people in the agriculture sector and the population between the ages of 16 and 65. Expressed as a percentage	SEPE	OK
Unemployment Rate in Services	Ratio between the number of unemployed people in the services sector and the population between the ages of 16 and 65. Expressed as a percentage	SEPE	OK
Unemployment Rate without previous employment	Ratio between the number of unemployed people without previous employment and the population between the ages of 16 and 65. Expressed as a percentage	SEPE	OK

Data on Companies and Activity Sectors

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Percentage of the Technical sector	Ratio between the number of companies categorised as: "Professional, scientific and technical activities and Administrative activities and auxiliary services" registered in the municipality in comparison with the total number of registered companies. Expressed as a percentage.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved
Percentage of the Industry sector	Ratio between the number of companies categorised as: "Extractive and manufacturing industries; Supply of electricity, gas, steam and air conditioning; Water supply, sewerage, waste management and remediation" registered in the municipality in comparison with the total number of registered companies. Expressed as a percentage	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved
Percentage of the Retail, Transport and Accommodation sectors	Ratio between the number of companies categorised as: "Wholesale and retail trade; repair of motor vehicles and motorcycles; Transport and storage; Accommodation" registered in the municipality in comparison with the total number of registered companies. Expressed as a percentage	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved
Percentage of the Education, Health and Social Services sector	Ratio between the number of companies categorised as: "Education; Health and social services activities" registered in the municipality in comparison with the total number of registered companies. Expressed as a percentage	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved
Percentage of the Construction sector	Ratio between the number of companies categorised under "Construction" and registered in the municipality in comparison with the total number of registered companies. Expressed as a percentage	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved. 3
Percentage of the Information and Communications sector	Ratio between the number of companies categorised under "Information and communications" and registered in the municipality in comparison with the total number of registered companies. Expressed as a percentage	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved
Percentage of the Finance and Insurance sector	Ratio between the number of companies categorised under "Financial and insurance activities" and registered in the municipality in comparison with the total number of registered companies. Expressed as a percentage	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Percentage of the Property sector	Ratio between the number of companies categorised under "Property activities" and registered in the municipality in comparison with the total number of registered companies. Expressed as a percentage.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved.
Percentage of the Other Services sector	Ratio between the number of companies categorised as: "Arts, recreational and entertainment activities and Other services" registered in the municipality in comparison with the total number of registered companies. Expressed as a percentage.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved.
Percentage of all of the Services sector	Ratio between the number of companies resulting from adding together all of the Technical, Education, Health and Social Services, Information and Communications, Finance and Insurance, Property, and Other Services categories registered in the municipality in comparison with the total number of registered companies. Expressed as a percentage.	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved.
Technical Companies per thousand inhabitants	Ratio between the number of Technical companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved.
Industry Companies per thousand inhabitants	Ratio between the number of Industry companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved.
Retail, Transport and Accommodation Companies per thousand inhabitants	Ratio between the number of Retail, Transport and Accommodation companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved.
Education, Health and Social Services Companies per thousand inhabitants	Ratio between the number of Education, Health and Social Services companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved.

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Construction companies per thousand inhabitants	Ratio between the number of Construction companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved.
Information and Communications Companies per thousand inhabitants	Ratio between the number of Information and Communication companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved.
Finance and Insurance Companies per thousand inhabitants	Ratio between the number of Finance and Insurance companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved.
Property Companies per thousand inhabitants	Ratio between the number of Property companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved.
Other Services Companies per thousand inhabitants	Ratio between the number of Other Services companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved.
Services Companies per thousand inhabitants	Ratio between the number of companies resulting from adding together the Technical, Education, Health and Social Services, Information and Communications, Finance and Insurance, Property, and Other Services categories registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved.
Companies per thousand inhabitants	Ratio between the total number of companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	INE	Available for all the municipalities and for all the years involved.
Autonomous Community - Difference in the Technical sector	Difference between the percentage of companies in the Technical sector in the municipality and the average percentage of companies in that sector in the municipality's Autonomous Community.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved.

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Autonomous Community - Difference in the Industry sector	Difference between the percentage of companies in the Industry sector in the municipality and the average percentage of companies in that sector in the municipality's Autonomous Community.	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved.
Autonomous Community - Difference in the Retail, Transport and Accommodation sectors	Difference between the percentage of companies in the Retail, Transport and Accommodation sectors in the municipality and the average percentage of companies in these sectors in the municipality's Autonomous Community.	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved.
Autonomous Community - Difference in the Education, Health and Social Services sectors	Difference between the percentage of companies in the Education, Health and Social Services sectors in the municipality and the average percentage of companies in these sectors in the municipality's Autonomous Community.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved.
Autonomous Community - Difference in the Construction sector	Difference between the percentage of companies in the Construction sector in the municipality and the average percentage of companies in that sector in the municipality's Autonomous Community.	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved
Autonomous Community - Difference in the Information and Communications sector	Difference between the percentage of companies in the Information and Communications sector in the municipality and the average percentage of companies in that sector in the municipality's Autonomous Community.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved
Autonomous Community - Difference in the Finance and Insurance sector	Difference between the percentage of companies in the Finance and Insurance sector in the municipality and the average percentage of companies in that sector in the municipality's Autonomous Community.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved
Autonomous Community - Difference in the Property sector	Difference between the percentage of companies in the Property sector in the municipality and the average percentage of companies in that sector in the municipality's Autonomous Community.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved
Autonomous Community - Difference in the Other Services sector	Difference between the percentage of companies in the Other Services sector in the municipality and the average percentage of companies in that sector in the municipality's Autonomous Community.	INE	Available for municipalities with > 5,000 inhabitants and for all the years involved

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Autonomous Community - Difference in the Services sector	Difference between the percentage of all of the companies across the whole Services sector in the municipality and the average percentage of companies in that sector in the municipality's Autonomous Community.	INE	Available for municipalities with > 1,000 inhabitants and for all the years involved
Employment Resilience Index	This indicates how resilient a municipality is to a crisis in a given sector, depending on the distribution of job vacancies between the sectors. THE CLOSER TO 0, THE BETTER	INE	Available for municipalities with between 1,000 and 100,000 inhabitants and for all the years involved
Employment Resilience Index - Complete	Similar to the previous one, but it also includes the sectors that only have data for municipalities with >5,000 inhabitants. THE CLOSER TO 0, THE BETTER.	INE	Available for municipalities with between 5,000 and 100,000 inhabitants and for all the years involved.
Companies with 50-250 Employees per thousand inhabitants	Ratio between the number of companies with 50-250 employees registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Companies with 10-50 Employees per thousand inhabitants	Ratio between the number of companies with 10-50 employees registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Companies with > 50 M Sales per thousand inhabitants	Ratio between the number of companies with a turnover of over 50 million euros registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved
Companies with 10 M - 50 M Sales per thousand inhabitants	Ratio between the number of companies with a turnover between 10 and 50 million euros registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved
Companies with 2 M - 10 M Sales per thousand inhabitants	Ratio between the number of companies with a turnover between 2 and 10 million euros registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Companies < 2 M Sales per thousand inhabitants	Ratio between the number of companies with a turnover of less than 2 million euros registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved
AIS - Agriculture, Livestock and Fishing Companies per thousand inhabitants	Ratio between the number of Agriculture, Livestock and Fishing companies registered in the municipality in comparison with the number of inhabitants of the the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved
AIS - Extractive Industry Companies per thousand inhabitants	Ratio between the number of Extractive Industry companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Manufacturing Industry Companies per thousand inhabitants	Ratio between the number of Manufacturing Industry companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Energy Companies per thousand inhabitants	Ratio between the number of Energy companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Water Companies per thousand inhabitants	Ratio between the number of Water companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Construction Companies per thousand inhabitants	Ratio between the number of Construction companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Retail Companies per thousand inhabitants	Ratio between the number of Retail companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Transport Companies per thousand inhabitants	Ratio between the number of Transport companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
AIS - Accommodation Companies per thousand inhabitants	Ratio between the number of Accommodation companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Information and Communications Companies per thousand inhabitants	Ratio between the number of Information and Communication companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved
AIS - Financial and Insurance Companies per thousand inhabitants	Ratio between the number of Finance and Insurance companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved
AIS - Property Companies per thousand inhabitants	Ratio between the number of Property companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Technical companies per thousand inhabitants	Ratio between the number of Scientific and Technical companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Administrative Companies per thousand inhabitants	Ratio between the number of Administrative companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Public Administration Companies per thousand inhabitants	Ratio between the number of Public Administration companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved
AIS - Education Companies per thousand inhabitants	Ratio between the number of Education companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Health and Social Services Companies per thousand inhabitants	Ratio between the number of Health and Social Services companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
AIS - Artistic Companies per thousand inhabitants	Ratio between the number of Artistic companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Other Services Companies per thousand inhabitants	Ratio between the number of Other Services companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Domestic Assistance Companies per thousand inhabitants	Ratio between the number of Domestic Assistance Employees registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
AIS - Extraterritorial Organisation Companies per thousand inhabitants	Ratio between the number of Extraterritorial Organisation companies registered in the municipality in comparison with the number of inhabitants of the municipality. Expressed in per thousand.	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved
Sales per company	Average sales volume among reporting companies	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Employees per company	Average number of employees among reporting companies	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved

Data on Social Security Affiliation

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Social Security General Scheme Percentage	Ratio between the number of affiliates to the Social Security General Scheme and the total number of affiliates. Expressed as a percentage.	Social Security	OK
Social Security Self-Employed Scheme Percentage	Ratio between the number of affiliates to the Social Security Self-Employed Scheme and the total number of affiliates. Expressed as a percentage.	Social Security	OK
Social Security Special Scheme	Ratio between the number of affiliates to the Social Security Special Schemes (Agricultural Workers Special Scheme Domestic Employees Special Scheme, Special Scheme for Seafarers, Special Scheme for Coal Workers) and the total number of affiliates. Expressed as a percentage.	Social Security	OK
Affiliates to the Social Security General Scheme per working population	Ratio between the number of affiliates to the Social Security General Scheme and the population at a working age. Expressed as a percentage.	Social Security	OK
Self-Employed Affiliates per working population	Ratio between the number of people affiliated to the Social Security Self-Employed Scheme and the population at a working age. Expressed as a percentage.	Social Security	OK
Affiliates to Social Security Special Schemes per the working population	Ratio between the number of people affiliated to the Social Security Special Scheme and the population at a working age. Expressed as a percentage.	Social Security	OK
Workers per working population	Ratio between the total number of Social Security affiliates and the population at a working age. Expressed as a percentage.	Social Security	OK

Data on Income

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Average Gross Income	Obtained from the annual income tax return	Tax Agency	Available for municipalities with > 1,000 inhabitants except for the Basque Country and Navarra, and for the years 2014-2017
Average Disposable Income	Obtained from the annual income tax return	Tax Agency	Available for municipalities with > 1,000 inhabitants except for the Basque Country and Navarra, and for the years 2014-2017
Average Income per Person	Obtained from the Household Income Distribution Atlas	INE	Available for municipalities with > 100 inhabitants and for the years 2015-2016, except for municipalities in Gipuzkoa, which data is only available for 2016.
Average Household Income	Obtained from the Household Income Distribution Atlas.	INE	Available for municipalities with > 100 inhabitants and for the years 2015-2016, except for municipalities in Gipuzkoa, which data is only available for 2016.
Income per household	Average net monthly household income	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Income per working household	Average net monthly income for working households	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Income per employed person in working households	Average net monthly income for working households per employed person	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Income per member in working households	Average net monthly income of working households per family member	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Income per retired households	Average net monthly income of retired households	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Personal Income	Personal net monthly income	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Annual spending on communications per household	Average annual household spending on communications	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.
Annual educational expenditure per household	Average annual household spending on education	AIS-Group	Available for municipalities with less than 30,000 inhabitants and for all the years involved.

Data on Variations in the Population

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Registrations from the same Autonomous Community, but a different province per thousand inhabitants	Ratio between the number of registrations from the same Autonomous Community but from a different Province and the municipality's population. Expressed in per thousand.	INE	OK
Registrations from the same province per thousand inhabitants	Ratio between the number of registrations from the same province and the municipality's population. Expressed in per thousand.	INE	OK
Registrations from another Autonomous Community per thousand inhabitants	Ratio between the number of registrations from a different Autonomous Community and the municipality's population. Expressed in per thousand	INE	OK

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Registrations of foreigners per thousand inhabitants	Ratio between the number of registrations of foreigners and the municipality's population. Expressed in per thousand.	INE	OK
Total registrations per thousand inhabitants	Ratio between the total the number of registrations and the municipality's population. Expressed in per thousand.	INE	OK
Deregistrations moving to the same Autonomous Community, but a different province per thousand inhabitants	Ratio between the number of deregistrations from the same Autonomous Community but from a different Province and the municipality's population. Expressed in per thousand	INE	OK
Deregistrations moving to the same province per thousand inhabitants	Ratio between the number of deregistrations moving to the same province and the municipality's population. Expressed in per thousand.	INE	OK
Deregistrations moving to another Autonomous Community per thousand inhabitants	Ratio between the number of deregistrations moving to a different Autonomous Community and the municipality's population. Expressed in per thousand.	INE	OK
Deregistrations moving abroad per thousand inhabitants	Ratio between the number of deregistrations moving abroad and the municipality's population. Expressed in per thousand.	INE	OK
Total number of deregistrations per thousand inhabitants	Ratio between the total the number of deregistrations and the municipality's population. Expressed in per thousand.	INE	OK
Population balance per thousand inhabitants	Ratio between total populational balance (registrations and deregistrations) and the municipality's population. Expressed in per thousand.	INE	OK

Telefónica

Telco Data

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
White Zone Percentage	Percentage of population residing in white zones, calculated based on the 2014 population as the total. Expressed as a percentage.	Secretary of State for Digital Advancement	Available for all the municipalities and from 2016-2018.
White Zone Indicator	Indicator that designates whether the municipality has > 0% of the population living in white zones (1) or no inhabitants in white zones (0).	Secretary of State for Digital Advancement	Available for all the municipalities and from 2016-2018.
FTTH from Telefónica Indicator 2019	Indicator that designates whether the municipality had FTTH coverage from Telefónica in 2019 (1) or not (0).	Telefónica	Available for all municipalities
FTTH Coverage Percentage 2019	Percentage of the municipality with FTTH coverage from Telefónica in 2019. It is unknown whether this data has been calculated based on the population, property units, or another indicator.	Telefónica	Available for all municipalities
Fixed Network Coverage over 2 Mbps Percentage	Percentage of households that had fixed network coverage with a speed of > 2 Mbps in 2017 and 2018	Secretary of State for Digital Advancement	Available for all municipalities and from 2017-2018.
Fixed Network Coverage over 10 Mbps Percentage	Percentage of households that had fixed network coverage with a speed of > 10 Mbps in 2017 and 2018.	Secretary of State for Digital Advancement	Available for all municipalities and from 2017-2018.
Fixed Network Coverage over 30 Mbps Percentage	Percentage of households that had fixed network coverage with a speed of > 30 Mbps in 2017 and 2018.	Secretary of State for Digital Advancement	Available for all municipalities and from 2017-2018.
Fixed Network Coverage over 30 Mbps and with NGA quality Percentage	Percentage of households that had fixed network coverage with a speed of > 30 Mbps and with NGA (Next generation Access) Technology in 2017 and 2018	Secretary of State for Digital Advancement	Available for all municipalities and from 2017-2018
Fixed Network Coverage over 100 Mbps Percentage	Percentage of households that had fixed network coverage with a speed of > 100 Mbps in 2017 and 2018.	Secretary of State for Digital Advancement	Available for all municipalities and from 2017-2018.
3G HSPA Coverage Percentage	Percentage of households that had 3G mobile coverage in 2017 and 2018.	Secretary of State for Digital Advancement	Available for all municipalities and from 2017-2018

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
4G LTE Coverage Percentage	Percentage of households that had 4G mobile coverage in 2017 and 2018.	Secretary of State for Digital Advancement	Available for all municipalities and from 2017-2018.
FTTH minimum coverage Date	Indicates the month in which at least 30% of property units were provided with Telefónica FTTH coverage. Expressed in DD-MM-YYYY format.	Telefónica	Available for all municipalities <50,000 inhabitants. Missing values imply that 30% coverage has not been reached in the municipality.
"Maximum FTTH coverage" (percentage of property units covered by FTTH)	Indicates the maximum percentage of property units covered with Telefónica's FTTH and reached during the year. Expressed as a percentage	Telefónica	Available for all municipalities <50,000 inhabitants and all the years involved.

Other Data

Data	Description	Source	Availability for all municipalities and for the whole time range from 2014 to 2018
Patents applied for per thousand inhabitants	Ratio between the number of patents applied for per municipality of the first applicant and the population of the said municipality. Expressed in per thousand.	Spanish Patent and Trademark Office	OK
Percentage of valid votes	Ratio of votes counted compared with the total number of people registered on the electoral roll. Expressed as a percentage.	INE, El País	Available for all municipalities, in 2016 and 2018 (2019 elections).
Percentage of abstentions	Inverse percentage of the Percentage of valid votes. Expressed as a percentage.	INE, El País	Available for all municipalities, in 2016 and 2018 (2019 elections)
Percentage of invalid votes	Ratio between the votes considered invalid and the counted votes. Expressed as a percentage.	INE, El País	Available for all municipalities, in 2016 and 2018 (2019 elections)
Percentage of blank votes	Ratio of votes considered blank the counted votes. Expressed as a percentage.	INE, El País	Available for all municipalities, in 2016 and 2018 (2019 elections)
School Indicator 2019	Indicator that designates whether a municipality had a school in 2019 (1) or not (0).	Scholarum Website	Available for all municipalities. 11
Hospital Indicator 2019	Indicator that designates whether a municipality had a hospital in 2018 (1) or not (0).	National Hospital Catalogue 2017	Available for all municipalities.
Train Station Indicator 2019	Indicator that designates whether a municipality has a train station in 2019 (1) or not (0).	ADIF	Available for all municipalities.

5.2 Glossary

Machine Learning
Techniques that develop algorithms and statistical models which enable an IT system to perform tasks based on patterns and inferences without using specific instructions.

Broadband
There is no single agreed definition of the access speed that defines broadband. This study considers it as a minimum of 30 Mbps using FTTH technology.

CNAE
Spanish acronym for the National Classification of Economic Activities. This classification assigns a numerical code to each economic activity and defines an aggregation hierarchy at various levels.

Consistency
In statistics, this refers to the fact that the estimated value converges towards the true value when the number of samples converge towards infinity in a model.

Panel Data
A statistical technique widely used in econometrics which works with data that has a temporal dimension (years, months, etc.) as well as a cross-sectional dimension (individuals in the sample).

Box-plot diagram
These are representations of data distributions based on 5 metrics represented in the figure below:

- First quartile (Q1, or 25 percent): value that includes 25% of the values in the dataset below it.
- Third quartile (Q3, or 75 percent): value that includes 75% of the values in the dataset below it.
- Average: the central value of the dataset (coincides with Q2, or 50 percent).
- Minimum: Q1 - 1.5*IQR.
- Maximum: Q3 + 1.5*IQR.
- Values of less than the minimum or greater than the maximum are considered to be values that are exceptionally far from the majority of the values in the dataset (“outliers”).

Where IQR = Q3 - Q1

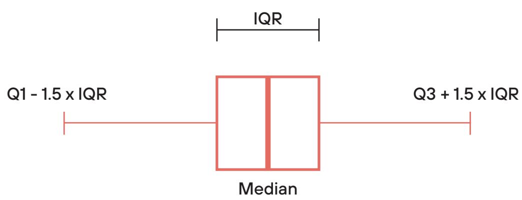


Image source: user Jhguch at en.wikipedia

ECV
Encuesta de Condiciones de Vida (Living Conditions Survey) conducted annually by the INE.

Fixed effects
In the panel data technique and for the regression being estimated, fixed effects are considered to be the fact of assuming that each individual in the sample may have a different starting point (ordinate at the origin).

Individual effects
In the panel data technique, individual effects are those that affect each of the individuals in the sample differently, with the said effects being invariant over time.

Time effects
In the panel data technique, time effects are those that are dependent on time and which affect all the individuals in the sample.

Efficiency
In statistics, this refers to the fact that an estimator may have more variance (which would be less accurate) than another.

Endogeneity
In a model that seeks to explain Y with X_k variables, endogeneity occurs when either of the X and Y variables are bidirectionally related. For example, in a supply-demand model, if one tries to predict price (Y) from demand (X), endogeneity will arise, since producers adjust their price based on the demand (X causes changes in Y) but consumers also change their demand in response to price (Y causes changes in X).

EPA
Encuesta de Población Activa (Working Population Survey) carried out by the INE on a quarterly basis.

EPF
Encuesta de Presupuestos Familiares (Household Budget Survey) carried out by the INE on an annual basis

Exogeneity
When there is no two-way relationship between one (or more) independent variable(s) and the dependent variable. In short, an absence of endogeneity.

Forward selection
This is a variable selection technique to reduce the initial set of variables to a smaller set while controlling the possible loss of representativity of the model.

FTTx
The scope of fibre-optic based connectivity can be to the subscriber's own home (FTTH for Fibre-to-the-home) or to the building connections (FTTB for Fibre-to-the-building or FTTP Fibre-to-the-premises for normal or business subscribers' buildings respectively).

Degrees of freedom
The degrees of freedom express the amount of information from the data that can be used to calculate the variability of the estimates. The greater the degrees of freedom, the more reliable the estimates obtained will be.

Unobservable heterogeneity
In the panel data technique, unobservable heterogeneity refers to the pattern of variables which, due to their nature, are unobservable and therefore cannot be explicitly included in the model (for example, attitudinal aspects of the population).

INE
Instituto Nacional de Estadística (National Statistics Institute).

Resilience Index
The Resilience Index is an indicator created by researchers at the Innovation Centre in Technology for Development, belonging to the Polytechnic University of Madrid. This indicator summarises the composition by employment sectors into a single value in the said municipality compared to the average of the other Spanish municipalities. This gives a measure of a municipality's vulnerability to a crisis in a particular sector resulting from the low diversification of employment.

Multicollinearity
Multicollinearity occurs when, in a model whose objective is to explain the pattern of a Y variable from another set of explanatory X_k variables, there is correlation between several or all the explanatory variables.

Rural Municipalities
According to Law 45/2007 on the Sustainable Development of the Rural Environment, Rural Municipalities are those with less than 30,000 inhabitants and with a population density of less than 100 inhabitants per km^2.

Significance level
In statistics, an effect is statistically significant if it is unlikely to be due to chance. This is associated with a hypothesis check that provides a p-value which determines that the lower the p., the greater significance of the result. Thus, a p-value which is less than 0.01 indicates a significance level of 1% (99% confidence level) and in this study is indicated with ***. If the p-value is between 0.01 and 0.05 it indicates a significance level of between 1% and 5% (confidence level between 99% and 95%) and is indicated with **. If the p-value is between 0.05 and 0.1 it indicates a significance level of between 5% and 10% (a confidence level of between 95% and 90% respectively) and is indicated by *.

SMEs
Small and Medium Enterprises. According to Annex I of the Commission Regulation (EU) Nº 651/2014, these are companies with fewer than 250 employees, and with a turnover of no more than €50 M or an annual balance sheet total of no more than €43 M.

R^2
A coefficient that characterises the quality of a statistical model in determining the proportion of variation in the results which can be explained by the said model.

Bayesian Networks
A technique that allows the creation of models to represent different variables related to each other through their conditional probabilities. This can be represented by a Directed Acyclic Graph (DAG).

Regression
In statistics, a technique for estimating relationships between a dependent variable and one or more independent variables. A linear regression assumes that the dependent variable is a linear combination of the independent variables.

ICT
Information and Communication Technologies.

Two Stage Least Squares (TSLS)
A technique applicable to panel data which uses instrumental variables in two stages: the first estimates the variable to be substituted from the instrumental variable, and the second one uses these estimates to estimate the original dependent variable.

Dependent and independent variables
In a model whose objective is to explain the pattern of a Y variable from another set of explanatory variables X_k, Y is called the dependent variable and X_k variables are called independent ones.

Voice over IP
A set of resources and protocols that allow voice calls to be made over the Internet.

xDSL
A set of technologies that provide high-speed connection using traditional copper lines. DSL stands for Digital Subscriber Line and the 'x' stands for the possible variants depending on the technology used (A for 'asymmetrical', 'S' for 'symmetrical', H for 'high bit-rate', V for 'very high bit-rate', etc.).

White Zones
Areas that do not have coverage of new generation broadband networks, nor forecasts their provision by any operator within 3 years, based on credible investment plans.

Bibliography

[1] BAI, Yang. *The faster, the better? The impact of internet speed on employment.* Information Economics and Policy, 2017, vol. 40, p. 21-25.

[2] Spanish Government, Spanish Digital Agenda, website: https://www.lamoncloa.gob.es/documents/agendadigital150213.pdf, last time accessed: 24/10/2019

[3] LAPOINTE, Paul. *Does speed matter? The employment impact of increasing access to fiber Internet.* Washington Academy of Sciences. Journal of the Washington Academy of Sciences, 2015, vol. 101, no 1, p. 9.

[4] FORMAN, Chris; GOLDFARB, Avi; GREENSTEIN, Shane. The Internet and local wages: A puzzle. American Economic Review, 2012, vol. 102, no 1, p. 556-75.

[5] Whitacre, Brian, Roberto Gallardo, and Sharon Strover. *"Does rural broadband impact jobs and income?* Evidence from spatial and first-differenced regressions." The Annals of Regional Science 53.3 (2014): 649-670.

[6] CRANDALL, Robert W.; LEHR, William; LITAN, Robert E. *The effects of broadband deployment on output and employment: A cross-sectional analysis of US data.* 2007.

[7] SINGER, H. J.; WEST, J. D. *Economic effects of broadband infrastructure deployment and tax incentives for broadband deployment.* White paper, Fiber-To-The-Home Council, 2010.

[8] COLOMBO, Massimo G.; CROCE, Annalisa; GRILLI, Luca. *ICT services and small businesses' productivity gains: An analysis of the adoption of broadband Internet technology.* Information Economics and Policy, 2013, vol. 25, no 3, p. 171-189.

[9] SINGER, Hal; CAVES, Kevin; KOYFMAN, Anna. The *Empirical Link Between Fibre-to-the-Premises Deployment and Employment: A Case Study in Canada.* Economists Incorporated, 2015.

[10] Shideler, Dave, and Narine Badasyan. *"The economic impact of broadband deployment in Kentucky."* TPRC, 2007

[11] KOLKO, Jed. *Broadband and local growth.* Journal of Urban Economics, 2012, vol. 71, no 1, p. 100-113.

[12] IVUS, Olena; BOLAND, Matthew. *The employment and wage impact of broadband deployment in Canada.* Canadian Journal of Economics / Revue canadienne d'économique, 2015, vol. 48, no 5, p. 1803-1830.

[13] HASBI, Maude. *Impact of Very High-Speed Broadband on Local Economic Growth: Empirical Evidence*. 2017.

[14] FISHER, Simon. *Next generation access-the latest strategy on FTTX.* In 2008 3rd IET Seminar on Next Generation Networks. IET, 2008. p. 1-11.

[15] CZERNICH, Nina. *Does broadband internet reduce the unemployment rate?* Evidence for Germany. Information Economics and Policy, 2014, vol. 29, p. 32-45.

[16] KRETSCHMER, Tobias. *Information and communication technologies and productivity growth*. 2012

[17] GRUBER, Hätönen; HÄTÖNEN, Jussi; KOUTROUMPIS, Pantelis. *Broadband access in the EU: An assessment of future economic benefits*. Telecommunications Policy, 2014, vol. 38, no 11, p. 1046-1058

[18] TOPORKOFF, Sylviane. ICT convergence and Europe's digital agenda 2010–2020. The Real Issues of the Middle East and the Arab Spring. Springer, New York, NY, 2013. p. 315-330.

[19] GILLETT, Sharon E., et al. *Measuring the Economic Impact of Broadband Deployment*. Final report, National Technical Assistance, Training. Research, and Evaluation Project, 2006, no 99-07, p. 13829.

[20] JAYAKAR, Krishna; PARK, Eun-A. *Broadband availability and employment: An analysis of county-level data from the National Broadband Map*. Journal of Information Policy, 2013, vol. 3, p. 181-200.

[21] MACK, Elizabeth A.; REY, Sergio J. *An econometric approach for evaluating the linkages between broadband and knowledge intensive firms*. Telecommunications Policy, 2014, vol. 38, no 1, p. 105-118.

[22] MCCOY, Daire, et al. *The impact of local infrastructure on new business establishments*. 2016.

[23] SOSA, David. Early evidence suggests gigabit broadband drives GDP. Analysis Group, 2014.

[24] CRANDALL, Robert W.; JACKSON, Charles L. *The $500 billion opportunity: The potential economic benefit of widespread diffusion of broadband Internet access*. Washington, DC: Criterion Economics, LLC, 2001.

[25] LEHR, William Herndon, et al. *Measuring broadband's economic impact*. 2006.

[26] SURIÑACH, Jordi; ROMANÍ, Javier; TERMES, Montserrat. *¿Afecta la banda ancha al crecimiento económico? Evidencia sobre agentes y territorio.* Investigaciones Regionales-Journal of Regional Research, 2007, no 10, p. 207-235.

[27] ANTONELLI, Cristiano. *The digital divide: understanding the economics of new information and communication technology in the global economy. Information Economics and Policy, 2003*, vol. 15, no 2, p. 173-199.

[28] MARTÍN, Juan Diego Otero. *Competencia y regulación en los mercados de comunicaciones electrónicas: banda ancha y despliegues de nueva generación*. Boletín económico de ICE, Información Comercial Española, 2019, no 3111, p. 19-30.

[29] ROJAS, Edwin Fernando, et al. *Conectados a la banda ancha: Tecnología, políticas e impacto en América Latina y España*. 2012.

[30] KATZ, Raúl L. *Impacto de la regulación de las telecomunicaciones en España*.

[31] KOUTROUMPIS, Pantelis. *The economic impact of broadband on growth: A simultaneous approach. Telecommunications Policy, 2009*, vol. 33, no 9, p. 471-485.

[32] KATZ, Raul L., et al. *Estimating broadband demand and its economic impact in Latin America*. Proc. 3rd ACORN-REDECOM, 2009, p. 1-20.

[33] CALZADA, Joan, et al. *Fiber deployment in Spain*. Journal of Regulatory Economics, 2018, vol. 53, no 3, p. 256-274.

[34] The State of Broadband 2018: *Broadband catalyzing sustainable development, 2018*, ISBN: 978-92-61-26431-4

[35] CAVA-FERRERUELA, Inmaculada; ALABAU-MUNOZ, Antonio. *Broadband policy assessment: A cross-national empirical analysis*. Telecommunications Policy, 2006, vol. 30, no 8-9, p. 445-463.

[36] KATZ, Raul L.; AVILA, Javier G. *The impact of broadband policy on the economy*. In Proceedings of the 4thd ACORN-REDECOM Conference Brasilia May. 2010. p. 1.

[37] KENNEDY, Jessica, et al. *A review of the economic impact of broadband on businesses in regional and rural communities. 2006*.

[38] BERTSCHEK, Irene, et al. *The Economic Impacts of Broadband Internet: A Survey*. Review of Network Economics, 2015, vol. 14, no 4, p. 201-227.

[39] KEISER, Gerd. *FTTX Concepts and Applications*. 2006.

[40] MAYORGA, Mauricio; MUÑOZ, Evelyn. *La técnica de datos de panel una guía para su uso e interpretación*. Central Bank of Costa Rica. Economic Research Department, 2000.

[41] LABRA, Romilio; TORRECILLAS, Cecilia. *Guía CERO para datos de panel. Un enfoque práctico*. UAM-Accenture Working Papers, 2014, vol. 16, no 1, p. 57.

[42] HSIAO, C. (2014). *Analysis of Panel Data (Nº 54).* Cambridge University Press.

[43] BABAALI, Nadia. *Are governments fulfilling their role in meeting the Digital Agenda?* FTTH Council Europe, http://www.ftthcouncil.eu/documents/Opinions/2012/Government_Final.pdf, 2012.

[44] RENDON SCHNEIR, Juan; XIONG, Yupeng. *Strategic and economic aspects of network sharing in FTTH/PON architectures*. 2012.

[45] MANSELL, R. (2014). *Here Comes the Revolution — The European Digital Agenda*. The Palgrave Handbook of European Media Policy (pp. 202-217). Palgrave Macmillan, London.

[46] ARELLANO, Manuel; BOVER, Olympia. *La econometría de datos de panel.* En Investigaciones Económicas (Segunda época) .Vol XIV, nº 1, (pp. 3-45). 1990.

[47] BOIX, Carles; RIBA, Clara. *Las bases sociales y políticas de la abstención en las elecciones generales españolas: recursos individuales, movilización estratégica e instituciones electorales.* Reis 90/00 (pp. 95-128)

[48] DERKSEN, Shelley; KESELMAN, H. J. *Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables.* British Journal of Mathematical and Statistical Psychology (1992) 45 (pp. 265-282)

[49] SENLIOL, Baris; GULGEZEN, Gokhan; et al. *Fast Correlation Based Filter (FCBF) with a Different Search Strategy*. 2008 23rd International Symposium on Computer and Information Sciences.

[50] MAO, K. Z. *Fast Orthogonal Forward Selection Algorithm for Feature Subset Selection*. IEEE Transactions on Neural Networks, 13(5). 2002 (pp. 1218-1224).

[51] Spanish Government. *Report: "Cobertura de Banda Ancha en España en el año 2018."* https://avancedigital.gob.es/banda-ancha/cobertura/Documents/Cobertura-BA-2018.pdf

[52] JORDÁN, V., Galperin, H., & Peres Núñez, W. (2010). *Acelerando la revolución digital: banda ancha para América Latina y el Caribe*.

[53] Ministry of Agriculture, Fisheries and Food, Spanish Government. *Annual Indicators Report. Agriculture, Fisheries and Food 2018*. https://www.mapa.gob.es/es/ministerio/servicios/analisis-y-prospectiva/informe_anual_agric_pesca_ali_mm_17finalweb221120182_tcm30-495934.pdf

[54] CABALLERO, J. M., & Artigas, J. M. C. (1997). *Redes de banda ancha*. Marcombo.

[55] FABRITZ, N. (2013). *The Impact of Broadband on Economic Activity in Rural Areas: Evidence from German Municipalities (Nº 166).* Ifo Working Paper.

[56] Ministry of Agriculture, Fisheries and Food, Government of Spain. *Estrategia de Digitalización del Sector Agroalimentario y Forestal y del Medio Rural.* 2019. https://www.mapa.gob.es/images/es/estrategia_digitalizacion_sector_agroalimentario_forestal_medio_rural_ve_tcm30-509645.pdf

[57] Mölleryd, B. G. (2015). *Development of High-speed Networks and the Role of Municipal Networks.*

[58] Kongaut, C., & Bohlin, E. (2014). *Impact of broadband speed on economic outputs: An empirical study of OECD countries*.

[59] Rohman, I. K., & Bohlin, E. (2013). *The impact of broadband speed on the household income: comparing OECD and Brics*.

[60] Rohman, I. K., & Bohlin, E. (2012). *Does broadband speed really matter for driving economic growth?* Investigating OECD countries.

[61] Gillett, S., Lehr, W., Osorio, C. A., & Sirbu, M. (2006). *Measuring the Economic Impact of Broadband Deployment. US Department of Commerce, Economic Development Administration*.

[62] 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.

[63] Whitacre, B. E. (2010). *Rural broadband availability and adoption in Oklahoma*. Choices, 25 (316-2016-7511), 1-7.

[64] Levin, S. L., Schmidt, S. R., & Scott, G. (2016). *Broadband adoption: Translating the digital divide literature into effective government policies and actions*. Available at SSRN 2911391.

[65] Dwivedi, Y. K., & Lal, B. (2007). *Socio-economic determinants of broadband adoption*. Industrial Management & Data Systems.

Telefónica

Experts who have participated in the preparation of the report

- Eduardo Merino Machuca.
- Silvia Hernández Martín.
- Esther Cuervo Fernandez.
- Javier Carro Calabor.
- Pedro Antonio de Alarcón Sánchez.
- Elena Quiroga Fernández.
- José María Bolufer Francia.