REM WORKING PAPER SERIES

Does road accessibility to cities support rural population growth? Evidence for Portugal for the 1991-2011 period

Patrícia C. Melo, Conceição Rego, Paulo Rui Anciães, Nuno Guiomar, José Muñoz-Rojas

REM Working Paper 0165-2021

March 2021

REM – Research in Economics and Mathematics
Rua Miguel Lúpi 20,
1249-078 Lisboa,
Portugal

ISSN 2184-108X

REM – Research in Economics and Mathematics

Rua Miguel Lupi, 20
1249-078 LISBOA
Portugal

Telephone: +351 - 213 925 912
E-mail: rem@iseg.ulisboa.pt

https://rem.rc.iseg.ulisboa.pt/

https://twitter.com/ResearchRem

https://www.linkedin.com/company/researchrem/

https://www.facebook.com/researchrem/

Does road accessibility to cities support rural population growth? Evidence for Portugal for the 1991-2011 period

Patrícia C. Melo[1], Conceição Rego[2], Paulo Rui Anciães[3], Nuno Guiomar [4], José Muñoz-Rojas[4]

[1] ISEG-Lisbon School of Economics and Management, University of Lisbon and REM/UECE, Lisbon, Portugal.
[2] Department of Economics and CEFAGE-U.É., Universidade de Évora, Portugal
[3] University College London, Centre for Transport Studies, London, United Kingdom
[4] LABSCAPE - Mediterranean Landscape Systems Lab, MED - Mediterranean Institute for Agriculture, Environment and Development, University of Évora, Portugal

Abstract

Transport investment is frequently advocated as having the double virtue of achieving both economic growth and territorial cohesion. The idea is that improving the accessibility of lagging regions to cities, increases the attractiveness of those regions for people and businesses. However, transport is only one of the factors affecting local development and there is no consensus on its net effect on population growth. The large scale of public funding allocated to motorway investment since the country joined the European Union in 1986 makes Portugal an ideal case study to examine the potential effect of improved road accessibility on the development of lagging rural areas. In this paper, we investigate the relationship between rural population change and road accessibility to the urban hierarchy (i.e. cities of different sizes) between 1991 and 2011. Regression analyses show that rural population growth is negatively associated with road distance and road travel time to the urban hierarchy, notably to medium-sized cities (i.e. 20,000-99,999 inhabitants). This suggests that medium-size cities play an important role in supporting population growth in their rural hinterlands. Robustness tests confirmed the validity of these findings. There is no evidence of nonlinearities in the magnitude of the effect between accessible and remote rural areas, which may be partially related to the relatively small size of the country.

Keywords: rural areas, population change, road accessibility, rural-urban linkages, spillover effects

JEL Classification*: R11, R12, J21*

1. Introduction

The nature of urban-rural relationships is complex and has changed over time in tandem with the improvements in information and communication technologies, leading also to an increasingly blurry distinction between what is urban and rural (Irwin et al., 2009). Whilst the theories and models explaining these linkages differ across disciplines (e.g. economics, geography, planning), there is a general belief that proximity to cities can benefit surrounding rural areas, by providing markets for goods produced in rural areas and a larger and more diversified pool of employment opportunities and services (e.g. public services, hospitals, schools). At the same time, rural areas supply cheaper housing and offer a better quality of life (e.g. less pollution, easier access to green space), favouring out-migration of urban workers to rural areas.

The nature and scope of urban-rural interactions, however, is likely to differ with the distance to the nearest city and with city size. Partridge et al. (2007) proposed a conceptual framework for analysing rural population change in the context of rural-urban linkages. This framework describes two types of urban-rural linkages. The first type is the urban distance discount, i.e. the effect of distance to cities on rural population change, regardless of whether those cities experience growth or decline. The second type of linkages is the spillover effect of urban growth on nearby rural areas. This effect can be positive (known as "spread effect") or negative (known as "backwash effect"). The prevalence of one type of effect over the other depends not only on the distance between rural and urban areas, but also on the size and characteristics of both areas. There is longstanding support, going back to the central place theory (Christaller, 1933), that the range and diversity of functions offered by cities of different size is proportional to their size. Larger cities provide a wider spectrum of economic and social functions (i.e., more specialized services and jobs), increasing even further the scope for rural-urban interactions.

In this paper, we investigate the relationship between rural population change and road accessibility to cities of different sizes (henceforth "the urban hierarchy") in mainland Portugal (excluding the

island regions of Açores and Madeira). We focus on the period between 1991 and 2011, covering 20 years of growing demographic and economic asymmetries between the urban coastal areas and the rural interior of the country. Road accessibility in Portugal was poor until the 1980s, but it has vastly improved since the country joined the European Union in 1986 and gained access to structural funds, a large proportion of which were allocated to improving road transport. According to Pereira and Pereira (2017), investment in roads grew from 0.74% of Portugal's GDP in the period 1980-1989, to 1.32% in 1990-1999 and then to 1.52% in 2000-2009. Investment on motorways alone corresponded to 0.07%, 0.30% and 0.59% of GDP, respectively, for the same periods. The result was a rapid expansion of motorways and dual carriageway roads (see Figure 3). In contrast, the rail network shrank during this period due to the closure of several railway lines and the reduction of services in others (Anciães, 2013). It was expected that the investment in roads would drive economic growth at the national level, while also fostering territorial cohesion by improving the connection between the less developed rural regions in the interior and the more developed urban regions along the coast. Despite this expectation, over the last 30 years there has been an even greater concentration of population and economic activities in cities in the coastal areas, and a decline in rural areas in the interior (Teixeira, 2006, Sousa et al., 2011).

Previous studies have looked at regional population change in mainland Portugal (e.g. Santos et al., 2013, Ribeiro and Silva 2011a, Ribeiro and Silva, 2011b, Fontes et al., 2014, Anciães, 2016). However, as noted in the next section, the evidence obtained in these studies is not conclusive. Furthermore, the studies used large spatial units (e.g., provinces), which are highly heterogeneous, and did not compare the effect of accessibility of rural areas to cities of different sizes. To tackle these limitations, the present study uses small spatial units and considers road-based accessibility to cities of different sizes. We use data from the 1991, 2001 and 2011 population census, combined with spatial data for the road network and other variables, to investigate how population change in rural

areas is affected by road accessibility to the urban hierarchy, capturing differences in the range of functions (goods, services, jobs, etc.) provided in cities of different sizes.

The paper makes three contributions to the literature on urban-rural linkages. The first contribution is to measure accessibility in terms of distance and time on the road network, thus accounting not only for geographic remoteness but also for the role of the road system in overcoming it. The second contribution is to consider the effects of road accessibility to the whole urban hierarchy, and not just to the nearest city: this allows us to test if and how city size affects population growth in rural hinterlands. The large and rapid investment in the motorway network makes Portugal a particularly interesting case to study the role of transport accessibility on the development of lagging rural areas.

The remaining of the paper is organized as follows. Section 2 is an overview of the literature on rural population change in the context of rural-urban linkages. Section 3 presents the data and the empirical strategy. Section 4 provides descriptive statistics for rural population change and road accessibility in Portugal over the period studied, while section 5 reports and discusses the results from the regression analyses. Section 6 summarises the lessons learnt, discusses policy implications, and offers directions for future research.

2. Overview of relevant literature

Table 1 is an overview of previous empirical studies of rural-urban linkages focusing on rural population change. It summarises existing evidence in terms of the types of data used, period studied, research methods, and main findings. Overall, these studies show positive (spread) effects of urban proximity on rural population growth. One of the major mechanisms underlying this result is rural out-commuting – i.e. individuals work in cities but live in nearby rural areas due to lower housing costs and preferences for natural amenities and better environmental quality (e.g. clean air, less noise, more appealing landscapes). The literature also tends to find that spread effects diminish to zero and may even turn negative with increasing remoteness. However, the tipping point beyond

which these effects prevail may differ with city size (Barkley et al., 1996, Ganning et al., 2013, Berdegué and Soloaga, 2018).

An alternative explanation for the presence of spread effects is business relocation from urban to rural areas due to the lower cost of land and real estate of the latter. The importance of these factors, however, is likely to differ across industries and be less salient for sectors with higher intensity of knowledge spillovers, human capital, and specialised input-output linkages, all of which are dependent on good access to urban agglomerations (e.g. Barkley et al., 1996). The literature suggests that the benefits spreading from production-side urban agglomeration economies towards accessible (i.e. commutable) rural areas render them less dependent on local job growth. The hypothesis that access to urban jobs can become a substitute for local rural jobs in sustaining local population levels is not novel (see Partridge et al., 2010 for a discussion). However, the extent to which out-commuting contributes to the growth of rural areas depends on where rural dwellers spend their wages. If the increased demand for goods and services takes place in the cities where the rural dwellers work, the benefits for local areas will be likely limited to the housing sector, with limited impact on the services and retail sectors. Rural areas may even end up having a role similar to suburban “bedroom communities”, albeit with lower population densities and less access to services (e.g. Lavesson, 2017).

Although out-commuting cannot be considered as a plausible growth strategy for remote rural areas, these areas may still enjoy population growth due to in-migration of households seeking lower house prices, rural amenities and better quality of life (Benson and O'Reilly, 2009). However, even with growing trends for remote working (intensified during the COVID-19 pandemic), and the expansion of information and communication technologies even in remote areas, lifestyle-driven rural in-migration is still limited because only a small proportion of the urban population can relocate to rural areas (e.g., workers with flexible occupations, freelancers, retirees).

There is very limited and inconclusive evidence for Portugal on the mechanisms underlying urban-rural linkages and the specific role played by road accessibility in that relationship. This is especially surprising given the massive improvement in the country's road network over the last 30 years. Some studies have found that the relationship between population change and accessibility is statistically insignificant (Santos et al., 2013), but others have found it to be negative (Ribeiro and Silva, 2011a, Ribeiro and Silva, 2011b), or indeterminate, depending on the model specification (Fontes et al., 2014, Anciães, 2016). With the exception of Anciães (2016), previous studies have focused on a specific subset of case study regions, rather than the full sample of rural areas. Furthermore, these studies did not consider the relationship between rural population change and accessibility to urban areas of different sizes, failing to capture heterogeneity resulting from differences in urban agglomeration economies. The present paper contributes to the existing literature by analysing urban-rural linkages from the perspective of road-based accessibility to urban areas. We take into account travel distances and travel times on the road network from rural areas to the full hierarchy of cities, which allows accounting not only for geographic isolation but also for the role of the road system in overcoming such isolation.

Table 1: Overview of empirical studies of rural population change and rural-urban linkages

Study	Data	Dependent variable & period	Methods	Main findings
Barkley et al. (1996)	Census tracts of FEA across three southern states of the U.S.	Change in rural population density 1980-1990	Population density-distance functions based on cubic spline regression models	Positive spillover effects on rural population growth especially for rural areas at the urban fringe, i.e. "spread through decentralization".
Henry, Barkley and Bao (1997)	Census tracts of FEA across three southern states of the U.S.	Change in rural population density between 1980-1990	Boarnet (1994) extension of Carlino and Mills (1987) model - estimation of a system of 2 equations for changes in population density and employment density	Positive spillover from employment growth in urban core and fringe; effect is stronger when urban fringe growth is greater than urban core growth. The effect of distance to the urban core alone is non-significant
Henry, Schmitt and Piguet (2001)	Communes (municipalities – LAU-1) of FER in eastern France	Change in rural population density 1982-1990	Same as Henry, Barkley and Bao (1997)	Positive spillover effect from employment growth in urban fringe on rural population change. Negative spillover (backwash) effect from employment growth in urban core.
Schmitt et al. (2006)	Cantons (NUTS-III) of FER in eastern France	Change in rural population density 1982-1990	Boarnet (1994) extension of Carlino and Mills (1987) model - estimation of a system of 3 equations for change in population density, employment density of exports sector, and employment density of services sector	Positive spillover effects from growth on urban service jobs over rural population change for FER with declining core. No effects for FER with growing core, nor from growth on urban exports jobs
Partridge et al. (2007)	Canadian census subdivisions, nationwide	Population change 1981-1991,1991-2001, 1981-2001	Cross-sectional regressions using OLS and SEM estimators. Separate analysis for samples of rural areas of different size (>1500 vs. <=1500 pop) and distance to urban areas (<50 km vs. >=50 km)	Strong positive effect of urban access (urban distance discount). In addition, spread effects dominate up to 175 km, and backwash effects dominate thereafter.
Partridge et al. (2008)	U.S. county data, nationwide	Population change 1950-2000	Population change regressions using GMM estimators, controlling not only for distance to nearest urban area but also incremental distance to higher-tier urban hierarchy	Negative effect of distance to the nearest urban area, increasing with size of urban area
Veneri and Ruiz (2016)	OECD TL3 regions in Europe, North America and South America	Population change 2000-2008	Cross-sectional regressions using OLS, spatial nonparametric approach PS-GAM, and SAR estimators	Spread effects outweigh backwash effects; spread effects decline with distance.
Berdegué and Soloaga (2018)	Rural localities (less than 15,000 people) in Mexico	Population change 2000-2010	Cross-sectional regressions using OLS and dummy variables for state-level effects, distance to the nearest urban location and urban locations of different sizes.	Positive effect of proximity to city on population growth, especially cities with 350,000–500,000 people; rural areas interact with multiple cities. Spread effects vanish at 2 (3) hr travel time from cities with 15,000-49.999 (350,000 plus) people.

Notes: FEA-Functional Economic Areas; FER-Functional Economic Regions; GMM-Generalized Methods of Moments; OLS-ordinary least squares; PS-GAM-p-spines generalized additive model; SAR-spatial autoregressive model; SEM-spatial error model; TL3-Territorial Level 3 regions.

3. Data and methods

3.1. Spatial units, scope of analysis, and variables

Spatial units and scope

Using data from the 1991, 2001, and 2011 Portuguese population censuses, we constructed a dataset at the level of *freguesias* (civil parishes). These are the smallest administrative units in Portugal, corresponding to Local Administrative Units of level 2 (LAU2) in the European NUTS regional classification system. To take account of changes in the boundaries of *freguesias* between census periods, we harmonised the data with reference to the situation in 2011, when there were 4050 *freguesias* in mainland Portugal. To delimit the scope of our analyses, we applied the administrative classification of urban-rural areas 'Typology of Urban Areas' (TIPAU) developed by the Portuguese National Statistics Institute (INE), which follows a similar logic to OECD's regional typology of small administrative areas (Brezzi et al., 2011). The TIPAU classification defines *freguesias* as "predominantly urban areas" (5,000 or more inhabitants), "moderately urban areas" (2,000-4,999 inhabitants), and "predominantly rural areas" (less than 2,000 inhabitants).[1] Our analysis focuses on the predominantly rural areas. In 2011, these areas accounted for 25% (i.e., 2078) of all *freguesias* in mainland Portugal. Their average population in 1991, 2001, and 2011 was 808, 744, and 663 inhabitants, respectively. The median population was 563, 495, and 423 inhabitants. Appendix A shows the TIPAU classification of freguesias in mainland Portugal.

Demographic, socio-economic, accessibility, and natural environment variables

The analyses used the variables below. The data sources are described in Appendix B.

- Demographic variables: population size and density; share of the population aged 65 years or more.

[1] http://smi.ine.pt/Versao/Download/10129.

- Socio-economic variables: share of population with higher education (treated in this paper as an indicator of human capital); share of employment in the tertiary sector; unemployment rate.
- Accessibility: road-based distance and travel time from rural areas to the nearest city of different size, number of railway stations.
- Natural environment variables: proportions of area designated as a site in the European Natura 2000 network (Special Areas of Conservation - Habitats and Special Protection Areas - Birds); and in the RAMSAR International Network of Wetlands; Shannon-Wiener index of diversity of natural landscapes and habitats (Spellerberg and Fedor, 2003); and the standard deviation of slopes steepness as an indicator of topographical irregularity.

Table 2 shows descriptive statistics of these variables. The average population density decreased over the period 1991-2011, and the share of elderly in total population increased. The average share of people employed in the tertiary sector remained stable over the period. The average unemployment rate was similar in 1991 and 2001, but nearly doubled in 2011. The average share of people with higher education increased substantially, while remaining at a relatively low level. Road accessibility improved considerably over the overall period, both in terms of road distance and driving times to nearby cities of different sizes. In contrast, there were no significant improvements in railway access to rural areas.

Table 2. Descriptive statistics of variables describing rural freguesias in 1991, 2001 and 2011

Variables/Descriptive statistics	1991			2001			2011		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Population density (people/km2)	45.49	34.69	35.46	41.61	30.02	34.54	36.87	25.76	32.28
Percentage of population with higher education	0.59	0.47	0.64	2.12	1.88	1.51	5.06	4.72	2.86
Percentage of population aged 65 plus	22.59	21.47	7.10	28.53	27.01	9.07	33.39	31.43	10.43
Percentage of employed in tertiary sector	52.85	52.00	17.59	54.01	54.05	14.04	53.41	54.08	12.78
Unemployment rate	5.92	4.10	6.39	7.72	6.30	5.86	12.36	11.63	6.35
Number of active train stations	0.10	0.00	0.39	0.09	0.00	0.35	0.06	0.00	0.29
Road distance to nearest small city [10,000-20,000[, kms	28	24	17	27	24	16	27	24	16
Road distance to nearest small-to-medium city [20,000-50,000[, kms	48	46	24	46	44	22	46	46	23
Road distance to nearest medium-to-large city [50,000-100,000[, kms	106	95	60	102	92	53	104	94	54
Road distance to nearest large city with at least 100,000, kms	128	120	70	122	118	60	123	118	61
Travel time to nearest small city [10,000-20,000[, min	28	24	17	26	23	14	25	23	13
Travel time to nearest small-to-medium city [20,000-50,000[, min	46	44	23	41	40	19	37	36	17
Travel time to nearest medium-to-large city [50,000-100,000[, min	100	89	56	78	69	39	67	61	31
Travel time to nearest large city with at least 100,000 people, min	120	113	65	91	88	44	82	81	37
Percentage of area in European Natura 2000 Network	18.25	0	33.21	18.25	0	33.21	18.25	0	33.21
Percentage of area in International Network for Wetlands	0.34	0	3.49	0.34	0	3.49	0.34	0	3.49
Shannon indicator of landscape diversity	1.60	1.64	0.31	1.60	1.64	0.31	1.60	1.64	0.31
Standard deviation of slopes	4.91	4.89	1.95	4.91	4.89	1.95	4.91	4.89	1.95

Number of freguesias in each year: 2,078. SD: Standard deviation.

3.2. Empirical strategy

We adopted the approach used by Veneri and Ruiz (2016) and Partridge *et al.* (2007) to estimate how population change in rural areas is associated with road accessibility to cities of different sizes controlling for local conditions that can also influence rural population growth. Equation (1) represents the baseline model specification. We estimate this model for each period separately (1991-2001 and 2011-2011) and simultaneously (i.e. pooling the two sub-periods together).

$$\%\Delta POP_{i,(t-0)} = \alpha + \beta' DEM_{i,0} + \gamma' ECO_{i,0} + \delta' NAT_{i,0} + \rho' ACC_{i,0} + \theta_r + \varepsilon_{i,(t-0)} \quad (1)$$

where the subscripts *i*, *t*, *0*, and *r* indicate, respectively, the rural freguesia, the final year of the period of analysis, the initial year, and the wider region containing each freguesia (Portuguese distritos). The dependent variable ($\%\Delta POP_{i,(t-0)}$) measures the percentage change in population, which is equivalent to the percentage change in population density because the area of freguesias remains constant). The rate of population change is regressed on a set of explanatory variables for the beginning of each period. The main explanatory variables are transport accessibility ($ACC_{i,0}$), a vector which includes road distances and travel times to cities of different sizes, and the availability of rail stations. The control variables (as described in Section 3.1) include demographic structure ($DEM_{i,0}$), economic structure ($ECO_{i,0}$) and the natural environment ($NAT_{i,0}$). The vector θ_r is a set of fixed effects for region r (Portuguese *distrito*), capturing commonalities in freguesias in the same region. Finally, $\varepsilon_{i,(t-0)}$ is the error term, which allows for heteroscedasticity and clustering on rural freguesias.

Most explanatory variables were entered in the model as logs. This was to make their statistical distribution more symmetrical, mitigating the effect of possible outliers, and to interpret the parameter estimates as relative marginal changes. Since the dependent variable is the growth rate of rural population, the parameter estimates can be interpreted as the change in the population

growth rate of *beta* percentage points associated with a 1% increase in the explanatory variable. The number of train stations and the variables for the natural environment were not entered as logs due to the presence of many zeros.

The baseline model specification in equation 1 assumes that the slope of the curve between population growth rate and road accessibility remains constant regardless of the distance a given freguesia is to the nearest city of a given size. In other words, the model assumes that the effect of changing road accessibility is constant for freguesias near cities (i.e. accessible rural areas) and for freguesias far from cities (i.e. remote rural areas). The assumption of a linear effect regardless of distance range may not hold true, for example, if there are diminishing returns to improved road accessibility as the network develops. We used two methods to test this hypothesis. The first method was to add interaction terms between the road accessibility variables and dummy variables defining accessible vs. remote - using the approach developed by Dijkstra and Poelman (2008). The second method was to estimate semi-parametric models allow for a non-linear functional form for the relationship between population change and travel time, whilst maintaining a linear functional form for the remaining explanatory variables. The second approach also has the advantage of overcoming the arbitrariness in define accessible vs. remote rural areas. The results are reported in Sections 5.2 and 5.3 respectively, whereas the results for the baseline model specification are reported in section 5.1.

The estimation of the model in equation 1 may also suffer from endogeneity bias, in particular due to potential reverse causation between the placement of roads and rural development: i.e., road investment may be a function of demand-side factors such as population size or population growth, instead of the opposite. In this case, the parameter estimates for road access to the urban hierarchy will be biased and inconsistent. Simple approaches to address this identification issue include using time lags between population change and road accessibility (ruling out reverse direction in the relationship), and replacing road-based accessibility with straight-line distances (as these do not

depend on the location of road infrastructure). More sophisticated approaches consist on the use of causal inference techniques based on comparisons between treated and control groups and - the most frequent method – the use of instrumental variables (IV) methods The rationale of this method is to find variables (i.e., *instruments*) that help explain road accessibility but do not affect population change directly. Among the most common instruments used in the literature are historical roads, road plans and geographic factors that influence the placement of roads (e.g. altitude range and variation). These instruments have been used in studies of the impact of motorways on economic performance and urban growth for Spain (e.g., Garcia-López et al., 2015) and Italy (e.g., Percoco, 2016). In a recent study for Portugal, Rocha et al. (2020) found that geography (altitude range and morphological slope variability) and presence of historical roads (Roman roads and XIX Century roads) helped explaining the spatial distribution of motorways between 1981-2011, controlling for demand-side factors. This suggests that historical roads and geographical factors may work well as instruments for road accessibility in our analysis.

To investigate whether our baseline model specification may suffer from endogeneity bias, we have therefore implemented some of these approaches, namely: i) replacing road-based proximity to the urban hierarchy with straight-line distances; ii) modelling population change between 2001-2001 as a function of road accessibility in 1991 and the change in road accessibility between 1991 and 2001 (to reduce the scope for reverse causation bias); and iii) estimating constrained versions of IV regressions due to the mismatch between the number of instruments and the number of endogenous variables. We report and discuss the results from these endogeneity checks in Section 5.4.

4. Descriptive statistics of rural population and accessibility

4.1. Evolution of rural population change

Figure 1 shows the rate of population change over the periods 1991-2001 and 2001-2011 for all freguesias in mainland Portugal. In both periods, there is a clear distinction between the coastal areas (in the West and South) and the country's hinterland, with population decline mostly occurring in the latter. Figure 2 plots these rates according to the typology of urban areas described in Section 3.1. Although there is considerable variation in population changes in each of the three types of areas, rural areas show a higher concentration of freguesias in the bottom left quadrant (i.e., population decline in both periods), compared with predominantly urban and moderately urban areas. In contrast, predominantly urban areas show a higher concentration of freguesias in the top-right quadrant (population growth in both periods).

Table 3 reports the mean-weighted population rates of change by type of area and period. Predominantly rural areas are characterized by negative population growth rates in all periods (-1.39% for 1991-2001, -1.65% for 2001-2011, and -3.14 for 1991-2011). In contrast, population in urban areas grew in all periods (but less in 2001-2011) and population in moderately urban areas grew slightly in 1991-2001 (0.54%) and declined in 2011-2011 (-0.18%).

Figure 1. Population change across freguesias: 1991-2001 (left panel), 2001-2011 (right panel)

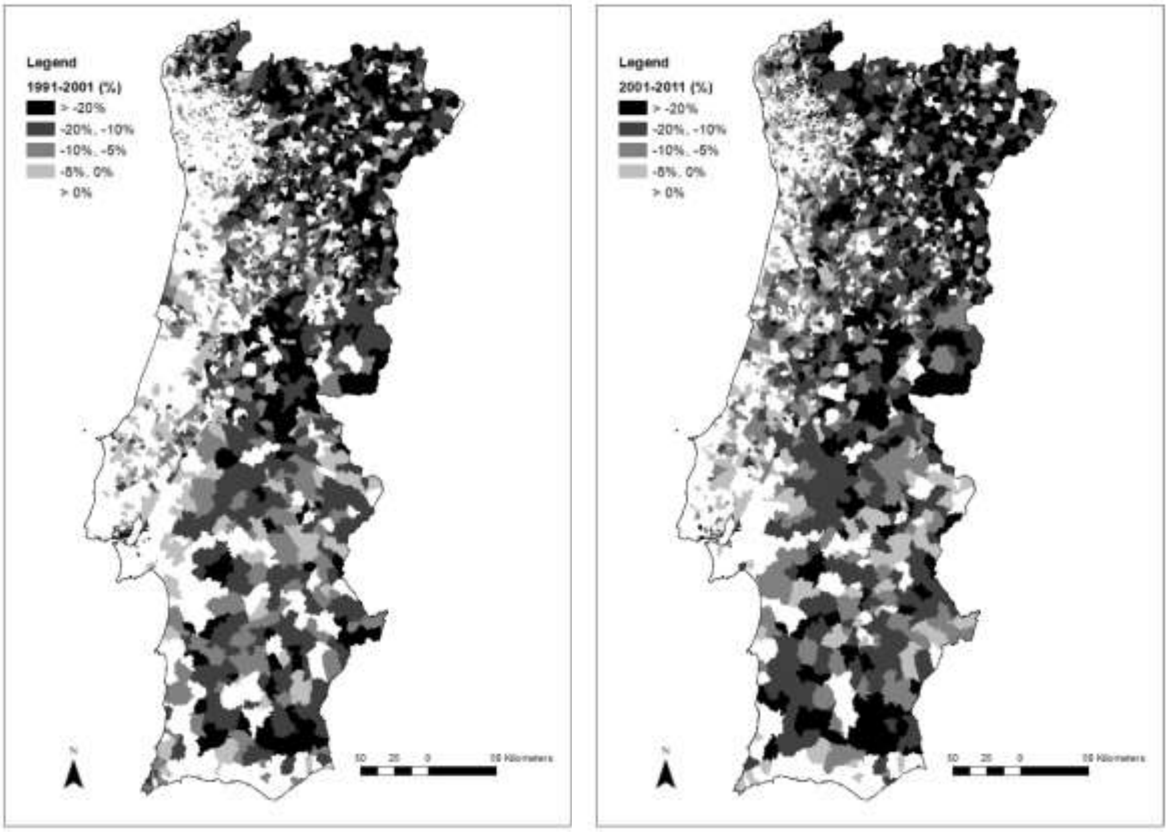


Figure 2. Population change in 1991-2001 and 2001-2011 across freguesias, by rural-urban typology

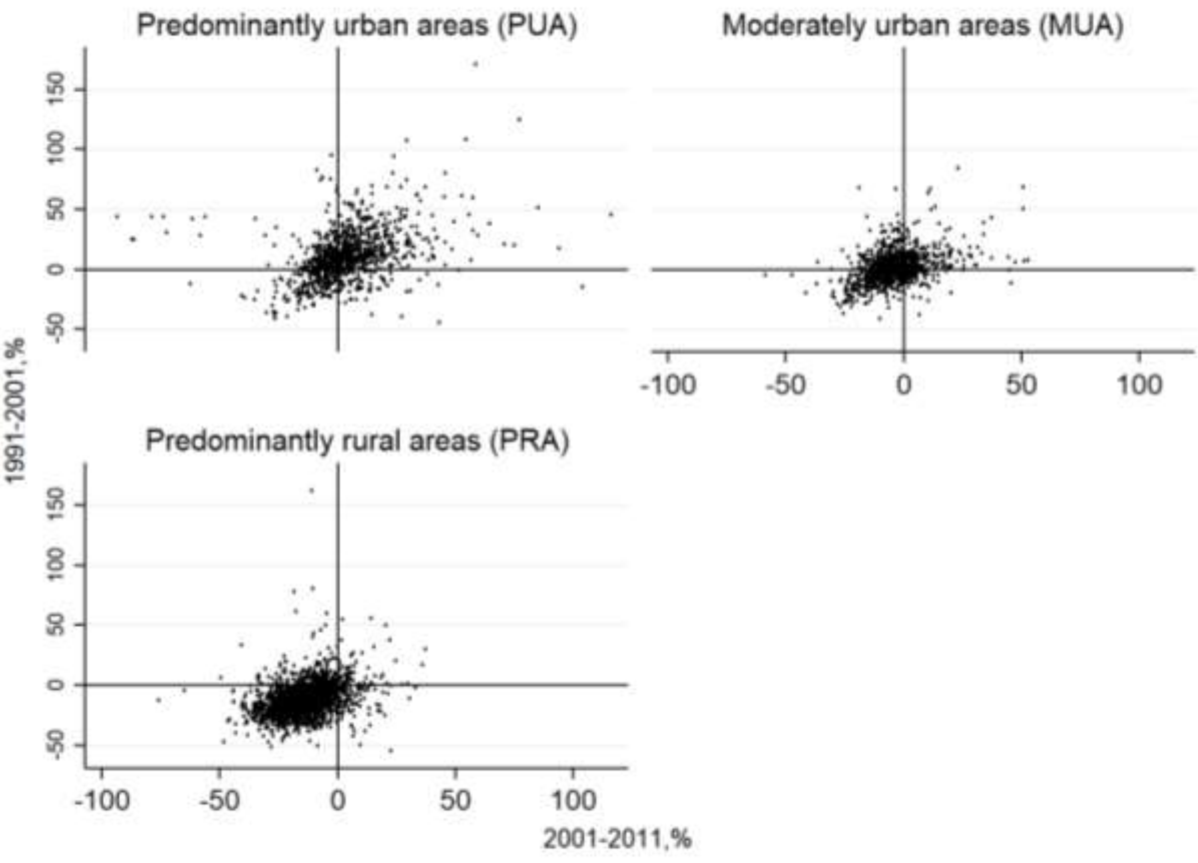


Table 3. Weighted population change across freguesias by rural-urban typology, in percentage

Periods	Predominantly Urban Areas (PUA)	Moderately Urban Areas (MUA)	Predominantly Rural Areas (PRA)	All freguesias
1991-2001	6.92	0.54	-1.39	6.07
2001-2011	0.46	-0.18	-1.65	-1.37
1991-2011	7.41	0.35	-3.14	4.62

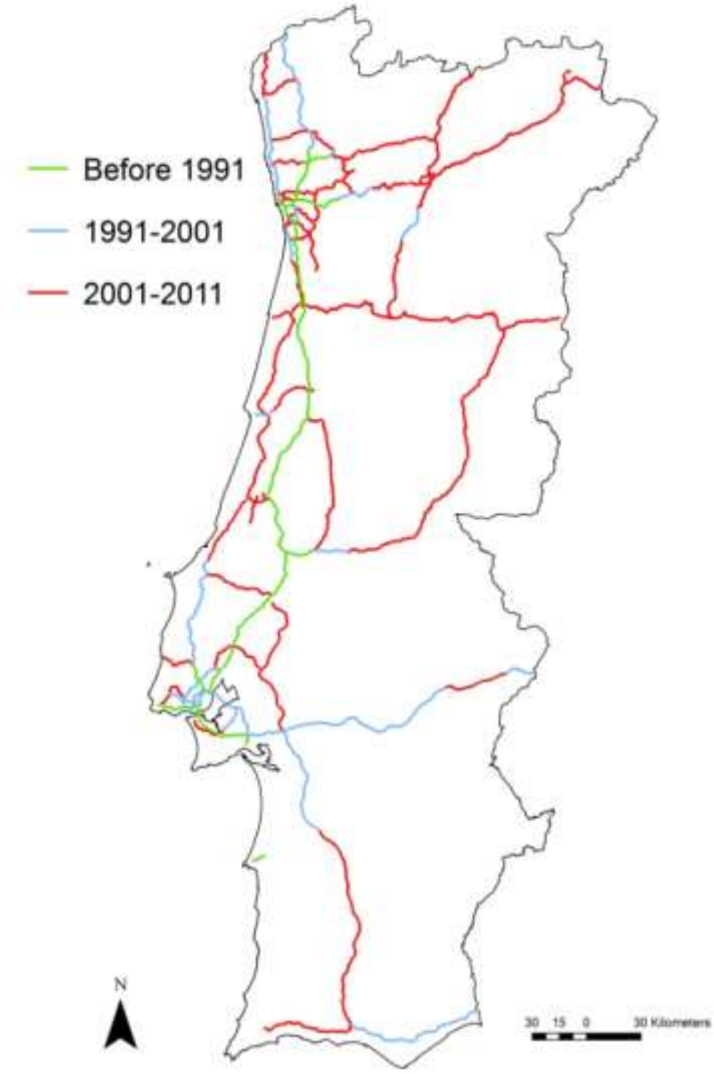
4.2. Evolution of road accessibility from rural areas to the urban hierarchy

Accessibility (the ease of accessing places) can be measured with different indicators, all with their own positive and normative assumptions (Paez et al., 2012). As mentioned in the previous section, in this paper we used road distance and travel times to the nearest city of a given size. We constructed bespoke road models for 1991, 2001, and 2011. These models used as a base the road map layer of the 1999 topographic map produced by the Portuguese Army Geospatial Information Centre (IGeoE, 1999). We then created the 1991 map by removing the motorways and dual carriageways that were built between 1991 and 1999, and the 2001 and 2011 maps by adding those built after 1999. The information on the evolution of the network of motorways and dual carriageways was extracted from commercial road maps and the maps published by the Portuguese Institute of Road Infrastructure (IEP, 1990-2004, EP, 2005-2011). We then assigned travel speeds to each road section based on speed limits for each type of road, using the classification in the National Road Plans of 1985 and 2000, and information on whether the road section cut across built-up areas. A shortest route algorithm was then implemented using *ArcGIS 10.6* to calculate the shortest road distances and travel times from each *freguesia* to the cities in mainland Portugal in 1991, 2001, and 2011. The georeferenced information for freguesias and cities was extracted, respectively, from the Portuguese Official Administrative Map (DGT, 2011) and a report by the National Statistics Institute (INE (2014). The locations of cities were identified as the capitals of the municipalities to which each of the freguesias belong (or the main urban nodes of each freguesia, when the city was not capital of any municipality).

Figure 3 shows the evolution of the Portuguese network of motorways and dual carriageways between 1991 and 2011. To illustrate how the changes in the motorway network affected road accessibility, Figure 4 shows the travel times to the two top tiers of the urban hierarchy (i.e., >=100,000 people in top panel, >=50,000 people in bottom panel) for all freguesias in 1991 (left panel), 2001 (middle panel) and 2011 (right panel). The reduction of travel times is clear.

We could not construct similar accessibility measures for rail transport. Rail travel times depend not only on the speed of each rail service, but also on the frequency of services at different times of the day and days of the week. Rail accessibility also depends on the possibility of making day return trips and on the availability of bus services to access rail stations. These variables could not be calculated due to the lack of information. As an alternative, we estimated an indicator of availability of rail transport as the number of rail stations in operation in each freguesia. The large majority of rural freguesias do not have access to railway services: 95.4% of rural freguesias did not have any functioning train stations in 2011, 3.4% had one station, and 1.2% had two or three stations. Although this variable does not measure accessibility, but only availability of stations in operation, it can still help capture the process of railway closures that occurred in the period studied (Anciães, 2013). The extent of railway lines in operation between 1991 and 2011 shrank by about 11%. If we consider the latest data for 2019, we observe that the network shrank by 19% since 1991.

Figure 3. Evolution of the Portuguese network of motorways and dual carriageway roads 1991-2011



The trends in rural population change presented so far consider freguesias as isolated spatial units. However, as already discussed, accessible rural areas tend to outperform remote rural areas in terms of demographic and economic growth by taking advantage of nearby urban areas. In the analysis that follows, we classified cities in four groups that provide a balanced distribution while capturing well the full range of city sizes: small (10,000-19,999 inhabitants); small-to-medium (20,000-49,999); medium-to-large (50,000-99,999); and large (>=100,000). Between 1991 and 2011, the number of large cities remained constant (i.e., 6 large cities), whilst it increased from 7 to 10 for medium-to-large cities, 30 to 38 for small-to-medium size cities, and 39 to 90 for small cities.

Figure 5 shows the percentage of rural freguesias within 30-, 45-, and 60-minutes travel time to the nearest cities. The charts suggest that the expansion of the motorway network during the 1990s and 2000s (Figure 3) contributed at least partially to the improvement in the accessibility of rural areas to cities of different sizes. The share of rural areas within 45 minutes to large and medium-to-large cities increased since 1991 but was still small in 2011 (20% and 27% respectively). If we consider the range of up to 30 minutes, the values are even lower (7% and 10% in 2011 for large and medium-to-large cities respectively).

The pairwise correlations between rural population change and accessibility to urban areas are all negative and statistically significant at the 1% level, suggesting that the rate of population decline increases with distance from cities. The correlation coefficient for the period 1991-2001 is equal to -0.27 for small cities, -0.28 for small-to-medium cities, -0.32 for medium-to-large cities, and -0.28 for large cities. The respective coefficients in the period 2001-2001 are slightly higher in absolute value (-0.35, -0.33, -0.35, and -0.28. The scatter plots of population change and accessibility to cities are provided in Appendix C.

Figure 4. Road travel times (minutes) to cities with at least 100,000 people (top panel) and at least 50,000 people (bottom panel), in 1991 (left), 2001 (middle) and 2011 (right)

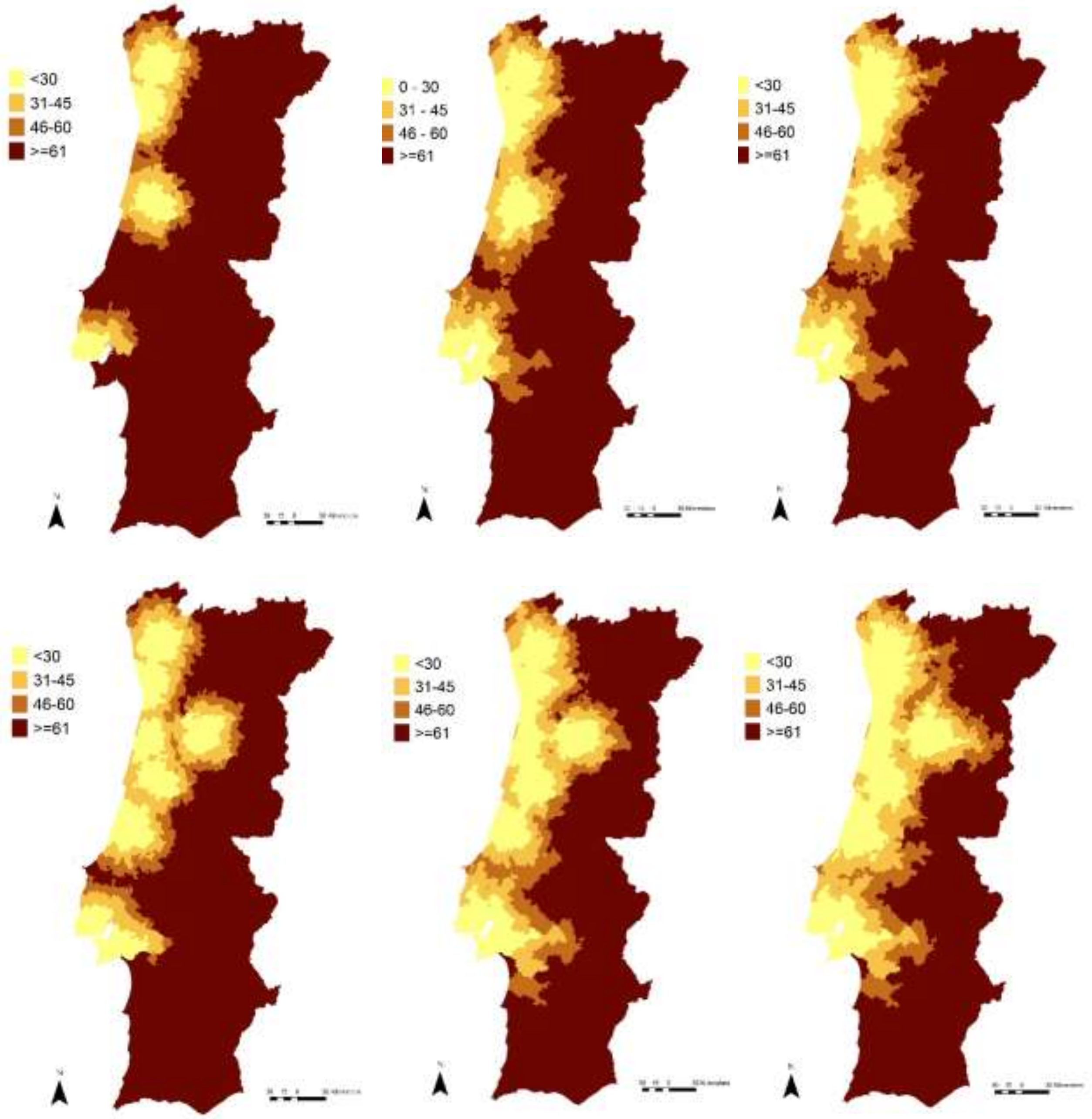
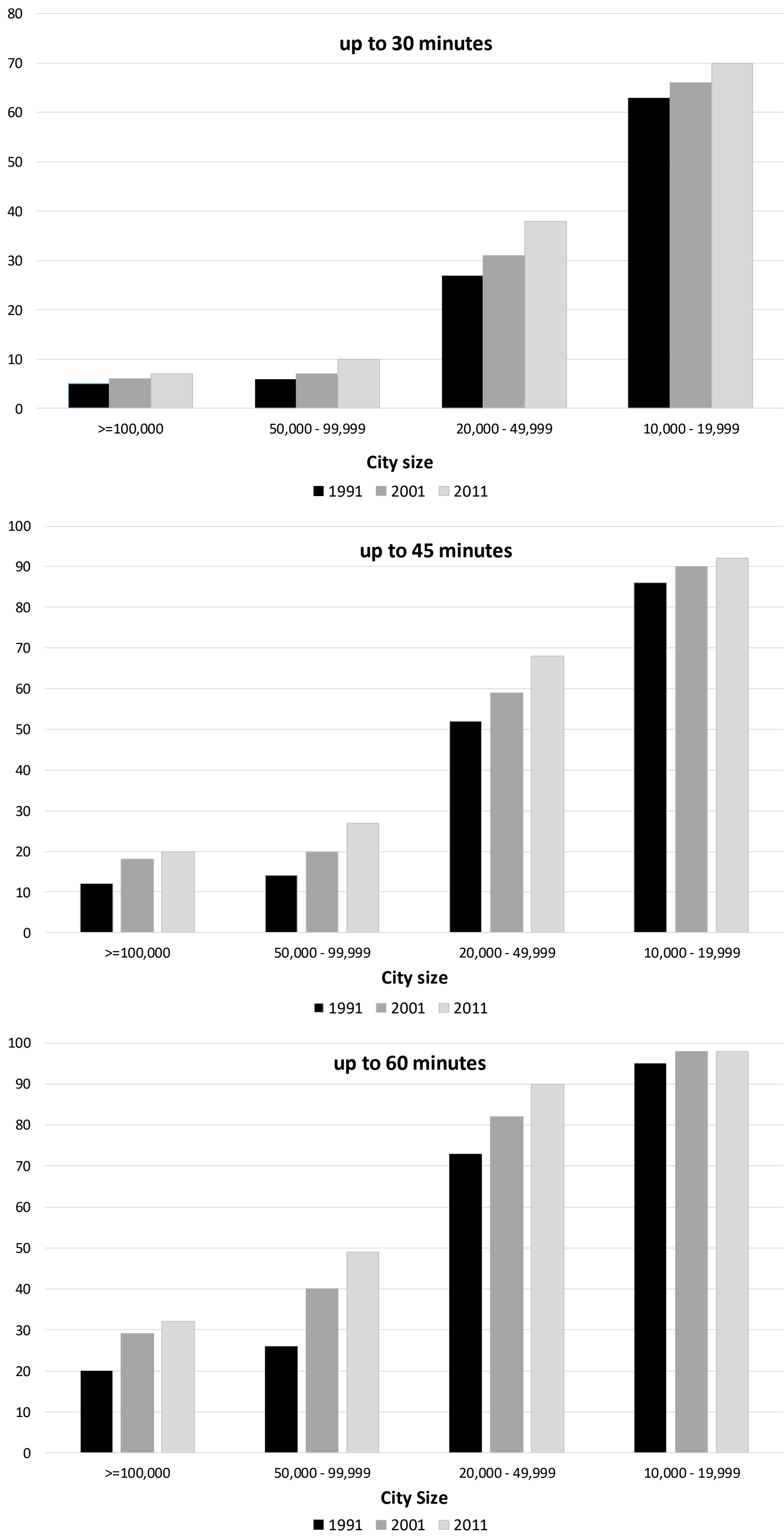


Figure 5. Share of rural freguesias within 30-, 45-, and 60-minutes road travel time to the urban hierarchy



5. Results and discussion

5.1.Population change across all rural areas

Table 4 shows the results from the ordinary least squares regression models pooling all rural freguesias. Models 1-3 use road distances while Models 4-6 use road travel times. Models 2 and 5 refer to the period 1991-2001; Models 3 and 6 refer to the period 2001-2011; and Models 1 and 4 pool those two periods together. The specifications used explain between 26%-28% of the variation in population change across rural freguesias.

Rural population growth is negatively associated with travel distance and time to medium-size cities (small-to-medium size and medium-to-large size) across all models. The evidence for other city sizes is weaker: rural population growth is negatively associated travel distance and time to small cities in 1991-2001, while it is positively associated with travel distance to large cities in 1991-2001. Considering the models that pool data for both periods, a reduction of 10% in road distance to small-to-medium size cities or to medium-to-large size cities is associated with an increase in the population growth rate of 0.31 percentage points (i.e. model 1). The corresponding figure for the models using travel time (i.e. model 4), we observe that a reduction of 10% in driving time is associated with an increase in the population growth rate of 0.33 percentage points.

The results described above for road accessibility to the urban hierarchy suggest that medium-size cities in Portugal are important in sustaining population growth in their rural hinterlands. The group of medium-size cities in Portugal includes most of the cities that are district capitals. Districts have served as the basis for the spatial distribution of major services of general interest (e.g. courts, post offices, health, education) since the second half of the XIX century. The presence of these services led not only to the creation of public sector jobs, but also greater population and private sector employment. Elsewhere, Berdegué and Soloaga (2018) also found that rural population growth is stronger for urban proximity to medium-size cities compared to smaller and larger cities in Mexico (although in absolute size, medium-size cities are much larger in Mexico than in Portugal).

As for the role of railway access, the results show that the effect of having active train stations is only significant in Model 5 (at 10% level). While this result suggests that the presence of rail services has little influence on rural population growth, it should be noted that this variable is an imperfect indicator of rail accessibility, since all stations were treated as being equal regardless of the level of service they offer to local communities, as mentioned in Section 3.1.

Considering the importance of the demographic and socioeconomic features of rural areas, the results obtained are also in line with expectations and previous studies. The rate of population change is positively associated with the initial level of population density, suggesting there is some persistence in population trends. Likewise, population growth is negatively associated with higher shares of older population (i.e., aged 65 or more) and positively associated with the share of population with higher education in each sub-period but not the whole period) and with the share of employment in the tertiary sector. In Models 2 and 5, population growth is negatively associated with higher initial values of unemployment rates.

With regards to the natural environment variables, landscape diversity has a positive and significant association with rural population growth in all models. Morphological unevenness, measured by the standard deviation of slope steepness, has a negative association with rural population growth, which may partially reflect the fact that population settlements tend to develop in locations with less rugged terrain. Finally, the results for the presence of high-value natural sites is not conclusive: the relation is not significant for the percentage of area in the European Natura 2000 Network, while it is significant and negative for the percentage of area in the International Network for Wetlands. While the presence of sites with natural interest may attract visitors, this does not necessarily mean a positive impact on population growth; in fact, the greater the share of land classified as having natural protection status, the lower the supply of land available for urbanization.

Table 4. Regressions of rural population growth rates

	Models using road distance			Models using travel time		
	(1) Pooled	(2) 1991-2001	(3) 2001-2011	(4) Pooled	(5) 1991-2001	(6) 2001-2011
Initial distance to nearest small city [10,000-20,000[(in log)	-0.0068 (0.0045)	-0.0161** (0.0073)	0.0023 (0.0047)			
Initial distance to nearest small-to-medium city [20,000-50,000[(in log)	-0.0312*** (0.0043)	-0.0315*** (0.0063)	-0.0317*** (0.0051)			
Initial distance to nearest medium-to-large city [50,000-100,000[(in log)	-0.0311*** (0.0065)	-0.0486*** (0.0107)	-0.0140** (0.0069)			
Initial distance to nearest large city with at least 100,000 people (in log)	-0.0075 (0.0066)	-0.0180* (0.0102)	-0.0031 (0.0073)			
Initial travel time to nearest small city [10,000-20,000[(in log)				-0.0079 (0.0050)	-0.0199** (0.0081)	0.0026 (0.0052)
Initial travel time to nearest small-to-medium city [20,000-50,000[(in log)				-0.0334*** (0.0045)	-0.0349*** (0.0066)	-0.0365*** (0.0055)
Initial travel time to nearest medium-to-large city [50,000-100,000[(in log)				-0.0327*** (0.0071)	-0.0511*** (0.0110)	-0.0167** (0.0078)
Initial travel time to nearest large city with at least 100,000 people (in log)				0.0092 (0.0073)	-0.0173 (0.0106)	0.0000 (0.0088)
Initial population density (in log)	0.0112** (0.0047)	0.0125* (0.0075)	0.0077 (0.0051)	0.0117** (0.0047)	0.0112 (0.0076)	0.0070 (0.0051)
Initial percentage of population with higher education (in log)	0.0039 (0.0037)	0.0252*** (0.0095)	0.0103* (0.0057)	0.0001 (0.0040)	0.0234** (0.0095)	0.0095* (0.0057)
Initial percentage of population aged 65 plus (in log)	-0.0882*** (0.0109)	-0.0689*** (0.0186)	-0.0963*** (0.0124)	-0.0929*** (0.0109)	-0.0660*** (0.0183)	-0.0957*** (0.0125)
Initial percentage of employed in tertiary sector (in log)	0.0137*** (0.0053)	0.0135** (0.0060)	0.0174* (0.0102)	0.0142*** (0.0053)	0.0132** (0.0059)	0.0175* (0.0102)
Initial unemployment rate (in log)	-0.0029 (0.0026)	-0.0073* (0.0042)	0.0021 (0.0036)	-0.0034 (0.0026)	-0.0078* (0.0042)	0.0020 (0.0036)
Initial number of active train stations	-0.0023 (0.0051)	-0.0107* (0.0065)	0.0038 (0.0064)	-0.0020 (0.0050)	-0.0113* (0.0065)	0.0033 (0.0064)
Percentage of area in European Natura 2000 Network	0.0000 (0.0001)	0.0000 (0.0001)	0.0000 (0.0001)	0.0000 (0.0001)	0.0000 (0.0001)	0.0000 (0.0001)
Percentage of area in International Network for Wetlands	-0.0016*** (0.0005)	-0.0017** (0.0008)	-0.0014*** (0.0005)	-0.0016*** (0.0005)	-0.0017** (0.0008)	-0.0014*** (0.0005)
Shannon indicator of landscape diversity	0.0235*** (0.0061)	0.0294*** (0.0083)	0.0175** (0.0080)	0.0210*** (0.0061)	0.0272*** (0.0083)	0.0162** (0.0079)
Standard deviation of slopes	-0.0074*** (0.0012)	-0.0072*** (0.0018)	-0.0074*** (0.0016)	-0.0069*** (0.0012)	-0.0065*** (0.0018)	-0.0069*** (0.0016)
Constant	0.3777*** (0.0701)	0.4723*** (0.1170)	0.2532*** (0.0701)	0.3405*** (0.0697)	0.5003*** (0.1204)	0.2619*** (0.0689)
Controls for district	YES	YES	YES	YES	YES	YES
Observations	4156	2078	2078	4156	2078	2078
Adjusted R^2	0.2582	0.2567	0.2787	0.2580	0.2603	0.2813

Notes: * p<0.1; ** p<0.05; *** p<0.01. Standard errors are reported in parentheses. All models include controls for regions (Portuguese distritos).

5.2.Population change across accessible and remote rural areas

To investigate whether the effect of road accessibility on population change varies with travel distance or time to cities, we estimated a new model that includes interaction terms between road distance or travel time and a dummy variable indicating whether a given rural freguesia is considered "accessible" or "remote". This approach corresponds to asking the question: How does the effect of a reduction in travel distance or time on population change differ between accessible and remote rural areas? We classified rural areas as accessible or remote using the definition proposed by Dijkstra and Poelman (2008). Accessible areas are those less than 30-, 45-, or 60-minutes away from the nearest city with over 50,000 people (corresponding to the "large" and "medium-to-large" cities in Portugal). Considering the 2078 rural freguesias that existed in 2011, using the cut-off points of 30 minutes, 45 minutes and 60 minutes results in sets of 35, 204, and 555 accessible rural freguesias, respectively.

Table 5 reports the results of the new models. The results are in line with the findings from the baseline models in the previous section, reinforcing the importance of proximity to medium-size cities for rural areas compared to both small cities and large cities. There are, nevertheless, some nuances on the nature of the relation between accessible and remote rural areas depending on the definition of the cut-off point. For the 30-minute cut-off point, none of the interaction terms is statistically significant, indicating that the average marginal effect of road accessibility to the urban hierarchy is the same for accessible and remote rural areas. For the 45- and 60-minutes thresholds, the only differences are for small-to-medium size cities: the magnitude of the effect of road accessibility on population change is smaller for accessible rural areas compared to remote rural areas. For the 45-minute threshold, reducing road distances (travel times) to small-to-medium size cities by 10% is associated with an average increase in the growth rate of population of 0.13 (0.01) percentage points for accessible rural areas and of 0.33 (0.36) percentage points for remote rural areas. The corresponding effect for the 60-minute threshold is 0.28 (0.17) percentage points for

accessible rural areas and 0.35 (0.38) percentage points for remote rural areas. Furthermore, there is weak evidence that isolation from larger cities may actually benefit rural population growth: the coefficient is only significant when remoteness is defined using the 60-minute threshold and for the travel time model. Isolation from large cities can function as a protection from competitive forces.

5.3.Testing for nonlinearities in the effect of travel time to cities on rural population change

To avoid selecting potentially arbitrary cut-off points for the definition of what should be considered a remote rural area, we re-estimated the models using semi-parametric regressions that allow the shape of the relationship between population change and travel time to be drawn from the data without making restrictive assumptions on its functional form. More specifically, we used Robinson's (1988) double residual semiparametric estimator and Hardle and Mammen's (1993) test that compares nonparametric with parametric specifications of the relationship between rural population change and road accessibility.[2]

Figure 6 shows the nonparametric fit of the relationship between rural population growth and travel time to the cities of different sizes, for the pooled sample covering the periods 1991-2001 and 2001-2011. The vertical axis in each graph shows the value of the nonparametric fit of rural population growth and the horizontal axis shows travel time to cities in the initial period. The shaded area is the confidence interval determined by the two standard error lines above and below the estimate of the curve. The shape of the estimated curves does not reveal significant nonlinear effects, and indeed we cannot reject the null hypothesis in the Hardle and Mammen's test (1993) that the nonparametric fit can be approximated by a parametric linear fit.

[2] The models were implemented using Stata's *semipar* command (Verardi and Debarsy (2012).

Figure 6. Nonparametric fit of the relationships between population change (vertical axis) and road travel time to urban hierarchy (horizontal axis)

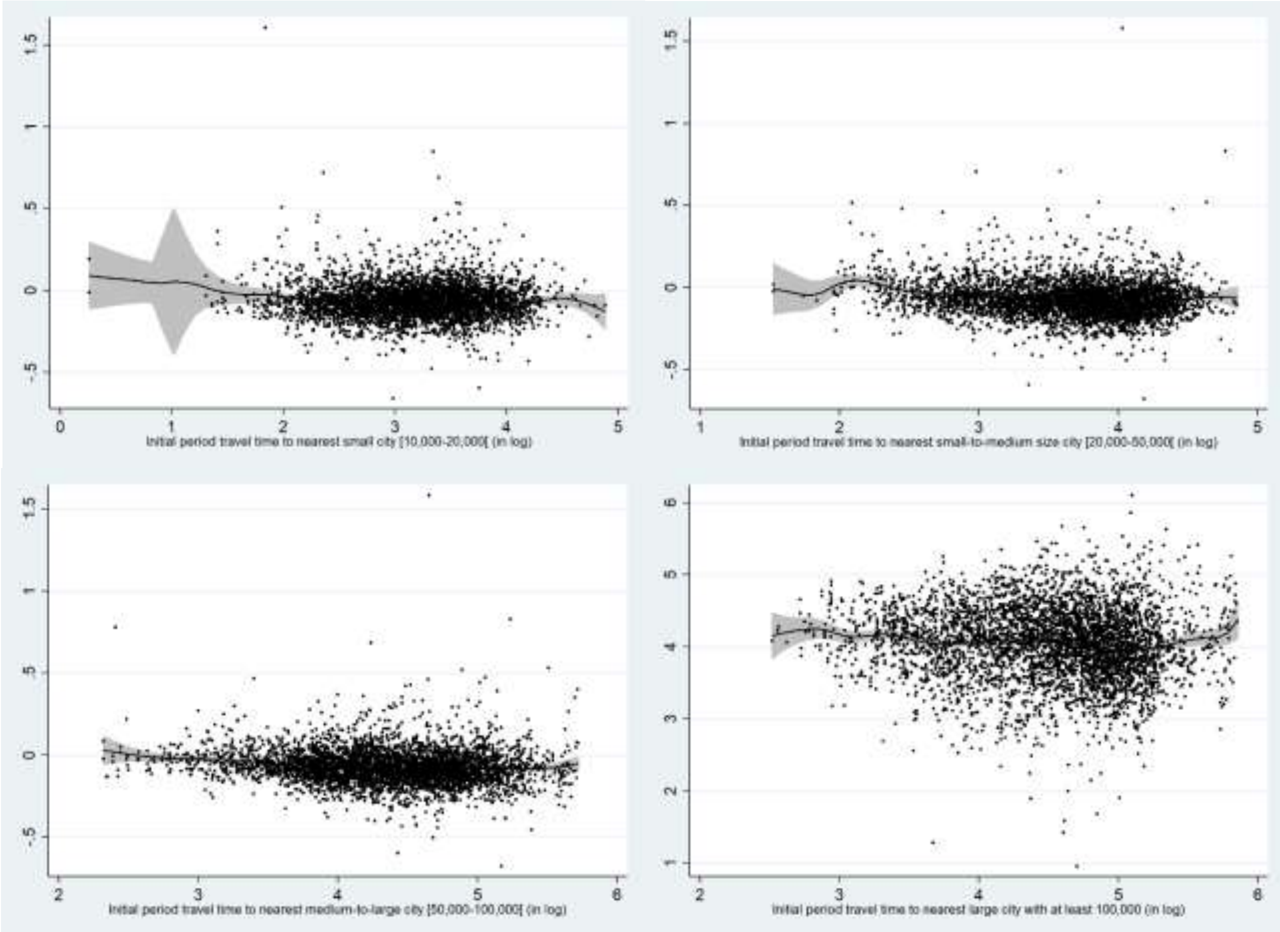


Table 5. Regressions of rural population change and road access to urban hierarchy (including interactions)

	Accessible vs. Remote (cut-off=30 minutes)		Accessible vs. Remote (cut-off=45 minutes)		Accessible vs. Remote (cut-off=60 minutes)	
	Using road distance	Using travel time	Using road distance	Using travel time	Using road distance	Using travel time
Initial period population density (in log)	0.0109**	0.0116**	0.0111**	0.0115**	0.0114**	0.0117**
Initial period road access to nearest small city (in log)	-0.0073	-0.0083	-0.0067	-0.0077	-0.0054	-0.0060
Accessible * initial period road access to nearest small city	0.0350	0.0275	0.0049	0.0064	-0.0043	-0.0070
Initial period road access to nearest small-to-medium city (in log)	-0.0318***	-0.0336***	-0.0327***	-0.0355***	-0.0353***	-0.0381***
Accessible * initial period road access to nearest small-to-medium city	0.0273	0.0259	0.0194**	0.0277***	0.0168**	0.0211***
Initial period road access to nearest medium-to-large city (in log)	-0.0309***	-0.0319***	-0.0299***	-0.0308***	-0.0309***	-0.0332***
Accessible * initial period road access to nearest medium-to-large city	-0.0268	-0.0895	-0.0109	-0.0136	-0.0105	-0.0078
Initial period road access to nearest large city (in log)	-0.0085	0.0090	-0.0066	0.0088	0.0029	0.0200**
Accessible * initial period road access to nearest large city	0.0298	0.0576	0.0072	0.0200	-0.0127	-0.0131
Initial period percentage of population with higher education (in log)	0.0039	0.0002	0.0038	0.0003	0.0039	0.0006
Initial period percentage of population aged 65 plus (in log)	-0.0884***	-0.0929***	-0.0882***	-0.0926***	-0.0883***	-0.0924***
Initial percentage of employed in tertiary sector (in log)	0.0136**	0.0139***	0.0135**	0.0141***	0.0137**	0.0145***
Initial period unemployment rate (in log)	-0.0029	-0.0034	-0.0028	-0.0033	-0.0026	-0.0031
Initial period number of active train stations	-0.0023	-0.0018	-0.0030	-0.0027	-0.0027	-0.0022
Percentage of area in European Natura 2000 Network	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Percentage of area in International Network for Wetlands	-0.0016***	-0.0016***	-0.0016***	-0.0016***	-0.0015***	-0.0015***
Shannon indicator of landscape diversity	0.0234***	0.0210***	0.0230***	0.0206***	0.0223***	0.0197***
Standard deviation of slopes	-0.0074***	-0.0069***	-0.0072***	-0.0067***	-0.0070***	-0.0064***
Accessible (vs. Remote) dummy	-0.1909	-0.0429	-0.0592	-0.1241	0.0543	0.0401
Constant	0.3890***	0.3424***	0.3700***	0.3332***	0.3312***	0.2910***
Controls for district	YES	YES	YES	YES	YES	YES
Observations	4156	4156	4156	4156	4156	4156
Adjusted R^2	0.2577	0.2576	0.2581	0.2587	0.2590	0.2593

Legend: * p<0.1; ** p<0.05; *** p<0.01. All models include controls for regions (Portuguese distritos).

5.4.Endogeneity checks

To address the concerns resulting from potential endogeneity bias between road location and population growth, we implemented the robustness analyses described in Section 3.2. The first analysis replaced road-distance and travel time with straight-line distances. The results are reported in Table 6 and are in line with those obtained for the baseline models using road distances and travel times (see Table 4), which suggests that endogeneity bias is not likely to affect our findings. In the second robustness analysis we re-estimated the baseline models for population change between 2001-2001 as a function of road distance and travel time in 1991 and the change in road distance and travel time between 1991 and 2001. The results (reported in Table 7) are similar to the original ones, suggesting again that endogeneity bias does not seem to affect our findings. In the case of the models using road travel times, rural population change between 2001-2011 is also positively affected by the change in road travel times to medium-size cities in the previous period (i.e. 1991-2001), besides the positive effect from proximity to medium-size cities in 1991. The differences may partially reflect the fact that travel times are a better proxy for real proximity than Euclidean road distances. The third robustness check consisted of estimating IV models using historical (i.e. density of Roman roads) and geographical (i.e. morphological slope variability) instruments. However, since we only have two instruments for four endogenous variables (i.e. road distance or travel time to each of the four groups of cities) in each model, we could only apply the IV approach separately for one endogenous variable at a time. In all cases, the model parameter estimates were not statistically significant and showed signs of weak instrument bias (i.e. the first stage regression had a very low goodness of fit). These results were not surprising given the poor correlation between the instruments and the endogenous variables. The pairwise correlation coefficients between the density of Roman roads and the proximity to the urban hierarchy range between -0.10 and 0.00 and

are not significant. The pairwise correlation coefficients between morphological slope variability and the proximity to the urban hierarchy vary between -0.15 and 0.24 and are generally not significant.[3]

Table 6. Rural population growth and proximity to the urban hierarchy based on straight-line distances

	Pooled	**1991-2001**	**2001-2011**
Straight line distance to nearest small city [10,000-20,000[(in log)	-0.0047	-0.0149*	0.0055
Straight line distance to nearest small-to-medium city [20,000-50,000[(in log)	-0.0308***	-0.0303***	-0.0305***
Straight line distance to nearest medium-to-large city [50,000-100,000[(in log)	-0.0293***	-0.0508***	-0.0072
Straight line distance to nearest large city with at least 100,000 people (in log)	-0.0136**	-0.0256**	-0.0038
Initial population density (in log)	0.0110**	0.0119	0.0087*
Initial percentage of population with higher education (in log)	0.0061*	0.0269***	0.0110*
Initial percentage of population aged 65 plus (in log)	-0.0874***	-0.0696***	-0.0985***
Initial percentage of employed in tertiary sector (in log)	0.0146***	0.0152**	0.0175*
Initial unemployment rate (in log)	-0.0033	-0.0082*	0.0018
Initial number of active train stations	-0.0016	-0.0092	0.0046
Percentage of area in European Natura 2000 Network	0.0000	0.0000	-0.0000
Percentage of area in International Network for Wetlands	-0.0017***	-0.0017**	-0.0015***
Shannon indicator of landscape diversity	0.0265***	0.0344***	0.0183**
Standard deviation of slopes	-0.0089***	-0.0090***	-0.0085***
Constant	0.3610***	0.4669***	0.2125***
Controls for district	YES	YES	YES
Observations	4156	2078	2078
Adjusted R^2	0.2564	0.2565	0.2763

Notes: * p<0.1; ** p<0.05; *** p<0.01. All models include controls for regions (Portuguese distritos).

[3] The results from the instrumental variables regressions estimated for each endogenous variable separately can be obtained from the authors upon request.

Table 7. Rural population growth in 2001-2011 and road access in 1991 and road access change in 1991-2001

	Using road distance		Using travel time	
	(1)	(2)	(3)	(4)
1991 distance to nearest small city (in log)	0.0028	0.0018		
1991-2001 change in distance to nearest small city		-0.0150		
1991 distance to nearest small-to-medium city (in log)	-0.0307***	-0.0318***		
1991-2001 change in distance to nearest small-to-medium city		0.0223		
1991 distance to nearest medium-to-large city (in log)	-0.0139**	-0.0133*		
1991-2001 change in distance to nearest medium-to-large city		0.0202		
1991 distance to nearest large city (in log)	-0.0017	-0.0022		
1991-2001 change in distance to nearest large city		0.0133		
1991 travel time to nearest small city (in log)			0.0023	0.0008
1991-2001 change in travel time to nearest small city				-0.0183
1991 travel time to nearest small-to-medium city (in log)			-0.0336***	-0.0360***
1991-2001 change in travel time to nearest small-to-medium city				0.0430***
1991 travel time to nearest medium-to-large city (in log)			-0.0143**	-0.0187**
1991-2001 change in travel time to nearest medium-to-large city				0.0368**
1991 travel time to nearest large city (in log)			-0.0008	0.0002
1991-2001 change in travel time to nearest large city				-0.0151
Initial population density (in log)	0.0077	0.0078	0.0071	0.0075
Initial percentage of population with higher education (in log)	0.0100*	0.0101*	0.0099*	0.0091
Initial percentage of population aged 65 plus (in log)	-0.0980***	-0.0960***	-0.0966***	-0.0952***
Initial percentage of employed in tertiary sector (in log)	0.0177*	0.0174*	0.0172*	0.0189*
Initial unemployment rate (in log)	0.0021	0.0023	0.0020	0.0020
Initial number of active train stations	0.0036	0.0032	0.0037	0.0047
Percentage of area in European Natura 2000 Network	-0.0000	0.0000	-0.0000	-0.0000
Percentage of area in International Network for Wetlands	-0.0016***	-0.0014***	-0.0016***	-0.0014***
Shannon indicator of landscape diversity	0.0173**	0.0173**	0.0162**	0.0160**
Standard deviation of slopes	-0.0077***	-0.0074***	-0.0075***	-0.0068***
Constant	0.2475***	0.2478***	0.2587***	0.2658***
Controls for district	Yes	Yes	Yes	Yes
Observations	2078	2078	2078	2078
Adjusted R^2	0.2772	0.2774	0.2779	0.2820

Notes: * p<0.1; ** p<0.05; *** p<0.01. All models include controls for regions (Portuguese distritos).

6. Conclusion

This study investigated the relationship between rural population growth and road access to the urban hierarchy in mainland Portugal. The large scale of public funding allocated to motorway investment since the country joined the European Union in 1986 makes Portugal an excellent case study to examine the role of improved road accessibility on the development of lagging areas, especially rural areas. By combining census data for 1991, 2001 and 2011 with GIS-based data for the road network and other relevant variables, we estimated regression models to investigate whether

population growth in rural areas is affected by road distances and travel times to cities of different size over the period.

The results show that road accessibility to the urban hierarchy positively influences rural population growth. The more interesting finding is that the positive effects of proximity are always statistically significant for medium-size cities (i.e. between 20,000 and 99,999 people), but not always for small and large cities. These results are valid for both measures of road accessibility (distance and travel times).

Our findings indicate that medium-size cities in Portugal play an important role in supporting population growth in their rural hinterlands. Furthermore, the results shed some light on the contradiction that road accessibility increased dramatically but population in rural areas declined in mainland Portugal since the 1980s. The models show that the decline in population is not explained by the increase in road accessibility. On the contrary, the increase in accessibility had a positive effect on population change. In other words, the increase in accessibility may have actually prevented population from declining even more. As shown in the models, population decline was mostly explained by demographic and socio-economic variables. Rural areas with low population density, high proportion of population aged 65+, low proportion of people with higher education, and low proportion of employment in the tertiary sector, showed a tendency to decline in their population. The decrease in road travel distance and travel time to cities attenuated that distance. Furthermore, this effect is linear, i.e. it particularly applies to rural areas that were already accessible to cities and to those that were remote. Both types of rural areas benefited by the increase in accessibility to cities allowed by road investment.

In both accessible and remote rural areas, it is possible that the mechanism through which accessibility contributed to rural population changes was the increase in the scope for commuting to cities. In the areas nearer to cities, this could be daily commuting, while in the areas farther away, it could be weekly commuting (i.e. people maintaining their residences in rural areas but working in

the cities during the week). This interpretation is consistent with the findings of (e.g. Lavesson, 2017), who found that rural areas can benefit from employment opportunities in nearby urban areas, mainly through local residential markets.

The results suggest that improving accessibility to cities may contribute to population growth in rural areas, or at least to mitigate population decline - achieving a positive change in population growth would require other policies to improve the local economy and stem the out-migration of younger population. Our results also indicate that supporting the population and economy of cities, especially medium-size cities, may indirectly support population growth in the rural areas that are more accessible to those cities. However, to achieve a balanced growth of both cities and rural areas, the interactions between the two types of areas need to be managed at a spatial level that is wider than the current administrative boundaries of city and rural regions in most Western countries, including Portugal (OECD, 2020).

The methods used in the paper could be further developed in future research. We used indicators of accessibility based on road distances and travel times to the nearest cities of different population sizes. This assumes that rural-urban linkages exist only with the nearest cities. However, rural areas have a range of cities that they can access, all at different distances and offering different opportunities. Indicators of transport accessibility could therefore include information about all cities weighted by distance and by their size. This would have the double advantage of including the full range of cities available to each rural area and avoiding classifying cities into a small number of classes (as the choice of cut-off points can influence results). Another possible improvement could be considering different definitions of the opportunities available in cities. In this paper, population was treated as an indicator of these opportunities, but other indicators (e.g. employment, income) could be used.

Due to lack of suitable data, we could not construct an indicator for railway services and used instead a simpler indicator for the presence of active railway stations, which we found to be an

insignificant predictor of population change. A priority for future search would be to develop accurate indicators of rail accessibility, taking into account travel times and service levels, and investigate the relative importance of rail and road accessibility to provide a more balanced view on the role of different modes of transport on population change.

Acknowledgements

The authors acknowledge financial Support from FCT – Fundação para a Ciência e Tecnologia (Portugal), national funding through the research grants UIDB/05069/2020, and UIDB/04007/2020 and UIDB/05183/2020. This work also received financial support from FCT as part of the project PTDC/EGEECO/28805/2017 – Transport Infrastructure and Urban Spatial Structure: Economic, Social and Environmental Effects (TiTuSS). The work was carried out also as part of the European Union's Horizon 2020 project RELOCAL - Resituating the Local in Cohesion and Territorial Development (GA727097).

References

ANCIÃES, P. R. 2013. Rethinking the death of the railway in the Portuguese countryside. Transport and Mobility in Rural Areas in the Context of Demographic Change. Band 14, 45-57. Erfurt, Germany: Institut Verkehr und Raum.

ANCIÃES, P. R. 2016. Population decline and accessibility in the Portuguese interior. *In:* GEURS, K. T., PATUELLI, R. & DENTINHO, T. (eds.) *Accessibility, Equity and Efficiency: Challenges for Transport and Public Services.* Cheltenham: Edward Elgar.

BARKLEY, D. L., HENRY, M. S. & BAO, S. 1996. Identifying "Spread" versus "Backwash" Effects in Regional Economic Areas: A Density Functions Approach. *Land Economics,* 72**,** 336-357.

BENSON, M. & O'REILLY, K. 2009. Migration and the Search for a Better Way of Life: A Critical Exploration of Lifestyle Migration. *Sociological Review,* 57**,** 608–25.

BERDEGUÉ, J. A. & SOLOAGA, I. 2018. Small and medium cities and development of Mexican rural areas. *World Development,* 107**,** 277-288.

BREZZI, M., DIJKSTRA, L. & RUIZ, V. 2011. OECD Extended Regional Typology: The Economic Performance of Remote Rural Regions. *OECD Regional Development Working Papers, No. 2011/06.* OECD Publishing.

CHRISTALLER, W. 1933. *Die zentralen Orte in Süddeutschland. ,* Jena, Gustav Fischer.

DGT 2011. Carta Administrativa Oficial de Portugal. Direcao-Geral do Territorio.

DIJKSTRA, L. & POELMAN, H. 2008. Remote Rural Regions: How proximity to a city infl uences the performance of rural regions. *Regional Focus* [Online]. Available: https://ec.europa.eu/regional_policy/sources/docgener/focus/2008_01_rural.pdf.

EP 2005-2011. Tráfego – Rede Rodoviária Nacional. Estradas de Portugal.

FONTES, M. J., RIBEIRO, A. & SILVA, J. 2014. Accessibility and Local Development: Interaction between Cross-border Accessibility and Local Development in Portugal and Spain. *Procedia - Social and Behavioral Sciences,* 111**,** 927-936.

GANNING, J. P., BAYLIS, K. & LEE, B. 2013. Spread and backwash effects for nonmetropolitan communities in the US. *Journal of Regional Science,* 53**,** 464-480.

GARCIA-LÓPEZ, M., HOLL, A. & VILADECANS-MARSAL, E. 2015. Suburbanization and highways in Spain when the Romans and the Bourbons still shape its cities. *Journal of Urban Economics,* 85**,** 52-67.

HARDLE, W. & MAMMEN, E. 1993. Comparing nonparametric versus parametric regression fits. *Annals of Statistics,* 21**,** 1926-1947.

HENRY, M. S., BARKLEY, D. L. & BAO, S. 1997. The Hinterland's Stake in Metropolitan Growth: Evidence from Selected Southern Regions. *Journal of Regional Science,* 37**,** 479-501.

HENRY, M. S., SCHMITT, B. & PIGUET, V. 2001. Spatial Econometric Models for Simultaneous Systems: Application to Rural Community Growth in France. *International Regional Science Review,* 24**,** 171-193.

IEP 1990-2004. Tráfego – Rede Nacional do Continente. Instituto de Estradas de Portugal (Portuguese Roads Institute).

IGEOE 1999. Carta militar itinerária de Portugal, 1/500000. Edição 1999. (Topographic map of Portugal at the 1:500000 scale, 1999 edition). Instituto Geográfico do Exército) (Portuguese Army Geographic Institute.

INE (INSTITUDO NACIONAL DE ESTATISTICA) 2014. Cidades Portuguesas: Um Retrato Estatístico. Lisboa: INE.

IRWIN, E. G., BELL, K. P., BOCKSTAEL, N. E., NEWBURN, D. A., PARTRIDGE, M. D. & WU, J. 2009. The Economics of Urban-Rural Space. *Annual Review of Resource Economics,* 1**,** 435-459.

LAVESSON, N. 2017. When and how does commuting to cities influence rural employment growth? *Journal of Regional Science,* 57**,** 631-654.

OECD 2020. Decentralisation and Regionalisation in Portugal What Reform Scenarios? OECD Multi-level Governance Studies. Paris: OECD Publishing.

PAEZ, A., SCOTT, D. M. & MORENCY, C. 2012. Measuring accessibility: positive and normative implementations of various accessibility indicators. *Journal of Transport Geography,* 25**,** 141-153.

PARTRIDGE, M., BOLLMAN, R. D., OLFERT, M. R. & ALASIA, A. 2007. Riding the Wave of Urban Growth in the Countryside: Spread, Backwash, or Stagnation? *Land Economics,* 83**,** 128-152.

PARTRIDGE, M. D., ALI, K. & OLFERT, M. R. 2010. Rural-to-Urban Commuting: Three Degrees of Integration. *Growth and Change,* 41**,** 303-335.

PARTRIDGE, M. D., RICKMAN, D. S., ALI, K. & OLFERT, M. R. 2008. Lost in Space: Population Dynamics in the American Hinterlands and Small Cities. *Journal of Economic Geography,* 8**,** 727-757.

PERCOCO, M. 2016. Highways, local economic structure and urban development. *Journal of Economic Geography* 16**,** 1035-1054.

PEREIRA, A. M. & PEREIRA, R. M. 2017. Is All Infrastructure Investment Created Equal? The Case of Portugal. *GEE Papers, Número 75, Agosto de 2017*

RIBEIRO, A. & SILVA, J. 2011a. Space, development and accessibility between Portugal and Spain: the last frontier. *Revista Portuguesa de Estudos Regionais,* 27.

RIBEIRO, A. & SILVA, J. 2011b. A Spatial analysis on the relation between accessibility and spatial development for Cross-border regions. *Investigaciones Regionales,* 21**,** 141-156.

ROBINSON, P. M. 1988. Root-n-consistent semiparametric regression. *Econometrica* 56**,** 931-954.

ROCHA, B. T., AFONSO, N., MELO, P. C. & DE ABREU E SILVA, J. 2020. What drives the allocation of motorways? Evidence from Portugal's fast-expanding network. *(manuscript submitted for publication)*.

SANTOS, B., SILVA, J. & GONÇALVES, J. 2013. O impacte de acessibilidade no desenvolvimento regional sustentável: O caso da região da Beira Interior. *7º Congresso Rodoviário Português-" Novos Desafios para a Atividade Rodoviária.*

SCHMITT, B., HENRY, M., PIGUET, V. & HILAL, M. 2006. Urban growth effects on rural population, export and service employment: evidence from eastern France. *The Annals of Regional Science,* 40**,** 779-801.

SOUSA, J. F., FERNANDES, A., GALIAN, S., ESTÊVÃO, M. & ANTUNES, G. 2011. A evolução dos transportes e acessibilidades e as transformações na organização do território. . *In:* ESPAÇO, I. D. D. D. (ed.). Universidade NOVA de Lisboa, Faculdade de Ciências Sociais e Humanas (FCSH).

SPELLERBERG, I. F. & FEDOR, P. J. 2003. A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the 'Shannon–Wiener' Index. *Global Ecology and Biogeography,* 12**,** 177-179.

TEIXEIRA, A. C. 2006. Transport policies in light of the new economic geography: The Portuguese experience. *Regional Science and Urban Economics,* 36**,** 450–466.

VENERI, P. & RUIZ, V. 2016. Urban-To-Rural Population Growth Linkages: Evidence from OECD Tl3 Regions. *Journal of Regional Science,* 56**,** 3-24.

VERARDI, V. & DEBARSY, N. 2012. Robinson's square root of N consistent semiparametric regression estimator in Stata. *The Stata Journal,* 12**,** 726-735.

Appendix A Urban-rural classification of freguesias

Source: Authors, based on TIPAU classification obtained from Portuguese National Statistics Institute (INE).

Appendix B Data sources

Variable description	**Source**
Population density	Population Census (1991, 2001, 2011), obtained from the National Statistics Institute INE
Population aged 65 plus (%)	
Population with higher education (%)	
Employment in tertiary sector (%)	
Unemployment rate (%)	
Number of train stations with operating services	Rail timetables 1991, 2001, 2011, Portuguese Railways: CP (Comboios de Portugal)(Portuguese Railways) Guia Horário Official Inverno 1990/91, Verão91, Inverno 2000/01, Verão2001. (Official Timetables, Winter 1990/91, Summer 1991, Winter 2000/2001, Summer 2001); CP (Comboios de Portugal) (Portugeuse Railways) Horários, www.cp.pt. 2)Location of stations (GIS data): REFER (2011) (Rede Ferroviária Nacional) (National Railway Infrastructure) Localização de Estações.(Location of Stations)
Road distance to urban hierarchy, kms Road travel time to urban hierarchy, minutes	1) GIS road model integrating information from: • Topographic map, Portuguese Army Geographic Institute: IGeoE (Instituto Geográfico do Exército) (Portuguese Army Geographic Institute). (1999) Carta militar itinerária de Portugal, 1/500000. Edição 1999. (Topographic map of Portugal at the 1:500000 scale, 1999 edition) • Maps included in the yearly Traffic Counting reports, National Road Institute: IEP (Instituto de Estradas de Portugal) (Portuguese Roads Institute) (1990-2004) Tráfego – Rede Nacional do Continente (Traffic – National Road Network, Mainland Portugal). IEP, Almada. Estradas de Portugal (Portuguese Roads Institute) (2005-2011) Tráfego – Rede Rodoviária Nacional (Traffic – National Road Network). Estradas de Portugal, Almada • Road classification included in the National Road Plan, Portuguese Government • Commercial road maps

Variable description	Source
	2) List of cities: Publication by National Statistics Institute: INE (Instituto Nacional de Estatística) (2014) Cidades Portuguesas: Um Retrato Estatístico. Lisboa: INE 3) Location of freguesias and cities: Official Administrative map, Portuguese Geographic Institute: DGT (Direcao-Geral do Territorio [Directorate-General for Territorial Development] (2011) Carta Administrativa Oficial de Portugal [Official Administrative Map of Portugal]
Sites of the European Natura 2000 Network (% area)	European Environment Agency, https://www.eea.europa.eu/data-and-maps/data/natura-11
Sites of the International Network for Wetlands (% area)	RAMSAR international network of wetlands, https://www.ramsar.org/
Shannon indicator of landscape diversity	Shannon-Wiener index of biological richness and heterogeneity (Spellerberg and Fedor, 2003)
Standard deviation of slopes	Own authors calculation based on the slopes of a 30 metre resolution grid for continental Portugal

Appendix C Rural population change and travel time to urban hierarchy