European Commission

JRC TECHNICAL REPORTS

The Determinants of Population Growth

Literature review and empirical analysis

Alvarez-Diaz, Marcos

D'Hombres Beatrice

Ghisetti, Claudia

Pontarollo, Nicola

Dijkstra, Lewis

2018

JRC Working Papers in Economics and Finance, 2018/10

Contact information
Name: Beatrice D'Hombres
Address: Via E. Fermi 2749, TP 361, Ispra (VA), I-21027, Italy
Email: Beatrice.DHOMBRES@ec.europa.eu
Tel.: +39 033278-3537

JRC Science Hub
https://ec.europa.eu/jrc

JRC112628

PDF ISBN 978-92-79-93404-9 ISSN 2467-2203 doi: 10.2760/513062

Luxembourg: Publications Office of the European Union, 2018

How to cite this report: Alvarez Diaz, M., D`hombres, B., Ghisetti, C., Pontarollo, N. and Dijkstra, L., *The Determinants of Population Growth*, Publications Office of the European Union, Luxembourg, 2018, ISBN 978- 92-79-93404-9, doi:10.2760/513062, JRC112628

Acknowledgements

The authors greatly acknowledge Michaela Saisana and the Competence Centre on Composite Indicators and Scoreboards (CC-COIN) of the Joint Research Centre for their useful comments and suggestions. We also thank LUISA team of the Joint Research Centre, in particular Carlo Lavalle, Filipe Batista and Chris Jacobs-Crisioni, for sharing data that are used in the empirical analysis. Marina Ferent-Pipas is also acknowledged for her assistance at the early stage of this project.

Authors

Marcos Alvarez-Diaz

Beatrice D'Hombres

Claudia Ghisetti

Nicola Pontarollo

Lewis Dijkstra

Abstract

This report studies population dynamics in Europe. Its purpose is threefold. First, the report offers a literature review of the main drivers of population growth. Second, an empirical analysis is carried out in order to unveil the determinants of population growth in EU sub-regions (NUTS3 level) over the period 2000-2010. Spatial econometrics is employed to account for spatial dependence among neighbouring regions. Third, the existing evidence on the long-run relationship between economic and population growth is discussed, followed by an empirical assessment of the relationship between these two aggregates in Europe over the period 1960-2010. Time-series econometric tools are used for this analysis. The main findings of both the litterature reviews and empirical analyses are discussed, along with their implications and future extensions.

Keywords: population dynamics; population growth; spatial econometrics; time-series econometrics; spatial dependence; regional development

Contents

List of tables

List of figures

1 Introduction

The population in the 28 European countries (EU28) was estimated at 511.8 million inhabitants in 2017, rising by 1.5 million people with respect to the previous year and by more than 105 million people in comparison to the EU28 figures in 1960 (Eurostat: Demographic Statistics). This overall growing trend hides substantial variations across EU Member States: whereas 18 countries experienced an increase of their population over the past year, the opposite is found for the remaining 10 EU countries. The largest population increases were recorded in Luxembourg and Sweden whilst the most important decreases were observed in Lithuania and Latvia.

When enlarging the focus to account for population gains and losses occurred at sub-national levels, the picture gets even more uneven. As a result of a natural reduction in population or in- and out- migration flows about one quarter of the regions in Europe defined at NUTS2 level have seen a decline of the size of the working age population between 2000 and 2010 (Rees et al. 2012). Regions in Western and Southern Europe have generally registered an increase of their population, whereas Eastern European regions have often experienced a population decline.

Population dynamics are intrinsically linked to patterns of economic and social convergence as well as, more broadly, to territorial cohesion across EU regions (Iammarino et al. 2017). EU regional policy is an essential component of the European Union (EC 2017a; 2017d), with more than one third of the EU budget dedicated to regional transfers to foster the competitiveness of lagging regions and reduce regional disparities in Europe. Where to allocate of EU investments in order to ensure the highest effectiveness of the cohesion policy is still open to debate (EC, 2017b; 2017c).In this context, shedding more light on the main drivers of population dynamics at sub-national levels of analysis might be very helpful.

The Directorate General for Regional and Urban Policy – European Commission (DG REGIO) in collaboration with the Competence Centre on Composite Indicators and Scoreboard (CC-COIN) of European Commission Joint Research Centre (JRC), has decided to embark on a research project specifically dedicated at assessing the determinants of population growth in Europe.

This report, which is the first output of this work, is organized as follows.

As a first step, a literature review has been conducted in order to document what are the main factors affecting population dynamics and whether the influence of these factors depends on the level of analysis considered (namely national, regional or local). The search of the relevant literature was done through citation databases of research studies, such as Google Scholar and Scopus. In total, more than 80 manuscripts were selected, 65 of which were revised and summarized (24 at the national and 41 at sub-national level of analysis). We provide a summary for each revised paper (in Appendix A) and an executive summary table giving information, for each paper, on the main focus, variables and methodological choices (in Appendix B).

The first part of Section 2 is dedicated to the literature review. The unit of analysis, country coverage and methodological approaches of the existing studies are discussed and the keys drivers of population growth are reviewed. The determinants of population growth are divided into 5 main dimensions, namely the **demographic characteristics** and **socio-economic conditions** of the territories under investigation, as well as their **transport accessibility**, **natural environment** and **land use policy**. The literature review also allows us to identify potential existing research gaps. The second part of section 2 offers an original empirical analysis on the population growth determinants in EU28 regions at NUTS 3 level. A spatial econometric model is employed over the period 2000-2010. Preliminary results support the main findings reported in the literature. In particular, the economic conditions play a key role in shaping population growth. Not only do the economic characteristics of the own regions matter, the economic features of neighbouring regions are also important drivers of population dynamics. This analysis is still preliminary and will be extended so as to unveil the determinants of population dynamics on a longer time period.

Section 3 is dedicated to the long term association between population and economic growth. Understanding the relationship between these two dimensions has been the subject of a long debate over the previous decades. The discussion is theoretically and empirically focused on exploring whether population affects and/or is affected by economic development. As for section 2, the first part of section 3 reviews the existing literature whilst the second part examines whether there is a long run relationship between population and economic development for the EU28 Member States over the period 1960-2010. Results, based on time series econometrics, support the existence of a bidirectional and positive relationship between population and GDP *per* capita.

Section 4 presents the main findings and lessons learnt from the literature review and empirical analysis carried out in Sections 2 and 3. Both empirical exercises are still preliminary and need to be further developed as discussed in the concluding part of the report.

2 Determinants of population growth: Literature review and preliminary empirical analysis

The objectives of this section are twofold. The first one is to offer a literature review on the determinants of population growth which covers the most important and most recent published articles. This review of the literature will help answering the following questions**:**

- What are the countries covered by the existing evidence?
- What is the unit of analysis commonly used to study population dynamic patterns?
- What are the main econometric or statistical methods used in the literature?
- What are the main determinants of population growth?

Addressing these questions allows us to paint a portrait of the current state of the empirical research on population growth, as well as to identify potential knowledge gaps. The second objective of this section is to carry out an empirical analysis to examine the determinants of population growth in EU28 regions over the period 2000-2010.

2.1 Literature review

2.1.1 Which countries are covered in the literature?

Most of the research on the determinants of population growth has, as main object of study, cities or counties in the United States (Glaeser et al., 1995; Clark and Murphy, 1996; Cullen and Levitt, 1999; Beeson et al., 2001; Glaeser et al. 2001; Huang et al. 2002; Glaeser and Saiz, 2004; Glaeser and Shapiro, 2003; Glaeser and Kohlhase, 2004; Boarnet et al., 2005; Burchifield et al., 2006; Rappaport, 2007; Partridge et al., 2008; Ellen and O'Regan, 2010; Chi and Ventura, 2011; Baum-Snow, 2007; Duranton and Turner, 2012; Chi and Marcouillier, 2013; González-Val, 2015; González-Val and Lanaspa, 2016; Lewis and Stanley, 2016; Rickman and Wang, 2017). According to González-Val and Lanaspa (2016) the United States is an interesting object of study because, in this country, cities' formation is a relatively recent phenomenon and the inhabitants are, on average, much more mobile than others; e.g. than European citizens. The United States have also experienced the largest population growth of its history over the period 1990-2000 (Rickman and Wang, 2017), naturally raising interest in the academic community.

In spite of the predominance of the studies on the United States, few papers examine population growth patterns in other countries such as France (Pirotte and Madre, 2011), Ireland (Lutz, 2001), Spain (García-López et al., 2015; Matori et al., 2014), Portugal (Barreira et al, 2017), Brazil (da Mata, 2007; da Silva et al., 2017), Korea (Lee et al., 2007; Sohn, 2012), Australia (Trendle, 2009), Japan (Fukuda, 2012), Colombia (Duranton, 2016) and Mexico (Garza-Rodriguez et al., 2016).

A small number of papers have a cross-country perspective. This is the case of Lehmijoki and Palokangas (2006), Bloom et al. (2009) and Huang and Xie (2013) with more than 50 countries included in their empirical analysis, or Cheshire and Magrini (2006), and Veneri and Ruiz (2016) which cover a set of OECD and European countries.[1]

Lesson learnt – *Most of the existing evidence is based on US data. Evidence on Europe with a cross-country perspective is limited.*

2.1.2 What is the unit of analysis commonly used in the literature?

Research on urban population growth relies on different spatial units. The majority of the studies employ data at **city**, **county** or **metropolitan** levels. Studies based on city data are numerous (e.g Barreira et al., 2017; Ellen and O'Regan, 2010; Fukuda, 2012; Lee et al., 2007; da Mata el al., 2007; Glaeser et al., 1995; González-Val, 2015; González-Val and Lanaspa, 2016; Ouesli et al., 2015). Cities present the advantage of being the smallest level of aggregation for which there is relatively abundant data for studying population growth patterns. However, according to Beeson et al. (2001), there are three main issues when working with city level data:

First, the population growth registered by cities might not be a proper indicator of the development and flourishing of urban areas. There is an increasing mismatch between the place where one lives and the place where one works. In other words, with the proliferation of suburban areas, many cities have experienced increasing levels of employment, even though the population has declined over the same period. Therefore, working at city level could not properly reflect urban population dynamics as the influence of cities goes beyond their administrative borders.

Second, there is a temporal bias in most of the studies. Poor or unavailable historical data led researchers to focus on a period beginning, at most, in the 1950s (i.e., postwar period). Focusing on one specific time period does not allow to study long-term trends. Exceptions to this temporal bias are the recent studies of Duranton (2016) for Colombian cities and González-Val and Lanaspa (2014) for US cities.

[1] Oueslati et al. (2015), Jacobs-Crisioni and Koomen (2017) also adopt a cross-country perspective. However their focus is on population density growth, rather than population growth.

Third, there is a selection bias intrinsic to the choice of cities for the empirical analysis. Most of the existing studies include the cities which were the largest in the latest period, i.e. cities that were particularly successful in their process of transformation and registered high population growth rates over time. In order to avoid this bias, González-Val and Lanaspa (2014) include in their analysis the biggest cities in 1860 or 1900. In this way, the selection should be minimized as the sample covers "winning" but also "losing" cities.

Given all these problems, many studies have considered the county level as unit of analysis (Rappaport, 2007; Beeson et al., 2001; Glaeser et al., 2001; Glaeser and Kohlhase, 2004; Partridge et al., 2008; Huang et al. 2002; Chi and Ventura, 2011). However, the use of county level data entails the problems that the geographic boundaries of some counties have changed over time, and that the population dynamics of counties within the same metropolitan area might be correlated (Beeson et al., 2001). In fact, metropolitan areas have been another geographical unit commonly used to analyze urban growth (Duranton and Turner, 2012; Burchfield et al., 2006; Pirotte and Madre, 2011; García-López et al., 2015; Martori et al., 2016). Some comprehensive studies complement the analysis of city level population with some empirical evidence at the metropolitan level (Glaeser et al., 1995; Glaeser and Saiz, 2004; Glaeser and Shapiro, 2003; Cullen and Levitt, 1999; Duranton, 2016).

Other geographical units have also been used such as the functional urban areas in Europe (Cheshire and Magrini, 2006), minor civil division level (Chi and Marcoullier, 2013) and census-defined places (Boarnet et al., 2005) in the United States or the minimum comparable areas in Brasil (da Silva et al., 2017). Recently, Veneri and Ruiz (2016) examine, in OECD countries, how population growth at NUT3 level is affected by distance to near urban centers.

Finally Lutz and Qiang (2002) adopt a cross-country perspective and examine the determinants of population growth for a set of 187 countries. Most of the literature, with the country as unit of analysis, exclusively focuses on the relationship between population growth and economic growth over a long time period. This literature includes studies from Bloom et al. (2009), Furuoka (2009, 2013), Chang et al. (2017), Pegou-Sibe et al. (2016), Huang and Xie (2013), Jung and Quddus (1986) which will be discussed in more details in section 3 of the report.

Lessons learnt – *Most of the reviewed literature focuses on cities, counties or metropolitan areas. To the best of our knowledge the sole study having a NUTS3 (or equivalent) approach and a broad country coverage (OECD countries) is Veneri and Ruiz (2016).*

2.1.3 What are the main econometric or statistical methods used in the literature?

Researchers have adopted different modeling techniques to explore what are the main determinants of population growth. The linear regression model estimated by the method of ordinary least squares (OLS) is by far the most common approach (Lee et al., 2007; Beeson et al., 2002; Rappaport, 2007; Glaeser et al., 1995; Glaeser et al., 2003).

Panel econometric techniques have also been employed to study the drivers of population growth. Barreira et al. (2017), who examine Portuguese city dynamics over the period 1991-2011 rely on cities random effects, whilst Ellen and O'Regan (2010) include cities fixed effects to analyze the population of US cities between 1980 and 2000.

In order to tackle endogeneity issues, arising from reverse causality or/and omitted variables (which plague most of the estimations of the determinants of population growth) some authors have also relied on instrumental variables estimation methods. Duranton and Turner (2012) are interested in the effect of transport infrastructures on the population growth of cities between 1983 and 2003 and instrument the variable related to highway density with measures of interstate highways and railroads in 1947 and 1998. García-López et al. (2015) examine the effect of highway on the population growth of Spanish metropolitan areas over the period 1960-2011 and use a similar approach (including similar exclusion restrictions) as Duranton and Turner (2012). Ellen and O'Regan (2010), who are specifically interested in the effect of criminality on the population growth of US cities, instrument the indicator of criminality with information on the severity of the criminal justice system. Furthermore, instrumental variable approaches are also adopted in Duranton (2016), Cullen and Levitt (1999); da Mata et al. (2007), Sohn (2012) or Huang et al. (2002).

The current literature on population growth has underlined the importance of location and spatial dependence; i.e., population dynamics in one area is dependent on population patterns in neighboring areas. The number of studies that account for spatial dependence has hence substantially increased in recent years. Lewis and Stanley (2016), who examine the determinants of population variation across counties in South Carolina over the period 1998-2012, estimate a model including spatially lagged explanatory variables. The empirical results show that urban counties are affected by changes happening in urban and rural local counties. Chi and Ventura (2011) analyze the drivers of population changes at municipal level from 1970-2000 in Wisconsin (United States), using a model with spatially lagged endogenous effects. Cheshire and Magrini (2006) arrive to the same conclusion using data at the level of functional urban area for 12

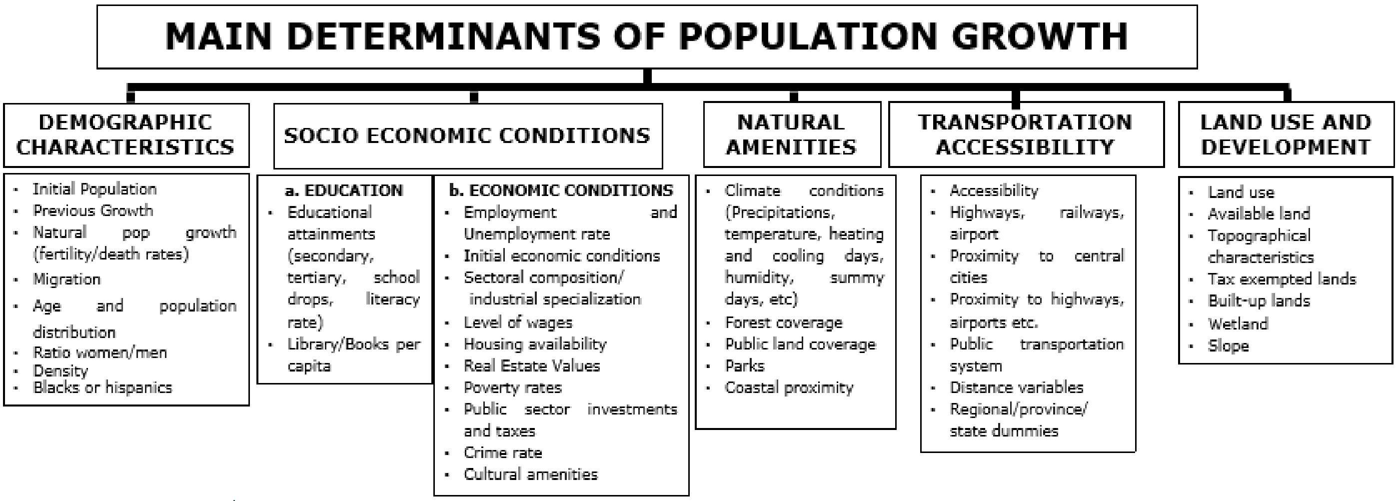
European countries. Da Silva et al. (2017) estimate several specifications which include spatial effects (Spatial Error Model, Spatial Autoregressive Model, Dynamic Spatial Durbin Models, etc.) to study the determinants of population growth in Brazilian minimum comparable areas over a 40 years period. The authors also check the robustness of their findings to alternative weighting matrices used to account for the spatial effects. Again, the results point to the importance of controlling for the influence of neighboring municipalities and, more broadly, for spatial dependence. Fukuda (2012) exploits spatial autoregressive and spatial error models to investigate the factors underlying cities' growth in Japan. Furthermore, also Chi and Marcoullier (2013), González-Val (2015) and Boarnet et al. (2005) are based on US data and account for spatial dependence through spatial econometric models.

Lesson Learnt - *The empirical evidence suggests that it is important to control for spatial effects, at least when working at a disaggregated level. Causal interpretation to findings based on OLS or spatial models that cannot account for reverse causality or omitted variable bias need a cautious interpretation. A classical but imperfect way to mitigate for this is to include time lagged covariates. Alternatively, instrumental variable methods are better suited when focusing on the effect of one specific covariate, provided that there is the possibility to identify exogenous instruments.*

2.1.4 What are the main determinants of population growth?

Population changes and more broadly population dynamics have been investigated by a multitude of disciplines, including geography and transport economics, urban planning and demographic studies as well as regional economics. Chi and Ventura (2011a) propose an interdisciplinary approach that draws on former theories developed by separate disciplines. In their framework, the determinants of population dynamics revolve around five components, namely the demographic characteristics and socio-economic conditions of the areas under scrutiny as well as its transportation accessibility, natural environment, and land use and development. In addition to these influential factors, Chi and Ventura (2011a) emphasize the importance of accounting for the spatial and temporal effects when studying population dynamics. In this section we review the empirical evidence on the main determinants of population growth. Following Chi and Ventura (2011a; 2011b), we have regrouped the determinants in 5 main categories (see Figure 1).

Figure 1: Main determinants of population growth.



Source: Grouping based on Chi and Ventura (2011a; 2011b).

The first category includes the **demographic characteristics** of the geographical area under scrutiny such as fertility, mortality and migration trends, population density and population composition in terms, for instance, of age, sex, races etc.

The second category covers the **socio-economic conditions** of the geographical areas and includes education-related measures, proxies for the economic conditions (income *per* capita, employment opportunities, sectoral structure of the geographical area, wages) as well indicators linked to housing conditions and local taxes.

The third category, **transportation accessibility,** captures the accessibility of the area in terms of, for instance, proximity to main airports and highways, the presence or effectiveness of public transportation systems, traffic levels as well as measures of the amount of local expenditures dedicated to transportation.

The fourth category, **natural environment,** measures the presence of attractive natural amenities such as favourable climate conditions, forest coverage or proximity to coasts.

Finally, the fifth category, **land use and development,** includes elements such as geophysical characteristics, built-up lands, cultural and aesthetical resources, and legal constraints to population growth posed by land development specific regulation.

2.1.4.1 Demographics characteristics

- **Rationale**

Demography is a salient determinant of population growth. In purely accounting terms, the population growth experienced by an area over a period of time greatly depends on demographic characteristics such as the natural rate of population growth (i.e., the number of births minus death plus) and net migration flows. Additional demographic characteristics have also an impact on population (e.g., previous population, age structure or population density). It is not surprising therefore that much of the empirical literature on the determinants of population growth has included some demographic variables as covariates.

- **Variables used**

Many different demographic variables have been added in the econometric models considered in the empirical literature. Some authors, such as Duranton (2016) and Lewis and Stanley (2016) directly include the ***natural rate of population growth*** as explanatory variable. However, the most common and, quoting Duranton (2016), "crucial" demographic variable is the ***initial level of population*** (e.g., Lee et al., 2007; Partridge et al., 2007; da Mata et al., 2007; Beeson et al., 2001; Glaeser et al., 1995, Glaeser and Saiz, 2004; Glaeser and Shapiro, 2003; Boarnet et al. 2005; Trendle 2009; Duranton, 2016; Baum-Snow, 2007; González-Val and Lanaspa, 2016; Glaeser and Shapiro, 2003). Other popular demographic variables are the value of the ***previous population growth*** (Cheshire and Magrini, 2006; Glaeser et al., 1995; González-Val, 2015; Ellen and O'Reagan, 2010) and ***population density*** (Lutz and Qiang, 2002; Rappaport, 2007; Partridge et al., 2007; Pirotte and Madre, 2011; Glaeser and Shapiro, 2003; Martori et al., 2016; Chi and Ventura, 2011; da Silva et al., 2017). Less frequently used variables are indicators of **migration flows** (Duranton, 2016; Glaeser et al., 1995; Ellen and O'Regan, 2010) and measures of ***ageing population*** (Lutz and Qiang, 2002; Glaeser and Shapiro, 2003; Lee et al., 2007; Partridge et al., 2007; Ellen and O'Regan, 2010; Sohn, 2012; Chi and Ventura, 2011; Lewis and Stanley, 2016; Veneri and Ruiz, 2016).

- **Findings**

Starting with the findings regarding the effect of ***initial levels of population***, the results reported in the empirical literature are rather mixed and inconclusive. On the one hand, Beeson et al. (2001), Glaeser and Saiz (2004), Boarnet et al. (2005), Lee et al. (2007), Da Mata et al. (2007) and Trendle (2009) report a negative association between population growth and initial levels of population. This finding implies that the most populated areas tend to grow less. In other words, it means there is a "***mean reverse***" or "***catching up"*** *effect*. On the other hand, Glaeser et al. (1995) claim that there is limited evidence that bigger cities have a lower population growth. González-Val and Lanaspa (2016) show that US city growth rates are independent to initial levels of population. On the other extreme, Duranton (2016) and Baum-Snow (2007) report to have found a positive effect associated with initials levels of population. This finding would suggest a possible pattern of **divergence** in population growth.

The ***population growth registered in the previous period*** is also used as possible determinant of population growth. The empirical evidence seems to indicate the existence of a ***persistent effect***. According to Glaeser et al. (1995), Cheshire and Magrini (2006) and Ellen and O'Reagan (2010), areas that experienced a substantial population growth in the recent past are more likely to also register a significant population growth in the following period.

The level of ***population density*** is included in the empirical studies to account for **agglomeration** effects (positive influence) or **congestion** effects (negative influence). The findings reported in Glaeser and Shapiro (2003), Partridge et al. (2007), González-Val (2015) and da Silva et al. (2017) reveal that low density cities grow faster than high density cities. Cheshire and Magrini (2006), who examine population growth of Functional Urban Regions (FUR) in the EU-12, found a negative but insignificant association between population density and population growth. At a country level, Lutz and Qiang (2002) observe that population density is positively associated with population growth only for the years 1960 and 1965. Their analysis is conducted on a pooled sample of 187 countries over the period 1960 to 2000, exploiting five-year sub-periods. On the contrary, Lehmijoki and Palokangas (2006) find a negative association between population density and the population growth of 69 low- and mid-income countries. Regarding **migration**, Duranton (2016) concludes that **in-migrant flows** have a significantly positive effect on the population growth of Colombian cities, while the opposite is found for **out-migrant flows**. Glaeser et al. (1995) argue that US cities that attracted more migrants in the past grow faster. On the contrary, Ellen and O'Regan (2010) do not find any significant effect of foreign population on the US city growth rates.

Finally, the empirical literature indicates that **ageing population** is a deterrent factor of population growth, while having a young population is associated with larger population growth. Specifically, Glaeser and Shapiro (2003) and Ellen and O'Regan (2010) find some evidence that those US cities with an ageing population present lower rates of population growth. Veneri and Ruiz (2016) show that rural OECD small rural regions with elderly population grow less. Lee et al. (2007) come to the same conclusion for Korean cities, while Sohn (2012) concludes that the percentage of population under-19 has a positive influence on the growth rates of small and medium cities in Korea.

2.1.4.2 Socio-economic conditions

Education

- **Rationale**

Several papers consider that human capital is among the strongest determinants of population growth (Glaeser et al., 1995; Glaeser et al., 2001; Glaeser and Shapiro, 2003; Cullen and Levitt, 1999, Glaeser and Saiz, 2004). The assumption is that geographical areas with a large concentration of highly educated people are expected to be more productive and hence economically more attractive. The concentration of skilled workers also creates positive externalities and fosters the generation of new ideas. This translates into higher wages and additional incentives to move into areas with a highly educated population. Additionally, the geographical areas with a large share highly educated people might also be better places to live as they might provide cultural goods and more generally local amenities that are associated to education (e.g., relatively free of social problems and offer better schools) and are less exposed to social problems linked to poor economic conditions. However, in spite of all these possible explanations, there is no consensus on the causes or implications of the relationship between human capital and population growth (Glaeser and Saiz, 2004).

- **Variables used**

Different variables have been used to approximate human capital. The **percentage of population with tertiary education** is the most common proxy for human capital (Glaeser and Saiz, 2004; Partridge, 2007; Trendle, 2009; Sohn, 2012; Chi and Ventura, 2011; Sohn, 2012; Duranton and Turner, 2012; Fukuda, 2012; González-Val, 2015; Duranton, 2016). Other proxies include **the average or median number of years of education** (Da Mata et al., 2007; Lee et al., 2007, Glaeser et al., 1995; Huang et al., 2002), the **percentage of population with secondary education** (Cullen and Levitt, 1999; Sohn, 2012; Chi and Ventura, 2011), **literacy rates overall or by gender** (Beeson et al., 2001; da Silva et al., 2017; Lutz and Qiang, 2002).[2] Finally, human capital has also been measured by "supply side indicators", i.e. related to the supply of educational facilities. For instance, Beeson et al. (2001) measure **educational infrastructures** with the number of libraries, colleges or books *per* capita, and Sohn (2012) relies on the **number of elementary school students *per* class.** Finally Clark and Murphy (2006) employ measures of **government expenditures on education.**

- **Findings**

Most of the papers investigating the link between population growth and education report a positive association. For instance, Glaeser et al. (1995) conclude that education exerts a positive and significant influence on US urban population growth for the period 1960-1990. Cullen and Levitt (1999) also find evidence of the positive effect of education on the population growth of US cities, in the period 1976-1993. Glaeser and Saiz (2004) examine the main determinants of population growth using a panel of US metropolitan areas and cities in the period 1970-2000 and conclude, even after having controlled for a large set of possible determinants (climate variables, sectoral variables, etc.), that human capital increases population growth. Glaeser and Shapiro (2003), studying the population growth of US cities and metropolitan areas in the decades of the 80s and 90s, arrive to similar conclusions. Beeson et al. (2001) find that educational infrastructures exert a positive influence on the population growth of US counties over the period 1840-1990. González-Val (2015) also concludes in favour of a positive influence of human capital on US cities' economic growth from 1990 to 2000. The positive impact of

[2] Martori et al. (2016) also analyze the effect of human capital but adopting a reverse direction. The authors use the **percentage of population with no educational diploma** to approximate the effect of low education on the growth rate of non-EU in-migrants to Barcelona.

education is also documented in Trendle (2009), Sohn (2012), Fukuda (2012), Duranton (2016), da Silva et al. (2017) using data respectively from Australia, Korea, Japan, Colombia, and Brazil. On the contrary, Huang et al. (2002) find a negative relationship between some educational variables and the population growth in rural counties of the US between 1950 and 2000, which reveals a brain drain from rural to urban counties for the best educated people.

Another interesting paper on the role of human capital on population growth is the one by Winters (2011). The paper studies the importance of student migration in affecting the relationship between human capital and population growth for nonmetropolitan counties in the United States over the period 1995-2000. Many students who move to an area for higher education might decide to stay in the area after their schooling is complete because of networks with local employers, which have been established during the studies, friendships or a taste for the local amenities. The ability to retain former students causes both population growth and an increase in the average level of human capital. To test this assumption, the authors estimate in/out/net migration growth equations with one of the explanatory variables being the share of the adult population with a college degree in the county. Winters (2011) indeed shows that the share of adults with a college education is positively associated with in-migration of higher students, whilst such an effect cannot be observed for in-migration of non-students. This result is obtained after having controlled for the demographic and economic characteristics of the nonmetropolitan counties as well as for distance related variables and several variables measuring the natural amenities of the counties.

Economic characteristics

- **Rationale**

As underlined in Chi and Ventura (2011), the economic characteristics of a geographical area are important determinants of population changes. Economic development is a proxy for employment opportunities and income conditions. Individuals are likely to live in places with low unemployment rates and more broadly good economic prospects and potential markets. In general, it is widely recognized in the literature that areas that are economically more dynamics tend to show higher rates of population growth (Glaeser et al., 1995; Duranton, 2016).

- **Variables used**

Various indicators have been used in the literature to proxy the economic dynamism of a geographical area. The level of **income *per* capita** is usually used as measure of the area's level of economic development (Glaeser et al., 1995; Clark and Murphy, 1996; Cullen and Levitt, 1999; Huang et al., 2002, Glaeser and Shapiro, 2003; Faria et al., 2006; Trendle, 2009; Baum-Snow 2007; Duranton and Turner, 2012; González-Val, 2015; Lewis and Stanley, 2016). Other authors, on the other hand, prefer including **unemployment rates** to proxy the economic attractiveness of the areas (Glaeser et al., 1995; Boarnet et al., 2005; Lee et al., 2007; Cheshire and Magrini, 2006; Fukuda, 2012; González-Val, 2015; Barreira et al., 2017). These two variables - income per capita and unemployment rates - are usually highly correlated.

Some authors have also shown the importance of accounting for the **sectoral structure of the area's economy**. Glaeser et al. (1995), Cullen and Levitt (1999), Lutz (2001) Beeson et al. (2001), Glaeser and Shapiro (2003), Boarnet et al. (2005), Duranton and Turner (2012), Lewis and Stanley (2016) or Barreira et al. (2017) include as covariates the percentage of employment in the manufacturing, agricultural and/or service sectors. Other related variables were also used, but to a lesser extent. For instance, González-Val (2015) and Duranton (2016) construct indices to reflect the **level of industrial specialization**. Duranton (2016) also points out that some important determinants such as the **level of wages** have been usually neglected by recent literature on population growth. The author includes the average city wage as one possible determinant of the population growth of the Colombian cities. **Housing availability** is taken into consideration in Cullen and Levitt (1999), Boarnet et al. (2005) and Barreira et al. (2017). The effect of the **intervention of the public sector**, through public investments and taxes, was addressed in Glaeser et al. (1995), Clark and Murphy (1996), Huang et al. (2002), Cheshire and Magrini (2006) and Fukuda (2012). These authors include different variables to approximate the intervention of the public sector with the underlying assumption that people move to cities with better and more efficient public policies.

- **Findings**

Regarding the use of **income *per* capita** as proxy of economic development and market potential the empirical evidence provides mixed results. Most studies, among others Glaeser and Shapiro (2003), Huang et al. (2002) and González-Val (2015), find the expected positive effect: the higher the levels of income, the higher the population growth. Conversely, a negative and significant effect was also reported in several studies. Faria et al. (2006), using a panel of 125 countries over the period 1950-2000, conclude in favour of a weak but negative relationship between population growth and *per* capita GDP: as income increases population expands at a slower rate. At the level of cities, Glaeser et. (1995) find a negative impact of initial levels of *per* capita income on the population growth of US cities in the period 1960-1990. Cullen and Levitt (1999) report a negative impact of median family income on the population growth for US cities and metropolitan areas employing data over the period 1976-1993. Baum-Snow (2007) observes that US city population decline was more rapid in decades that had faster income growth. Trendle (2009) also finds evidence of a negative influence of the income variable on the economic growth of Local Government Areas in Queensland (Australia) over the period 1996-2006. The author affirms that, at first sight, this might seem puzzling as higher income is generally associated with greater economic development. Lewis and Stanley (2016) also show their surprise by finding a negative influence of income on the population growth in the counties of South Carolina (US) in the period 1998-2012. Both Trendle (2009) and Lewis and Stanley (2016) argue that a possible explanation to this apparent contradictory finding is that the income variable is capturing the higher cost of living in high-income areas, disincentivizing in-migration (e.g., higher housing prices). Finally, Duranton and Turner (2012) exploit income as control variable to explore the impact of transport infrastructures on the population growth of US cities between 1983 and 2003 but the estimated coefficient is not reported.

All studies that include **unemployment rate** among the determinants of population growth arrive to the conclusion that this variable adversely affects population growth. Glaeser et al. (1995) report that high levels of initial unemployment reduce the subsequent population growth of large US cities over the years 1960-1990. Cullen and Levitt (1999) include the unemployment rate as control variable in their analysis of the effect of crime on population growth in the US cities and metropolitan areas over the period 1976-1993. They find some evidence of a negative effect in one of their model specifications. Ellen and O'Reagan (2010) report a negative effect of unemployment on the US city population growth in 1980-2000. González-Val (2015) also concludes that cities with initial high levels of unemployment in 1990 experienced lower rates of population growth in the period 1990-2000. Glaeser and Saiz (2004) find a negative but not significant effect of this variable on population growth for the US metropolitan areas and cities for the period 1970-2000.

The adverse effect of unemployment rate on population growth is not limited to cities and metropolitan areas of the U.S. The finding extends to other urban areas outside the US. For instance, Lee et al. (2007) show that unemployment rate had a negative effect on the population growth of Korean cities for the period 1980-2000, though the estimated coefficient was not statistically significant. Cheshire and Magrini (2006) reach a similar conclusion when studying the determinants of population growth of European Functional Urban Areas (FUA) over the period 1980-2000. Trendle (2009) concludes that the regional level of employment is one of the main drivers of the population growth of Local Government Areas in Queensland (Australia). Fukuda (2012) finds that initial levels of unemployment rate are negatively related to the population growth of the Japanese cities over the period 2000-2005. Barreira et al. (2017) suggest that having high levels of unemployment was a push factor for the population growth of the Portuguese cities in the period 1991-2011.

The **sectoral structure** of the local economy has a significant influence on population growth. Glaeser et al (1995) find that the population growth of large US cities over the period 1960-1990 was negatively and significantly related to the initial share of employment in manufacturing. Glaeser and Saiz (2004) report the same finding when examining the period 1970-2000. Glaeser and Kohlhase (2004) also reach similar conclusions when US counties are the level of analysis. Likewise, Chesire and Magrini (2007) report that the shares of employment in the manufacturing and agricultural sectors are negatively associated with the population growth of the EU12 Functional Urban Area. As explained in Glaeser and Kohlhase (2004), the cost of transport for goods has declined drastically whilst moving people remains costly. This implies that the labor force should be located in areas with a substantial share of employment in services and the opposite should be observed for areas with a large concentration of agricultural and industrial activities.

However, according to Beeson et al. (2001), there are often economies of scale and other positive externalities linked to industrial development (e.g., shared industry-specific services, technological spillovers within industries, comparative advantages associated with industrial diversity). This assumption is confirmed by several studies. For instance, Beeson et al. (2001) document that US counties with higher shares of employment in the manufacturing and commerce sectors in 1840 experienced faster population growth over the period 1840-1990. Da Mata et al. (2007) find that Brazilian cities with an initial high manufacturing ratio registered stronger growth over the period 1970-2000. Lee et al. (2007) conclude that the initial employment share of manufacturing industries positively affected population growth of Korean cities in the period 1980-2000, though the effect weakened over time. Jacobs-Crisioni and Koomen (2017) suggest that the weight of manufacturing sector relative to the service sector is positively associated with the population growth of European municipalities over the period 1961-2011. Barreira et al. (2017) affirm that the shares of employment in the manufacturing and service sectors acted as pull factors attracting population in the case of the Portuguese cities over the period 1991-2001. Da Silva et al. (2017) find that the share of employment in the manufacturing sector had a positive effect on the population growth of the Brazilian minimum comparable areas over the period 1970-2010; conversely, they find a negative effect of employment in agriculture due to the reduction of economic opportunities.

Worth being noted that a number of studies, such as Lutz (2001), Glaeser and Shapiro (2003), Glaeser and Saiz (2004) and Lewis and Stanley (2016), did not find any statistically significant association between population growth and the economic sectoral structure.[3] Other authors remark that patterns of specialization and diversity can be better proxies of the economic structure (Glaeser et al., 1992). Thus, some authors use different indices to reflect the effect of the **level of industrial specialization** on population growth. For instance, Duranton (2016) employ 2 indices of specialization and 2 indices of diversity as measures of the composition of the economic activity of the Colombian cities over the period 1993-2010. The specialization relative to the nation share in employment is the only index that shows a significant positive effect on population growth. González-Val (2015) defines an index of industrial diversity that approximates the diversity of the sectoral structure of the US cities in 1990, and uses it to explain their population growth in the period 1990-2000. He finds a positive significant impact of the index on the population growth of US cities.

A few other socio-economic determinants have been considered in the literature. Duranton (2016), for instance, reports a positive effect of average city wages on the population growth in Colombian cities. According to Barreira et al. (2017), housing availability could also be an important factor to account for. The authors show that the proportion of vacant middle-aged houses (10-30 years old) positively affects the population growth of the Portuguese cities in the period 1991-2011. On the contrary, Boarnet et al. (2005) do not find that percentage of housing stock ages is a significant determinant of the population growth in the Orange County (US) from 1980 to 1990. As suggested by Clark and Murphy (1996), local government spending and taxation could also explain population growth. However, this assumption is not confirmed by their empirical analysis carried out with US data at county level over the period 1981-1989. Glaeser et al. (1995) also examine the effect of the composition of government spending on the population growth of US cities and metropolitan areas between 1960 and 1990. Again, government spending does not seem to significantly influence population growth (except for the case of expenditure on sanitation). Huang et al. (2002), who study the effect of a set of local and state government expenditures (e.g. welfare, education, debt,

[3] Duranton and Turner (2012) use the share of manufacturing employment as control variable in their analysis of the impact of transport infrastructures on the growth of the US cities between 1983 and 2003, but they do not report the estimates of the effect.

highway and debt) on the population growth of rural US counties in the period 1950-1990, conclude that rural county government services financed by local taxes or debt have neutral or negative effects on population growth. Cheshire and Magrini (2006) show that the European integration process was a significant determinant that increased urban population growth in the EU-12. Fukuda (2012) finds that public indebtedness has negatively influenced the population growth of the Japanese cities over the period 2000-2005, whilst the fiscal strength of the city exerts a positive influence.

2.1.4.3 Natural Amenities

- **Rationale**

The presence or absence of natural amenities is a relevant determinant of population growth (or decline) at regional and/or other levels of analysis. The presence of natural amenities has long shaped the urban to rural population redistribution and "the improvement (or degradation) of natural amenities will attract migrants into (or drive residents out of) the region" (Chi and Ventura, 2011a: 557). Overall favourable natural amenities are expected to affect positively population growth.

- **Variables used**

Natural amenities have been measured through different proxies in the literature. Some authors, such as Duranton (2016), Duranton and Turner (2012), Fukuda (2012), Cheshire and Magrini (2006), Rapapport (2007), Beeson et al. (2001), Gonzalez-Val (2005), Partridge et al. (2007) or Winters (2011), mostly focus on **weather-climate conditions**, i.e., temperature, number of heating and cooling days/hours, number of sunny days/ hours, precipitations, humidity, altitude. Other studies consider **additional natural amenities** such as the presence or the proximity of a coast, the percentage of forest, water and/or wetland coverage, the percentage of public land coverage, the lengths of lakeshore/riverbank/coastline or view shed (e.g. Duranton, 2016; Cullen and Levitt, 1999; Chi and Ventura, 2011b; Chi and Marcouillier, 2013; Partridge et al., 2007; Winters, 2011; Glaeser et al., 2001; Glaeser and Saiz, 2004; Rickman and Wang, 2017). All in all, most of the studies combine climate/weather variables with additional natural amenities to have a broader measurement of the local environment.

- **Findings**

Most of the studies find that natural amenities are important drivers of location decision (e.g. Glaeser et al., 2011; García-Lopez et al., 2015; Duranton, 2016; Glaeser and Kohlhase, 2004; Clark and Murphy, 1996).

Certain studies also highlight the presence of spatial heterogeneities in the way natural amenities affect population growth (e.g. Chi and Marcouillier, 2013; Partridge et al., 2007; Rickman and Wang, 2017). In particular, Chi and Marcouillier (2013) test whether natural amenities influence vary with the characteristics of the region: urban, suburban, peri-urban or rural. The underlying assumption is that regional areas have specific characteristics that make the natural amenities more or less important for migration decisions. Urban areas offer other cultural opportunities that might be more important than natural amenities to drive location decisions. On the contrary, migrants to rural areas are likely to put more weights on natural amenities. Their results suggest that natural amenities have the largest importance on in-migration to rural areas adjacent to metro areas while no effect is observed for migration into urban areas. Similarly, Rickman and Wang (2017) found that, for metropolitan areas, the attractiveness of the city drives population growth more than economic productivity growth, leading the authors to conclude that "rather than agglomeration economies producing jobs that cause in-migration, it is the attractiveness of cities to households that appeared to spur population growth" (Rickman and Wang 2017: 89).

2.1.4.4 Transport- Accessibility

- **Rationale**

The more accessible is a geographical area, the higher is the probability that this area will experience further population flows. Transportation has played a central role in theoretical models of cities, but the empirical literature has often ignored urban transport as determinant of population growth. The reasons might be the lack of data and the difficulty of dealing with the simultaneous determination of population growth and transportation infrastructures in cities (Duranton and Turner, 2012).

- **Variables used**

Transportation accessibility should capture the accessibility of the area both for goods and people. Most of the papers have built or used *ad hoc* measures of accessibility (e.g. Chi and Ventura, 2011b; Duranton and Turner, 2012; Jacobs-Crisioni and Koomen, 2017; Beeson et al., 2001; Glaeser and Kohlhase, 2004). Other studies have captured accessibility mostly through the exploitation of regional/local dichotomous variables (e.g. Glaeser et al., 1995; Cullen and Levitt, 1999; González-Val, 2015).

Lastly, some papers have conceived accessibility in terms of inverted distance to core places, e.g. capital, airports, highways, closest urban centres (e.g. Chi and Ventura, 2011b; Lee et al., 2007; Da Mata et al., 2007; Lutz, 2001; Sohn, 2012; Winters, 2011).

A broad set of variables have been exploited in the literature to approximate *ad hoc* accessibility, including the presence or effectiveness of public transportation system, traffic levels, local capital expenditures on transportation, transportation costs, change in the transportation modes (from railways to cars, airplane), highway density, kilometres of highways and railroads, access to natural transportation networks such as oceans, rivers, lakes (e.g. Chi and Ventura, 2011b; Duranton and Turner, 2012; Beeson et al., 2001, Glaeser and Kohlhase, 2004). Most of these variables have been coupled with inverted distance variables to approximate for the distance to core places, be it the capital, the closest urban centre, the closest highway, the most important airport or train station (e.g. Chi and Ventura, 2011).

- **Findings**

The positive effect of transport accessibility on population growth has been several times documented in the literature (e.g. Duranton and Turner, 2012; Beeson et al., 2001; da Mata et al., 2007; Duranton, 2016).

Duranton and Turner (2012), whilst studying the role of roads in the growth of US cities between 1983 and 2003, estimate that an increase of one percent in a city's roadway leads to a 1.5 percent increase of employment and 1.3 percent larger population 20 years later.

However, depending on the accessibility measure and the characteristics of the areas under scrutiny (urban vs rural), transport accessibility can display mixed results. Transport accessibility can, on the one hand, influence positively population growth in sub-urban areas, and, on the other hand, the effect can turn negative in cities as people benefit of a better accessibility to out-migrate in favour of sub-urban or even rural areas (García-Lopez et al., 2015; Baum-Snow, 2007). Additional interesting evidence is discussed in Chi and Ventura (2011b), who argue that transportation accessibility is not found to affect population change directly but it rather acts as a facilitator of population flows. In other words, accessibility itself does not promote population change but it seems to strengthen the spatial lag effects of population change, thus having an indirect effect. Along this line, Jacobs-Crisioni and Koomen (2017) report a positive and significant effect of domestic accessibility on population density growth.

Glaeser and Kohlhase (2004) document the decline over the last century in the United States of the transportation cost for goods. Hence, we observe, as discussed earlier, a negative association between counties population growth and the counties share of employment in the agricultural or the manufacturing sectors. Over the same period, on the contrary, the transportation cost for people remained high and might have even increased because of the traffic congestion. This empirical fact, coupled with the change in the transportation modes (from railways to cars, airplane) have made more likely that cities and, more broadly dense areas, will host services.

A final element to be considered is the potential reverse causality between population growth and transportation investments. In this direction goes the recent study by Jacobs-Crisioni et al. (2016), which tests if population changes affect large scale investments in transport infrastructures. The authors conclude in favour of a bidirectional link between population growth and accessibility improvements. Also along this line, Duranton and Turner (2012) or García-López et al. (2015) instrument transport accessibility with long time lagged indicators of transport infrastructures in order to account for reverse causality.

2.1.4.5 Land Use and development

- **Rationale**

Land use and development refer to the quantity of land that a geographical area has available for its development. The development of a specific geographic area is determined by its geographic characteristics, built-up lands, vacant lands, cultural resources and legal constraints on the land use. **Geographic characteristics** and **land use** are often ignored in the empirical literature, though they are considered by some authors as important factors to explain the distribution and growth of the population (Chi and Ventura, 2011a; 2011b). On one hand, the importance of the geographic characteristics as determinants of population growth is discussed in Glaeser et al. (1995). People usually prefer living next to water and natural sources, on flat plain with extensive open areas and close to the coast. These authors point out that some dispersion observed in the rates of population growth can be greatly explained by geographic factors. On the other hand, the different use that is made of the land can also determine the growth and the distribution of the population in a geographical area. For instance, an area with a predominantly residential land use tends to display higher population growth than other areas more focused on agriculture.

- **Variables used**

Several measures have been used in the literature to capture the geographic characteristics of the area under scrutiny. Some studies collect topographical related information with variables such as the share of aquifers, elevation range and terrain ruggedness (Clark and Murphy, 1995; Beeson et al., 2001; Burchfield et al., 2006; Rappaport, 2007; Partridge et al., 2007; Winters, 2011; Duranton and Turner, 2012; García-López et al., 2015). Other studies include variables to proxy land use regulatory policy(Boarnet et al., 2005). Boarnet et al. (2005) include as proxies of land use the proportion of land devoted to single family residential, multi-family residential, mixed residential, services, industry and agriculture. Chi and Ventura (2011b) use the variables water, wetland, slope, tax-exempt lands (public and institutional land not available for development), and built-up lands to generate one index of developability, which approximates the proportion of developable lands.

- **Findings**

As observed in Clark and Murphy (1996) and Beeson et al. (2001), the presence of mountainous terrain has a negative influence on population growth. This suggests that the **topographical characteristics** of the geographical areas under scrutiny are to be accounted for when studying the determinants of population growth. Burchfield et al. (2006) also come to the interesting conclusion that mountains prevent sprawls while rugged terrain has the opposite effect. Interestingly enough, Winters (2011) notes that topography and the percentage of the area covered by water had significantly positive effects on the net migration of students to the nonmetropolitan US counties between 1995 and 2000. Other empirical studies indicate that **land use** variables are significant to explain population growth. For instance, Boarnet et al. (2005) provide evidence that the share of residential land were significant determinants to explain the population growth of the county of Orange (California, US), from 1980 to 1990. Chi and Ventura (2011) study the population growth at the minor civil division level in the state of Wisconsin (USA) from 1970-2000. The authors observe that the **constructed index of developability** is positively associated with population growth. Therefore, the more lands are available for building-up, the more likely population growth is to occur.

2.2 Preliminary empirical analysis: Determinants of population growth in Europe at regional level, NUTS3 level, 2000-2010

Drawing on the literature reviewed, the purpose of this section is to provide a preliminary assessment of the determinants of population dynamics in Europe at the sub-national level. To that end, we use the regional classification developed level by Eurostat in the 1970s, i.e., the Nomenclature of Territorial Units for Statistics (NUTS), which provides a single and uniform geographical division of the European Union into a hierarchical set of regions. The NUTS classification subdivides the economic territory of the Member States into three levels: NUTS1, NUTS2 and NUTS3. NUTS3 are subdivisions of NUTS2 which are themselves sub-divisions of NUTS1. NUTS1 represents major socioeconomic regions, NUTS2 captures basic regions used for the implementation of regional policies while

NUTS3 relates to smaller areas.[4] This is this lower level – NUTS3 – that we consider in the empirical analysis presented below.

We specify and estimate an econometric model that allows us to explain the population growth of the NUTS3 European regions over the period 2000-2010. As usually done in the literature, the explanatory variables of the model are measured at the beginning of the sample period (year 2000) to mitigate the possible problems of endogeneity or double directionality (Fukuda, 2012). The sample includes 1115 regions out of a total of 1315 European NUTS3 regions. We exclude NUTS3 regions (i) whose borders have changed over time because of a split or merge of the territories (ii) for which data were missing for the main variables of interest.

2.2.1 Data

The empirical analysis relies on the combination of several data sources. Data related to the socio-economic structure of the NUTS3 regions are drawn from Cambridge Econometrics' Regional Database (see https://www.camecon.com/european-regional-data/ for additional information). Cambridge Econometrics includes exhaustive information about the economic performance at sector, and regional level over time and for the entire set of EU countries.

Eurostat is the source of information for the remaining socio-economic, demographic and geographical variables included in the model. Those variables are listed and summarized in the next section, along with the model description.

2.2.2 Model description

The selected model belongs to the family of spatial econometrics models. This choice is led by the willingness to account for the presence of spatial interactions among NUTS3 that might affect their population growth.[5]

Following Vega and Elhorst (2015), we rely on the Spatial Lag of X model (SLX) to represent possible spatial interaction effects. In the literature, the use of the SLX is the point of departure when modelling phenomena that have a spatial component (Gibbons

[4] Note that two lower levels of Local Administrative Units (LAU) are also defined: the upper LAU level (LAU level 1, formerly NUTS level 4) and the lower LAU level (formerly NUTS level 5).

[5] The Moran's I test does not indeed reject the null hypothesis of absence of special dependence between NUTS3 regions, in terms of population growth. This supports for the need to account for spatial dependence in an econometric setting.

and Overman, 2012; Vega and Elhorts, 2015).[6] This specification has also been commonly used to model population growth (Boarnet et al., 2005; Cheshire and Magrini, 2006; Martori et al., 2014; Lewis and Stanley, 2016). The SLX model can be represented as follows:

$$Y = B \cdot \varphi + X \cdot \beta + C \cdot \delta + W \cdot Z \cdot \vartheta + U \qquad (1)$$

where $\boldsymbol{Y}$ is the population growth of the European regions over the period 2000-2010. Specifically, the growth is defined as the log difference of population in 2010 and 2000 ($Y = \log(Population_{2010}) - \log(Population_{2000})$).

The vector **B** includes variables capturing NUTS3 specific conditions in terms of degree of urbanization and localization. More precisely, we include a dummy variable named *Coastal* that equals one if the region has a sea border or at least half of its population within 50 km of the coast, zero otherwise. Recent data highlight that the population of the EU coastal regions increased more than the European Union as a whole (Eurostat, 2011). The degree of urbanization is captured through two binary variables[7]. The first one, named *Urban*, takes the value one if the region shows a percentage of population living in rural areas below 15 percent, and zero otherwise. On the contrary, the second variable, named *Rural,* equals one if the region shows a percentage of population living in rural areas higher than 50 percent and zero otherwise. Furthermore, we control for the fact that metro region or second-tier metro regions might exert specific population dynamics. This is captured by the two dummy variables *Capital* and *Second Tier.* The variable *Distance* controls for the distance to the closest capital or second-tier metro region.

The vector **X** contains a set of demographic and socio-economic variables which derive, again, from the literature reviewed above.[8] More precisely, the vector **X** includes a measure of the NUTS3 population density in 2000. The variable, named *Density,* controls for population pressure, and is defined as

$$Density = log\left(\frac{Population\ of\ the\ NUTS3}{Area\ of\ the\ NUTS3}\right) \qquad (2)$$

[6] Unlike other spatial models such as the spatial autoregressive model or the Durbin spatial model, the SLX model can be easily estimated by Ordinary Least Squares and the estimates can be directly interpreted, without any subsequent transformation.

[7] The rural-urban typology considered in this study is based on that one defined in Eurostat. The reference variable is another dummy variable called "intermediate" that takes value one if the share of population living in rural NUTS3 is between 15% and 50%.

[8] It is important to highlight that the framework of our analysis is based on Chi and Ventura (2011a; 2011b). However, one of the elements of originality of the current work - that our level of analysis is NUTS3 – constitutes a limitation when it comes to data availability. We could not find sufficient data for covering the category "Natural Environment", as no data are available for EU28 countries at NUTS3 level.

According to previous results reported in the literature review, we expect a negative, although small, influence of this variable on regional population growth rates. We also include the squared term of this variable (reported in the tables as $Density^2$) to capture potential non-linear effects. A priori we expect that low values of population density strengthen population growth until reaching a specific level of saturation. Beyond this level of saturation, increases in population density will lead to declines in the rate of population growth. We approximate transport accessibility, thereafter named *Accessibility,* with the logarithm of the potential accessibility index defined in Jacobs-Crisioni and Koomen (2017). In line with the convergence or catching-up effect hypothesis, we also include the logarithm of NUTS3 total population (henceforth named *Population*) at the beginning of the period (year 2000). The demographic composition of the NUTS3 is accounted for with a variable measuring the percentage of population aged 65 and above in 2000. Regions with an ageing population should experience lower subsequent population growth comparatively to regions with a younger population.

The vector **X** also includes three variables approximating the economic characteristics of the regions. More specifically, we include the employment growth of the European regions over the period 2000-2010 as well as the logarithm of the Gross Domestic Product (GDP) *per* capita of the regions in 2000 (in Euros at 2005 constant prices) to control for the economic dynamism of the regions. Following the existing literature, we also account for the structure of the regional economy with the inclusion of a variable measuring the weight of the industrial section in each region. This variable corresponds to the percentage of workers in the industrial sector in the year 2000 with respect to the total number of employees. These three variables are respectively named *GDP per capita, Empl Growth* and *Industrial Sector* in the rest of the paper. Lastly, we control for the stock of human capital with a variable measuring, in 2000, the percentage of population having a tertiary education. This variable, *Higher Education,* is not available at NUTS3 level, so we approximate it with its value at NUTS2 level.

The vector **C** includes country fixed effects. The country fixed effects will pick up the time invariant heterogeneity inherent to each country.

W is a weighting matrix which accounts for spatial effects incurring among neighboring regions. The weighting matrix used in this analysis is constructed following the indications given in Dall' Erba and Le Gallo (2008)[9]. Specifically, the matrix is based on the geographical distance between regional centroids and is defined as:

---

[9] Dall'Erba and le Gallo (2008) recommend the use of this weighting matrix in the European context due to the existence of islands. If a simple contiguity matrix had been used in the analysis, the weighting matrix would have rows and columns with only zeros for the islands. The use of this weighting matrix allows us to avoid this problem.

$$w_{ij} = \begin{cases} 0 & if\ i = j \\ \frac{1}{d_{ij}^2} & if\ d_{ij} \leq D \\ 0 & if\ d_{ij} > D \end{cases}$$

where d_{ij} is the is the great circle distance between centroids of region *i* and *j*. **D** is a subjective cut-off value. According to Dall' Erba and Le Gallo (2008), the lower quartile, the median and the upper quartile of the great circle distance distribution could be used as cutting values. The matrix **W** is row-standardized as $w_{ij} = \frac{w_{ij}}{\sum_{i=1} w_{ij}}$ which ensures that the sum of each row of **W** is equal to one.

The vector **Z** is a subset of the vector **X**; namely, population density, transport accessibility, GDP *per* capita and employment growth. **W·Z** corresponds to the average values of the neighboring regions for the variables included in vector **Z.**

In view of the literature reviewed before, Table 1 summarizes the expectations regarding the signs of the above mentioned variables in shaping population dynamics.

Table 1: Summary of the explanatory variables and expected effects.

	Explanatory Variable	Definition	Expected Effect
MATRIX X	**GDP *per* capita**	Logarithm of regional GDP per capita in 2000 (at 2005 constant prices)	+
	Empl Growth	Regional employment growth over the period 2000-2010	+
	Industrial Sector	Percentage of workers in the industrial sector in 2000.	-
	Higher Education	Percentage of population in the region with tertiary education in 2000.	+
	Accessibility	Logarithm of the potential accessibility index.	+
	Population	Logarithm of regional population in 2000.	-
	Pop over 65	Percentage of population aged 65 and above in 2000.	-
	Pop Density	Logarithm of the regional population density in 2000.	+
	Pop Density^2	Square of the variable log(DENSITY) in 2000	-
MATRIX B	**Coast Region**	Dummy variable that takes the value 1 if the region has coast, zero otherwise.	+
	Urban Region	Dummy variable that takes value 1 if the percentage of population of the region living in rural areas is below 15%, zero otherwise.	+
	Rural Region	Dummy variable that takes value 1 if the percentage of population living in rural areas is higher than 50%, zero otherwise.	-
	Capital Region	Dummy variable that takes value 1 if the region is capital metro region, zero otherwise.	+
	Second Tier Region	Dummy variable that takes value 1 if the region is second-tier metro region, zero otherwise.	+
	Distance	Distance to the closest capital or second-tier region (in Kilometers).	-

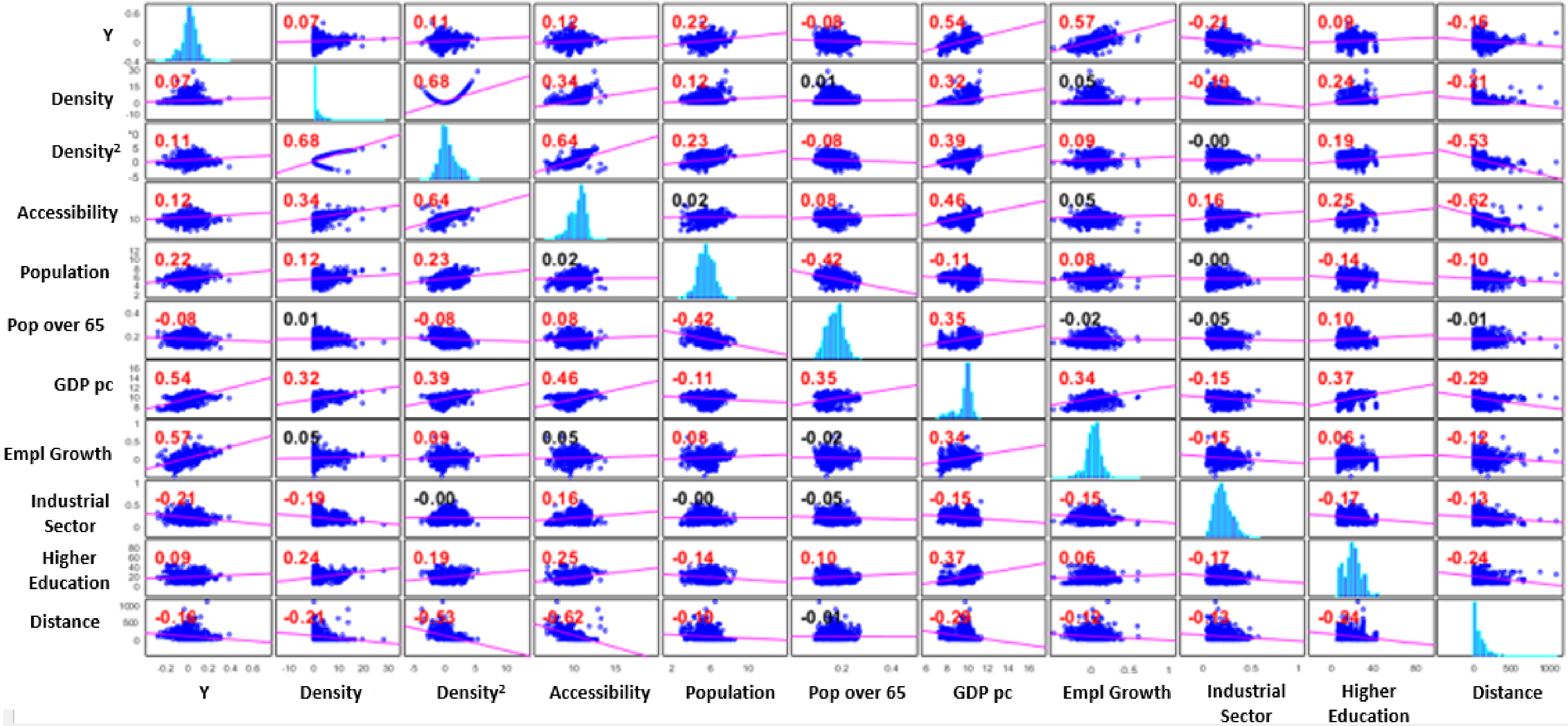
Table 2 reports the main descriptive statistics of the variables used in our model, and Figure 2 displays the pairwise correlation matrix. All independent variables are significantly correlated with population growth, and show the expected direction. Additionally, the independent variables are not highly linearly related among them, suggesting that there are no problems of multicollinearity. As expected *a priori,* the only exceptions are the relatively high correlations between the variable distance and transport accessibility variables (-0.62) as well as between population density and its squared value (0.68).

Table 2: Main Descriptive statistics of the variables used in the econometric analysis

	Mean	Standard Deviation	Min	Max
Population growth	0.0142	0.0751	-0.2755	0.383
GDP per capita	5.57	0.84	2.96	8.60
Empl growth	0.03	0.11	-0.49	0.60
Industrial sector	0.21	0.09	0.03	0.60
Higher Education	19.76	7.67	5.40	42.40
Accessibility	10.47	0.95	6.73	13.70
Population	5.57	0.84	2.96	8.60
Pop over 65	0.17	0.03	0.09	0.30
Pop Density	0.51	1.31	-3.71	5.30
Pop Density^2	1.98	3.25	0.00	28.20
Distance	90.11	107.96	0.00	1093.20
	Frequency			
Coast Regions	31.75%			
Urban Regions	24.48%			
Rural Regions	38.21%			
Capital Regions	6.01%			
Second Tier Regions	12.11%			

Source: Cambridge Econometrics' Regional Database and Eurostat data

Figure 2: Histogram of the variables and pairwise cross-correlation matrix (Pearson correlation coefficient).



Note: The correlation coefficients statistically significant are marked in red. The main diagonal of the pairwise matrix depicts the histogram of the variables

2.2.3 Empirical results

Results for the whole sample

Table 3 shows the estimated coefficients of the SLX models when different cut-off values **D** are used for the weighting matrix (lower quartile, median and upper quartile). The estimation is carried out by using ordinary least squares. The estimation process is robust to the choice of the cut-off value **D**; i.e., these estimated coefficients are not sensitive to the value of **D**. For that reason, hereinafter, we focus our explanation to the estimates when **D** is equal to the median of the "great circle distance" distribution.

The estimation of the SLX model provides information on the main determinants of regional population growth. The Adjusted-R^2 is around 0.71, which means that most of the variance of population growth over the period 2000-2010 is explained by the model. The main findings are summarized below.

The positive and significant effect of the **capital** and **second-tier dummies** indicates that the presence of **big cities** in a given region affects its population growth. To be more precise, metro capital regions have registered a population growth 3.26 percentage points greater than other NUTS3 regions. The effect of being second-tier metro region is relatively smaller (1.25 percentage points) in comparison with the capital metro regions.

The **distance to the closest capital or second-tier metropolitan area** has a negative effect on population growth: the greater the distance from the nearest big metropolitan area, the lower the growth of the population. This result corroborates the findings reported in other empirical studies (Lee et al., 2007; Partridge et al, 2008; Veneri and Ruiz, 2016). Our estimate suggests that population growth is reduced by 0.82 percentage points for every 100 kilometres away from the biggest metropolitan regions.

The **density of population** is not statistically different from zero whilst its square value is negative and significant. This finding suggests the existence of a **nonlinear relationship** between **population density** and **population growth**. Population growth decreases more than linearly with increasing population density. The negative effect of population density on population growth is also reported in Glaeser and Shapiro (2003).

The **level of accessibility** is positively associated with population growth but fails to reach a conventional level of statistical significance (*p-value* = 0.14).

Ageing population is associated with lower population growth. A one percentage point increase of population aged 65 or over in 2000 is associated with a population growth 0.37 percentage points lower in the following decade. This finding confirms results previously reported in the literature (Glaeser and Shapiro, 2003; Lee et al., 2007; Sohn, 2012).

Economic wealth is a driver of population growth. This result supports the idea that high-income areas experience more demographic dynamism (Glaeser and Shapiro, 2003; González-Val, 2015), and it contrasts with studies reporting a negative effect (Trendle, 2009). The estimated elasticity of growth with respect to the economic wealth suggests that a one percent increase of the GDP *per* capita in 2000 is associated with 0.06 percentage points higher population growth during the next decade.

The **regional rate of employment growth** has also a positive and significant impact on population growth. The most economically dynamic areas offer more opportunities of employment, attracting more in-migrants. Our estimate suggests that if the employment growth in one region increases by one percentage point, then its rate of population growth will be over the same period 0.17 percentage points higher.

Regions with a large **share of the employees working in the industrial sector** have experienced lower population growth over the period considered in the empirical analysis. The negative effect of employment in the manufacturing sector has been widely reported in previous literature (see, for example, Glaeser et al., 1995). As explained earlier, economic developing led, over the previous decades, to structural changes in the economies of most of the European regions. In addition, the economic crisis during the late 2000s and early 2010 hit more strongly the regions depending to a large extent on the industrial sector.

A priori, it would be expected a positive influence of the variable **tertiary level of education**. Instead, the influence is negative and significant. Keeping in mind that we already control for the level of economic dynamics of a region through several economic indicators, higher levels of education, *ceteris paribus*, might be associated with a delay in the age of marriage and in the age at first birth as well as a higher cost for raising children. This result needs however to be interpreted with a pinch of salt given that this variable is only measured at NUTS2 level.[10]

The spatial effects as measured by **W·Z** are statistically different from zero. In other words, the population growth in one region is not only affected by its own regional characteristics, but also by those of the neighbouring regions. More specifically, the rate of population growth of one specific region is enhanced if its neighboring regions are densely populated and display favorable economic conditions, both in terms of GDP *per* capita and employment opportunities. The magnitudes of the estimated coefficients

[10] Some lines of research are open in this point. First, we have to check if there is a bias due to the fact that our data to approximate education is at NUTS2 level, while our analysis is focused at NUTS3 level. Another interesting point is to explore the effects of the increasing access of women in the labor market.

suggest, for instance, that the population growth in a given region is fueled by 0.023 and 0.050 percentage points with increases of the two economic proxies, respectively GDP *per* capita and employment opportunities, by one percentage point. On the contrary, the attractiveness of a given region seems to be reduced if its neighboring regions are more easily accessible.

Table 3: Determinants of regional population growth in Europe, 2000-2010 - SLX specification

	VARIABLES	Cut-off values D for the weighting matrix		
		Lower Quartile	**Median**	**Upper Quartile**
MATRIX X	**GDP *per* capita**	0.059***	0.059***	0.060***
	Empl Growth	0.172***	0.173***	0.174***
	Industrial Sector	-0.077***	-0.078***	-0.078***
	Higher Education	-0.001***	-0.001***	-0.001***
	Accessibility	0.005*	0.005	0.004
	Population	0.0002	0.0001	-0.0001
	Pop over 65	-0.371***	-0.371***	-0.370***
	Pop Density	-0.001	-0.001	-0.002
	Pop Density^2	-0.003***	-0.003***	-0.003***
MATRIX B	**Coast Region**	-0.112	-0.0278	-0.003
	Urban Region	-1.196***	-1.197**	-1.198***
	Rural Region	-0.528	-0.527	-0.533***
	Capital Region	3.231***	3.260***	3.278***
	Second Tier Region	1.221***	1.247***	1.255***
	Distance	-0.008***	-0.008***	-0.008***
MATRIX Z	**W·Density**	0.008**	0.009*	0.009**
	W·Accessibility	-0.024***	-0.023***	0.023***
	W· GDP *per* capita	0.047***	0.050***	0.051***
	W· Empl Growth	0.201***	0.210***	0.212***
	Adjusted R^2	0.71	0.71	0.71
	Number of Observations		1115	

*Note: The symbols ***, ** and * indicate significance at the 1, 5 and 10 percent levels. The country dummies are not reported in this table to save space, but they are all statistically different from zero. The estimates associated to the variables in logarithms were already rescaled as in Wooldridge (2012) to allow for a direct interpretation.*

Heterogeneous effects across regions

Table 4 shows the estimation of equation (1) when the regions are divided according to their levels of income and urbanization. Starting with the grouping by level of income, some salient characteristics are worth mentioning. Within the group of the 25 percent poorest regions, the regions close to the coast experienced a significant higher rate of population growth. The growth in these regions was also strongly fueled by being capital metro region or, to a lesser extent, a second-tier metro region. The denser regions show a lower dynamism in terms of population growth, but they benefit from being surrounded by densely populated neighboring regions. In addition, we find that the initial level of population exerts a moderate but significant positive influence on population growth. This finding brings to light the existence of a divergence effect in terms of population growth among the poorest regions of the EU28. The population growth of the 25 percent richest regions is strongly influenced by their economic performance. Quantitatively speaking, the effect of employment opportunities is indeed much more important for the richest regions than for the poorest ones. Related to this, the employment growth in the neighboring regions also positively influences regional population growth.

Table 4: Determinants of Regional Population Growth, 2000-2010. Heterogeneous effects - Income levels and degree of Urbanization

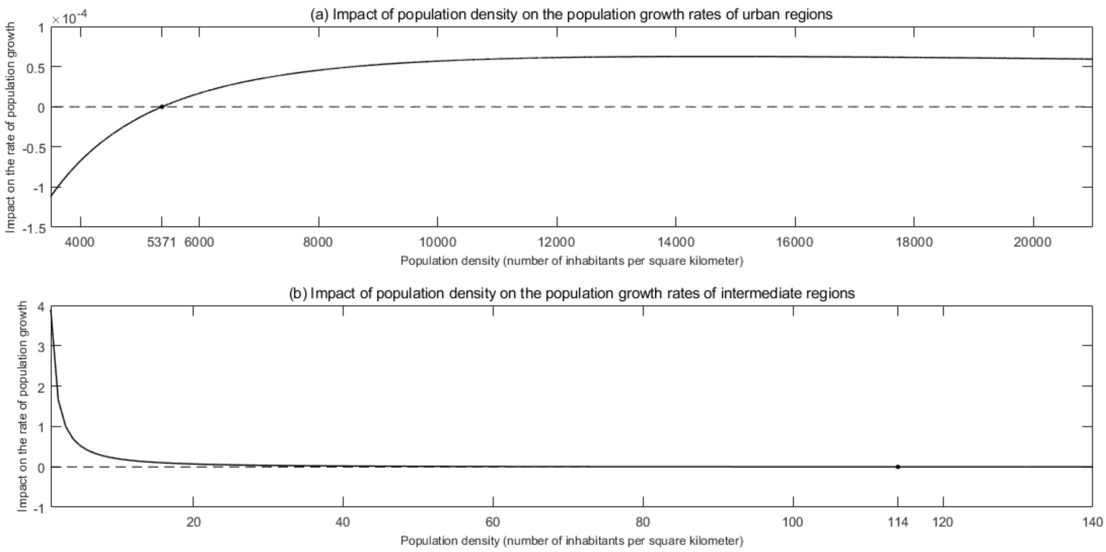
VARIABLES	Grouping by Income		Grouping by level of Urbanization		
	25% Poorest Regions	25% Richest Regions	Urban Regions	Intermediate Regions	Rural Regions
Coastal Region	3.548***	-0.802	-1.037	1.071*	0.053
Urban Region	0.068***	-0.387	-	-	-
Rural Region	0.006	-1.097	-	-	-
Capital Region	7.895***	1.148	3.160***	5.246***	5.983***
Second Tier Region	1.760*	0.678	0.042	2.368***	3.232**
Pop Density	-0.004	0.002	-0.078**	0.039*	0.057***
Pop Density^2	-0.006***	-0.002	0.005**	-0.004**	-0.006***
Accessibility	-0.005	-0.010	0.003	0.021**	-0.002
Population	0.013**	-0.005	-0.004	0.003	0.005
Pop over 65	-0.6311***	-0.569***	-0.416***	-0.410***	-0.392***
GDP *per* capita	0.033**	0.037***	0.046***	0.060***	0.092***
Empl Growth	0.070***	0.212***	0.114***	0.223***	0.139***
Industrial Sector	-0.087**	-0.041	-0.037	-0.124***	-0.072**
Higher Education	-0.001	0.002***	-0.001	-0.001***	-0.001
Distance	-0.003	-0.006	-	-0.007***	-0.010***
W· Density	0.025***	-0.001	0.004	0.006	0.002
W· Acessibility	-0.0012	0.010	-0.011**	-0.016	-0.034
W· GDP *per* capita	0.021*	-0.022	0.044***	0.063***	0.019
W· Empl Growth	-0.095	0.223**	-0.026	0.280***	0.396***
Adjusted R^2	0.78	0.63	0.60	0.72	0.80
Number of Observations	279	279	273	416	426

*Note: The symbols ***, ** and * indicate significance at the 1, 5 and 10 percent levels. Country dummies are included in all specifications.*

There are also some interesting results when the estimation of equation (1) is split by levels of urbanization. First, urban regions benefit from having a large city but only when this city is the capital of the country. Second, intermediate regions that are close to the coast present significant higher rates of growth (an increase of 1.07 percentage points with respect to non-costal regions). Third, the effect of employment opportunity is much more important for intermediate regions than for urban and rural regions. Fourth, the population density and its squared term are statistically associated with population growth for both the urban and intermediate regions. However, the effect of density is quite different for these groups of regions. The latter finding deserves a more detailed explanation. Figure 1 graphically depicts the relationship between population growth and population density for urban regions (Figure3.a) and intermediate regions (Figure3.b), while holding all other variables constant.

Looking first at Figure 3.a, we observe that for urban regions with low density levels, the relationship of the link between population growth and density is negative though at a decreasing rate. When density reaches a certain value (5371 inhabitants per square kilometer), the relationship turns positive and stable around a value slightly greater than zero. Knowing that 98 percent of the urban regions (268 out of 273 urban regions) show a population density lower than the threshold value, this implies that, in general, population density exerts a significant negative effect on population growth. By contrast, according to this turning point of the density variable, the region of Paris (FR101) and its neighboring metro regions Hauts-de-Seine (FR105), Seine-Saint Denis (FR106), as well as for the regions of Brussels (BE100) and Bucharest (RO321) should have experienced an agglomeration effect in the period 2000-2010. Figure 3.b shows the relationship between population density and population growth for regions classified as intermediate. In this case, the relationship between population density and population growth is close to zero, except at very low levels of population density.

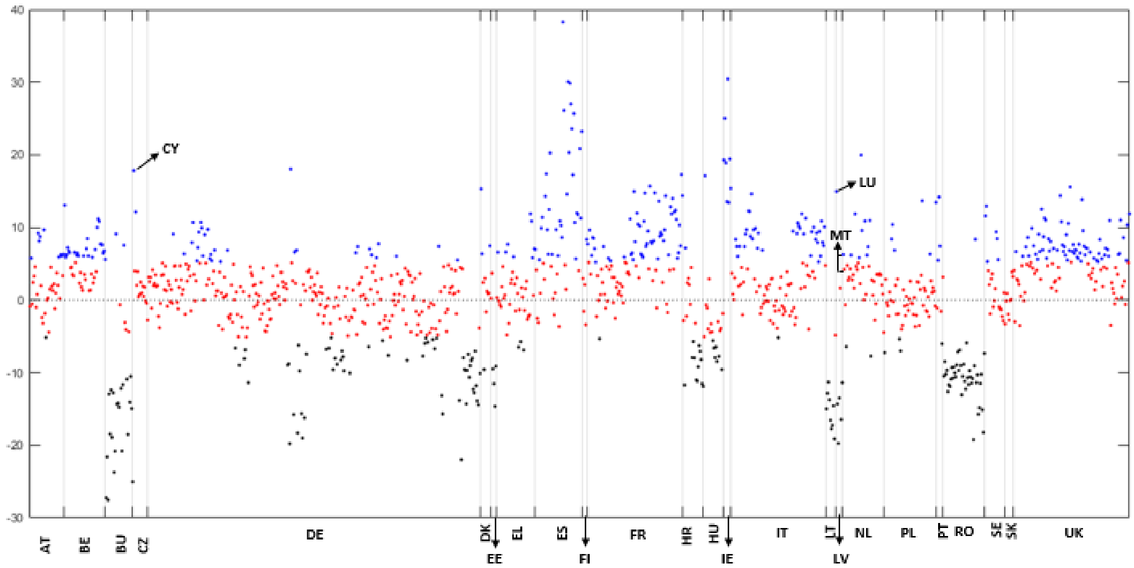
Figure 3: Population density and population growth: urban and intermediate regions



Note. The figures are based on the estimations reported in Table 4.

Figure 4 represents the regions grouped according to the speed of population growth over the period under study. The regions are classified into three groups: high-growth (blue dots), stagnation (red dots) and negative-growth (black dots). The classification was performed by using the k-means clustering method.[11] There is clearly a country clustering with, for instance, the regions of Romania, Bulgaria Lithuania and Latvia presenting all moderate rates of population growth (i.e., they belong to the group of negative-speed and/or stagnation). Conversely, regions in countries such as Belgium, Spain, France, Ireland, Italy and UK show high rates of population growth.

Figure 4: Classification of the regions according to their population growth rate over the period 2000-2010.



Note: The clustering process was done by using the K-means method considering the squared Euclidean distances.

The analysis is then extended to test whether the explanatory variables have heterogeneous effects across the three groups of regions defined by the cluster analysis. Table 5 displays the results. Some interesting differences emerge. First, being a coastal region is positively associated with population growth for the sub-group of negative-growth regions, while it does not affect the population growth of the two other types of regions. Second, capital regions have always experienced greater population growth but the magnitude of the estimated coefficients suggests that the effect was much stronger for the regions included in the negative-growth group of regions than for the other regions. Third, second-tier metropolitan regions do not seem to have benefitted from a larger population growth, and this, irrespective of the group of regions. Fourth, urban regions have displayed lower population growth compared to the intermediate regions, irrespectively of the type of regions. Fifth, the negative and significant coefficient associated with the population size in 2000 for the regions belonging to the high-growth group indicates that there is –only for this group- a convergence effect (main reverse or

[11] Martinez and Martinez (2015) explain in detail the k-means method used in our classification, as well as its implementation in Matlab. In our case, we assume as distance to be minimized the squared Euclidean distance.

catching up effect). The effect is the opposite for negative-growth regions, i.e. there is a divergence effect. Sixth, the effect of population density is statistically negative for regions in the high-growth and stagnation groups, but approaches zero as population density increases. Seventh, tertiary education has a positive and significant effect on the population growth of the intermediate regions, while the effect is negative for the high-growth group of regions and not significant for the stagnation group.

Table 5: Determinants of Regional Population Growth, 2000-2010. Heterogeneous effects – Population Growth Intensity

	High Growth	**Stagnation**	**Negative Growth**
Coastal Region	5.886	-3.829	2.387*
Urban Region	-1.298***	-0.094***	-3.747*
Rural Region	0.526	-0.707***	-0.019
Capital Region	1.914**	0.022***	7.660***
Second Tier Region	0.796	0.264	-0.0484
Density	0.008*	0.001	-0.004
Density^2	-0.002*	-0.002***	0.001
Accessibility	-0.009	0.002***	0.057**
Population	-0.008***	0.002	0.015*
Population over 65	-0.230***	-0.227***	-0.220
GDP per capita	0.0315***	0.0250***	0.008
EMPL growth	0.122***	0.046***	-0.014
Industrial Sector	-0.006	-0.054***	0.001
Higher Education	-0.001*	0.001***	-0.004***
Distance	-0.001**	-0.001	0.011*
W· Density	-0.001	0.003	0.006
W· Accessibility	-0.003	-0.002	-0.015
W·GDP per capita	0.039*	0.019***	0.005
W· Empl Growth	-0.081	0.115***	0.053
Adjusted-R2	0.44	0.36	0.54
Number of Observations	310	633	172

*Note: The symbols ***, ** and * indicate significance at the 1, 5 and 10 per cent levels.Country dummies are included in all specifications.*

Finally, Table 6 shows the estimation of equation (1) when the sample is split into two sub-samples: (i) countries having joined the EU before 2004 (EU15) and (ii) countries which became EU members since the 2004 enlargement (EU 13).[12] The comparison of

[12] The first group of countries includes Austria, Belgium, Germany, Denmark, Greece, Spain, Finland, France, Ireland, Italy, Luxembourg, the Netherlands, Portugal, Sweden and the UK. The second group of countries is made of Bulgaria, Croatia, Cyprus, Czech Republic, Estonia, Hungary, Lithuania, Latvia, Malta, Poland, Romania, Slovakia and Slovenia.

the estimates between both groups reveals some interesting differences. First, being a coastal region is only significant for the regions belonging to the EU13. Second, the rural regions of the EU15 show significantly lower rates of population growth than intermediate regions. Third, the effect of being capital is positive and significant for both groups of regions; however, the effect is stronger for EU13 regions (5.02 vs. 3.27 percentage points). Fourth, the economic conditions are important drivers of population growth both for the EU15 and EU13 regions but the magnitude of the coefficient is larger for the former groups of regions. Finally, it seems that conditional on the other covariates, education has a negative influence on population growth in the EU15 whilst the opposite is found for the EU13.

Table 6: Estimates of the SLX model to explain population growth for the EU-28 regions according to the membership to the EU before and after the enlargement of 2004.

VARIABLES	Old Member States (entrance before 2004)	New Member States (entrance after 2004)
Coastal Region	-0.084	3.795***
Urban Region	-1.1554***	0.573
Rural Region	-0.928**	1.355*
Capital Region	3.272***	5.021***
Second Tier Region	1.141**	1.503
Density	-0.001	0.001
Density^2	-.002***	-0.008***
Accessibility	0.003	-0.003
Population	-0.001	0.029***
Population over 65	-0.323***	-0.563***
GDP per capita	0.057***	0.032**
EMPL growth	0.208***	0.051**
Industrial Sector	-0.049**	-0.090**
Higher Education	-0.001***	0.002**
Distance	-0.001***	-0.001
W· Density	0.002	0.032***
W· Accessibility	-0.014*	-0.026
W·GDP per capita	0.049***	0.022
W· Empl Growth	0.332***	-0.066
Adjusted-R2	0.63	0.80
Number Observations	905	210

*Note: The symbols ***, ** and * indicate significance at the 1, 5 and 10 per cent levels. The country dummies are not shown in this table to save space, but they all showed a positive sign and were statistically significant at the standard levels of significance.*

2.3 Conclusions and future further research avenues

Studying the drivers of population growth is a prolific research topic that has aroused the interest of many researchers in the last decades. The purpose of this empirical section was to take part to this debate by shedding some light on the main determinants of population growth of the NUTS3 European regions over the period 2000-2010. Our preliminary results support the main findings reported in the literature. More specifically, the economic conditions play a key role in shaping population growth. Not only do the economic characteristics of the own regions matter, we also find that the economic features of neighboring regions are also important drivers of population dynamics.

Ceteris paribus, capital or second-tier metro regions have experienced larger population growth than the other EU28 regions over the period 2000-2010. Also, along this line, the further is the closest capital or second-tier metro region, the lower is the regional population growth. Urban regions have experienced lower population growth in comparison to regions with a larger share of rural population. Finally, as expected, ageing and denser regions experienced less population growth.

As a complementary analysis, we also investigate if there are some heterogeneous effects, i.e., if the determinants of population growth differ across sub-groups of regions. For this, the population growth equation has been estimated on sub-samples of regions classified according to their income, level of urbanization, patterns of population growth as well as whether they belong to the EU15 or EU13. In general, the results of this analysis support our previous findings. However, there are also some important divergences that deserve to be mentioned. First, although having a coast was not a significant determinant of regional population growth on the full sample, it turns out to be asset when the analysis is restricted to the 25 percent poorest regions or to the EU13 regions. Having a coast is also beneficial for the sub-sample of regions having experienced a negative population growth during the period under scrutiny.

Second, whilst education has a significant negative effect on population growth on the full sample – a result that needs further analysis and which should be interpreted with cautious given that the variable is measured at NUTS2 level - the opposite is found for the poor regions as well as for the ones part of EU13. Third, the effect of population density is negative, except for the regions classified as urban.

The results shown in this section are preliminary and call for additional empirical analysis as explained below:

The **time coverage** of the empirical analysis is too short. We have to consider blocks of 5-10 years and re-estimate the population growth equation over a long time period, i.e., from 1990 to 2010. If possible, we should even try to cover a 25 years period (1990-2015). However, we must pay attention, when defining the time-frame of our sample, to the possible distorting effects resulting from the economic crisis.

We should also employ **spatial panel data econometrics** if we cover more than one time period and compare the results with those obtained with the cross-section approach used in the present study. The SLX model, which was the specification selected for the empirical analysis, is recommended as the starting point in the literature on spatial econometrics. This is also the most common approach of the existing papers on the determinants of population growth when spatial effects are included in the analysis. Nevertheless, we need to try other spatial specifications such as the Spatial Autoregressive (SAR) model, the Spatial Error (SE) model, the Spatial Durbin (SD) model and the Spatial Durbin Error (SDE) model. Once estimated all these models, we must select the one that best fits the data.

The **weighting matrix** used in our study has several advantages that justify its use, such as reducing the risk of not including islands. However, a robust econometric exercise requires checking the sensitivity of our results to the use of alternative weighting matrices (i.e., contiguity matrix, distance-based matrices and nearest-neighbour matrices).

The **geographical unit for the analysis** is defined at NUTS3 level. Working at this level is not without problems since NUTS3 regions are quite heterogeneous. As suggested in previous exchanges with DG REGIO, we should aggregate all the NUTS3 belonging to the same metro region into a single NUTS3 region.

The variable used to measure **population density**; i.e., the number of inhabitants *per* square kilometre, should be reconsidered. Even though this is the most common measure of population density, it is not clear if it is the most suitable variable when it comes to estimate the relationship between population density and population growth. The reason is that this measure of density includes areas that are not appropriate for human settlement. The number of inhabitants *per* inhabited area or *per* arable land might be better indicators in this context.

In addition to the variable measuring the importance of the industrial sector, **the share of the population in the agriculture** should also be among the covariates. Other potential explanatory variables to be considered include **meteorological** variables (e.g., temperature and precipitations), additional proxies for capturing the transport accessibility (e.g., airports) and demographic characteristics of the region (e.g., ratio males/females) as well as for instance indicators of the regional level of **social capital** (e.g., quality of the institutions, voting turnout, trust).

The estimated effect of the **education** related variable is negative on the full sample. These counter intuitive findings should be also further explored.

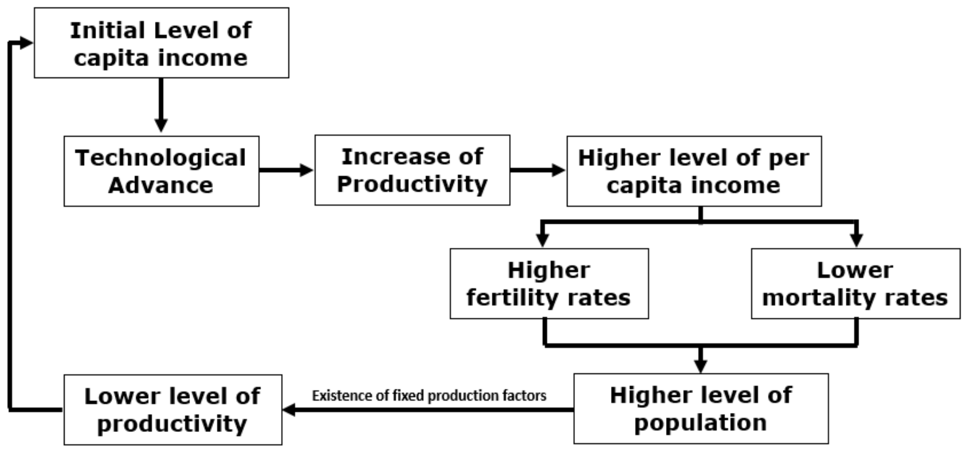
3 Long run relationship between population growth and economic growth: Literature review and preliminary empirical evidence

3.1 Introduction

Economic development and population are crucial pillars of both economic and demographic studies. Understanding the relationship between these two dimensions has been the subject of a long debate over the previous decades. The discussion is theoretically and empirically focused on exploring whether population affects and/or is affected by economic development. From a theoretical point of view, there are at least three alternative assumptions behind the relationship between population and development (Darrat and Al-Yousif, 1999).

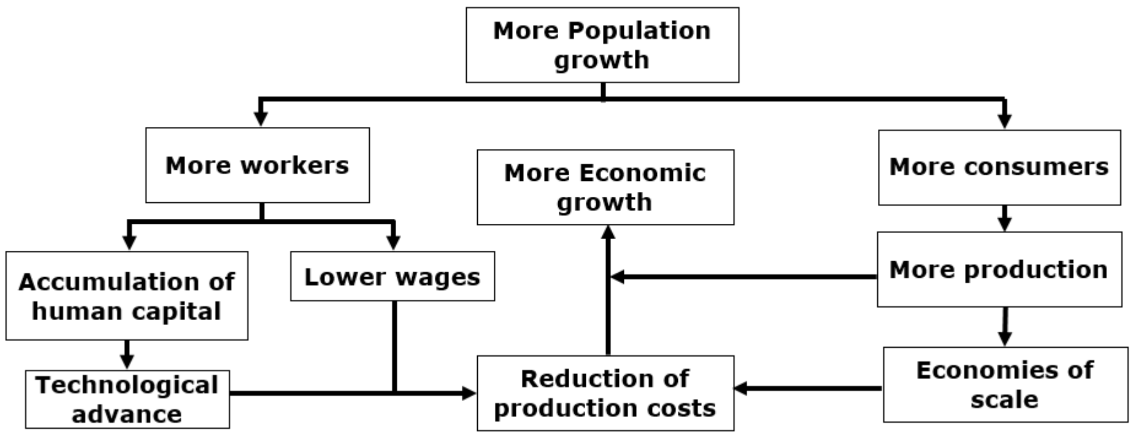
First, the "Orthodox" or "Malthusian" assumption states that there is a negative impact of population on economic development. Figure 5 graphically explains the mechanisms through which an increase in population growth leads to a decrease in income levels. According to this view, any technological advance causes a rise in productivity levels, involving an improvement in the initial levels of *per* capita income. The increase in *per* capita income leads to higher population through higher fertility and lower mortality. However, given that some production factors are fixed (e.g., land), higher levels of population is also associated with lower marginal productivity of labor, which translates into higher marginal costs and lower *per* capita income. Income *per* capita thus decreases until reaching its level before the technological progress. This line of reasoning suggests to limit fertility in developing and over-populated countries.

Figure 5: Theoretical relationship between population and economic growth according to the Malthusian approach



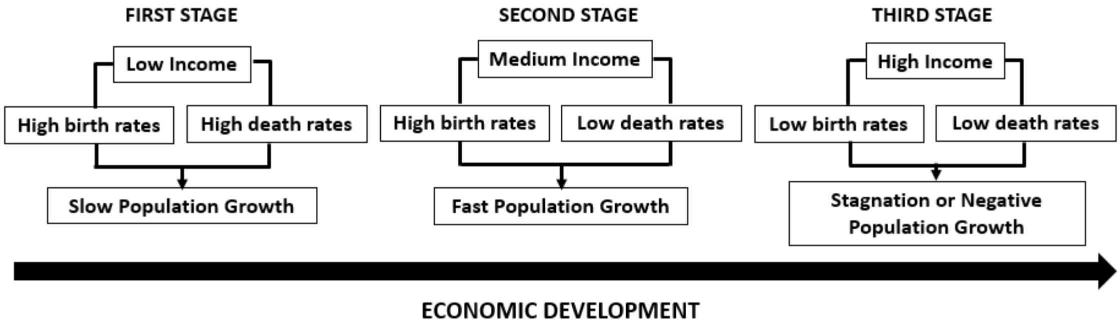
Contrary to the postulates of the Malthusian view, the "revisionists" do not consider that there is a clear and robust negative effect of population growth on economic growth. The rationale behind this view is twofold. On one hand, population growth increases the labor force, reducing wages and production costs. Additionally, having more workers implies more human capital accumulation and, therefore, more possibilities to innovate and spurring technological advances. On the other hand, population growth raises the number of potential consumers. Firms produce more and, due to the existence of economies of scale, the average production cost decreases. Figure 6 graphically summarizes the main arguments that support the revisionist view.

Figure 6: Theoretical relationship between population and economic growth according to the Revisionist approach.



The "demographic transition" theory is an alternative theoretical approach to the Malthusian and revisionist ones that reverses the causal link. This theory postulates that changes in income drive, at least partially, population growth. The impact of income will be positive or negative depending on the phase of economic development. In the first stage of development, characterized by low income levels, there will be high birth rates and high mortality rates, resulting in a slow population growth. In the second stage, as economic conditions improve, the population grows since the birth rates continue to be high and the death rates are reduced. In the final stage of economic development, societies have reached high levels of income. The low fertility and mortality rates lead to stagnation or even negative population growth. Figure 7 summarizes and describes the different stages of the relationship between population and economic growth.

Figure 7: Theoretical relationship between population and economic growth according to the demographic transition theory.



In the following section we review the most important articles that explore the relationship between population growth and economic development. Afterwards, we also perform a preliminary econometric analysis based on a three-step procedure to study whether there is a long run relationship between population and economic development for the EU28 Member States over the period 1960-2010.

3.2 Literature review

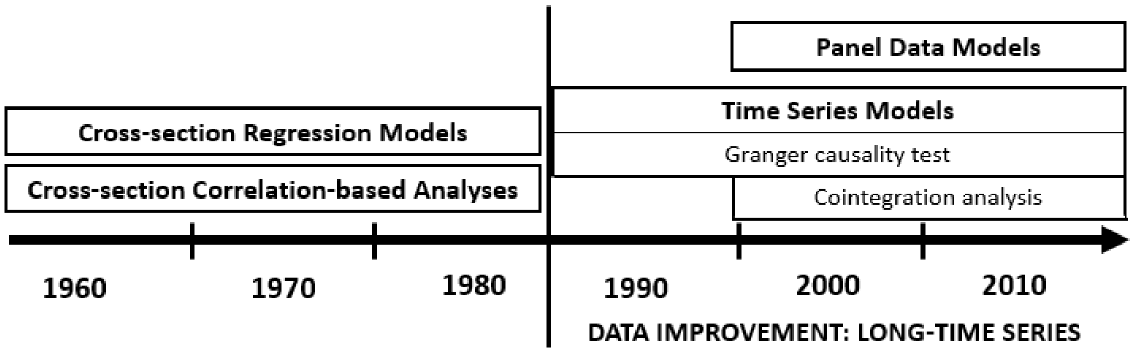
The relationship between population growth and economic development has been the subject of many empirical studies over the last decades (Thirlwall, 1994; Furuoka, 2013). The existing empirical evidence offers rather contradictory findings.

Most of the empirical research focuses on studying the relationship between population and economic development for the less developed countries (LDC) and overpopulated countries such as China (Hasan, 2015), Bangladesh (Nakibullah, 2010), India (Dawson and Tiffing, 1999),Thailand (Furuoka, 2009), Indonesia (Furuoka, 2013), countries of Central Asia (Savas, 2008), Mexico (Garza-Rodríguez, 2016) and other Latin-American countries (Thornton, 2001). In addition, a few other studies are interested in testing the relationship between population and economic development whilst relying on a large group of countries (Jung and Quddus, 1986; Faria et al., 2006; Bruecker and Schwandt, 2015; Chang et al., 2017).

Regarding the methodology, Figure 4 graphically displays the different methods used over time to study the relationship between population and economic development. The first studies date back to the 1960s and 1970s when the relationship has been explored using cross sectional data. These studies were based on simple correlation coefficients or on regression models estimated by OLS whilst controlling for a set of covariates (e.g., illiteracy rates, Gini's coefficient percent of urban population, among other variables). The empirical results, based on data from the 1960s and 1970s, could not find statistically significant association between population and economic development (Kelley, 1988). However, some researchers reopened the debate using data from the 1980s, as they found some evidence of significant negative correlation (Kelley and Schmidt, 1994) between these two dimensions. The studies based on regression estimates provided mixed results, with some concluding in favor of a positive relation, whilst other observed a negative association. Jung and Quddus (1986) offer a detailed overview of the literature based on cross-section regression models. The authors are, however, critical about the findings reported in these studies because none of them ponder the direction of the causal relationship between population growth and economic development. Moreover, and even more importantly, cross-sectional evidence is used to derive conclusions about what it is essentially a dynamic phenomenon (Simon, 1977).

This situation changed in the 1990s when researchers got access to long time series data. Jung and Quddus (1986) was one of the first studies that employed a time-series regression analysis to properly examine the relationship between population and economic development over time. Specifically, the authors use the Granger causality test to assess the causal direction of the relationship. The most recent studies tend to rely on the Granger causality test only as complementary tool to the cointegration analysis. Cointegration analysis consists in exploring whether there is a long term relationship between population and economic growth.[13] The Granger causality test has been used together with the Engle-Granger cointegration test (Hasan, 2010), the Johansen cointegration test (Dawson and Tiffin, 1999; Thornton, 2001; Furuoka, 2013), the Auto Regressive Distributed Lag bounds testing approach to cointegration (Furuoka, 2009; Savas, 2008), the Breitung-cointegration test (Furuoka, 2013) and the Gregory-Hansen cointegration test (Garza-Rodríguez et al., 2016). In addition, having longer time series data not only fostered the use of time series models but also allowed researchers to apply panel data econometrics (Brander and Dowrick, 1994; Faria et al., 2006; Huang and Xie, 2013; Bruecker and Schwandt, 2016; Sibe et al., 2016).

Figure 8: Representation of the methodologies used to check the relationship between population and economic growth.



The most recent empirical evidence provides also mixed conclusions on the link between population and economic development. This might be driven by the fact that these studies cover different set of countries and time periods and/or are being based on different econometric methods. While some authors report a statistically significant relationship, many others conclude that such relationship does not exist. Among those studies claiming a statistically significant relationship, the direction of the effect is very often contradictory. As explained by Furuoka (2009), there are four possible types of results:

[13] Some recent studies such as Kapuria-Foreman (1995) and Nakibullah (2010) and Chang et al. (2016) still use the Granger causality test to examine causality. Chang et al. (2017) combine Granger causality test with panel data and bootstrapping to determine the linkage between population and economic growth for 21 countries over the period 1870-2013.

1. Bi-directional causality; i.e., there is a causal relationship both from population growth to economic development and, from economic development to population growth.
2. No relationship between population growth and economic development.
3. Population-driven economic growth which means that there is only a unidirectional causal link from population growth to economic development. A positive relationship supports the revisionist view, while a negative one would be in line with the Malthusian approach.
4. Economic development-driven population growth by which there is only a unidirectional causal relationship from economic development to population growth. Such a finding would corroborate the demographic transition theory.

Table 7 summarizes the main findings of the literature according to the four possible results described in Furuoka (2009). The table reports the direction and the significance of the relationship. As we can see, some recent studies document a bi-directional positive causal relationship between population and economic development. Specifically, Garza-Rodríguez et al. (2016) finds evidence of positive bi-directionality for Mexico. Savas (2010) also finds a positive bi-directionality for five countries of Central Asia namely Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan and Uzbekistan while Pegou-Sibe et al. (2016) reach the same conclusion for a set of thirty populated countries. These findings support the revisionist view and, second, suggest that these countries are in the second stage of the demographic transition.

The finding of a positive bi-directional contrasts with the results reported by Jung and Quddus (1986), Huang and Xie (2013), Chang et al. (2017) and Thornton (2001) in a cross-country setting, as well as by Dawson and Tiffin (1998) using data from India. All these studies indeed found that population and *per* capita income neither stimulate nor stifle each other.

Several studies also conclude in favor of a unidirectional causal relationship flowing from economic development to population growth. Kapuria-Foreman (1995) and Nakibullah (2010) reach this conclusion using data respectively from India and Bangladesh whilst Brueckner and Schwandt (2015) report findings along the same line with a panel of 139 countries.[14] In contrast, some studies such as Kapuria-Foreman (1995), Darrat and Al-Yousif (1999) and Hasan (2010) find that economic development reduces the population growth of the countries under scrutiny, namely Sri Lanka, Syria, and China.

Overall empirical studies reflect the different stages of demographic transition and they seem more in line with the view that population growth does not harm economic growth.

[14] See also Chang et al. (2017) who conclude that economic development exerts a positive influence on the rate of population growth in 7 out of the 21 countries considered in their study.

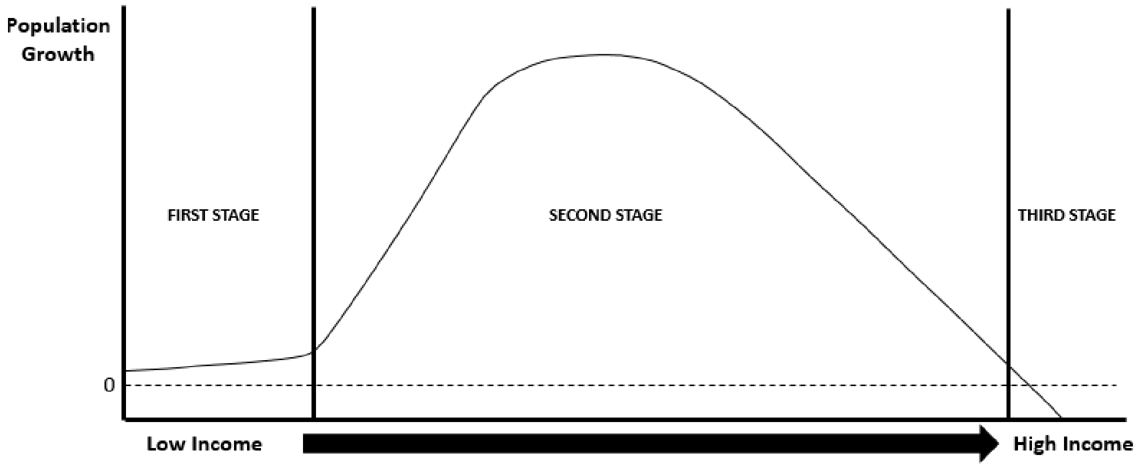
For instance, Darrat and Al-Yousif (1999) report that population growth stimulates economic growth in more than half of the twenty developing countries covered by their empirical analysis. Furuoka (2009, 2013) also shows that population growth has a positive impact on economic development in Thailand and Indonesia. The study of Hasan (2010) report that the growth of population in China was a stimulus for its economic development through the realization of favorable economies of scale and low labor costs. Similarly, Kapuria-Foreman (1995) finds a positive impact of population on economic development for China, as well as for Guatemala, Turkey, Chile and Mexico; but the effect seems to be negative for India.

A simple look at Table 7 gives evidence of the lack of consensus in the literature. The lack of consensus might be partly driven by the fact that the relationship between population and economic development is nonlinear. Using a panel of 125 countries over the period 1950-2000, Faria et al. (2006) indeed find an inverted U-shape relation between the two time series. Population increases as income increases, but at a decreasing rate. At a certain threshold level, the increase in income leads to a decline in population growth. Figure 9 shows the non-linear relationship proposed in Faria et al. (2006) and its connection with the different stages of the demographic transition theory.

Table 7: Summary of the empirical literature on the relationship between population and economic growth.

		Relationship	Reference	Countries	Period
POSITIVE-BIDIRECTIONAL CAUSALITY		E —+→ P P —+→ E	Savas (2010)	Kazakhastan, Kyrgyzstan, Tajikistan, Turkmenistan and Uzbekistan	1989-2007
			Sibe et al. (2016)	30 most populated countries	1960-2013
			Garza-Rodríguez et al. (2016)	México	1960-2014
			Chang et al. (2017)	Japan	1952-2013
INDEPENDENCE		P —?→ E E —?→ P	Dawson and Tiffin (1998)	India	1950-1993
			Thornton (2001)	Argentina, Brazil, Chile, Colombia, Mexico, Peru and Venezuela	1900-1994
			Huang and Xie (2013)	90 countries	1980-2007
			Chang et al. (2017)	Austria, Brazil, Canada, Finland, Italy, The Netherlands, Norway, Portugal, Sweden, UK, USA	1952-2013
ECONOMIC-DRIVEN POPULATION: DEMOGRAPHIC TRANSITION THEORY	Second Stage	E —+→ P	Nakibullah (2010)	Bangladesh	1959-1990
			Bruecker and Schwandt (2015)	139 countries	1960-2007
			Kapuria-Foreman (1995)	India	1950-1990
			Chang et al. (2017)	Belgium, Denmark, France, Spain, Germany, New Zealand and Switzerland	1870-2013
	Third Stage	E —-→ P	Kapuria-Foreman (1995)	Sri-Lanka and Syria	1950-1990
			Darrat and Al-Yousif (1999)	China and India	1950-1997
			Hasan (2010)	China	1952-1998
			Chang et al. (2012)	Uruguay	1952-2013
POPULATION-DRIVEN ECONOMIC GROWTH: REVISIONIST VIEW		P —+→ E	Darrat and Al-Yousif (1999)	Argentina, Brasil, Chile, Ghana, Guatemala, Indonesia, Mexico, Morocco, Pakistan, Peru, Phillippines, Syria, Turke and Uruguay.	1950-1997
			Kapuria-Foreman (1995)	China, Guatemala, Turkey, Chile and Mexico	1950-1990
POPULATION-DRIVEN ECONOMIC GROWTH: MALTHUSIAN VIEW		P —-→ E	Kapuria-Foreman (1995)	India	1950-1990
			Chang et al. (2017)	Sri-Lanka	1952-2013

Figure 9: Representation of the non-linear relationship between population and economic growth according to the findings in Faria et al. (2006).



3.3 Preliminary empirical analysis: population growth and economic development, EU28, 1960-2010

The empirical analysis presented in this section aims at providing some preliminary results on the relationship between population growth and economic development in EU28 over the period 1960-2010. The empirical strategy consists in the following three-steps procedure:

1. Search for a long run equilibrium between the two aggregates by means of the Auto Regressive Distributed Lag (ARDL) bounds testing approach;
2. Study of the causal relationships, and direction of such relationship, using the Granger causality test;
3. Exploration of the sign of the relationships between population and economic development.

The data come from the World Bank database (https://data.worldbank.org/region/european-union). The variable population (POP) measures the number of inhabitants in the EU28, while economic development (GDPpc) is approximated by the EU28 GDP *per* capita (constant 2010$). We rely on annual data over a period going from 1960 to 2010. Figure 10 depicts the time plot of the two variables and Table 9 reports some descriptive statistics. Both variables display an upward trend. The series are transformed into logarithms to reduce the variability and asymmetry in their distribution.

Figure 10: Time plot: population and GDP *per* capita, 1960-2010

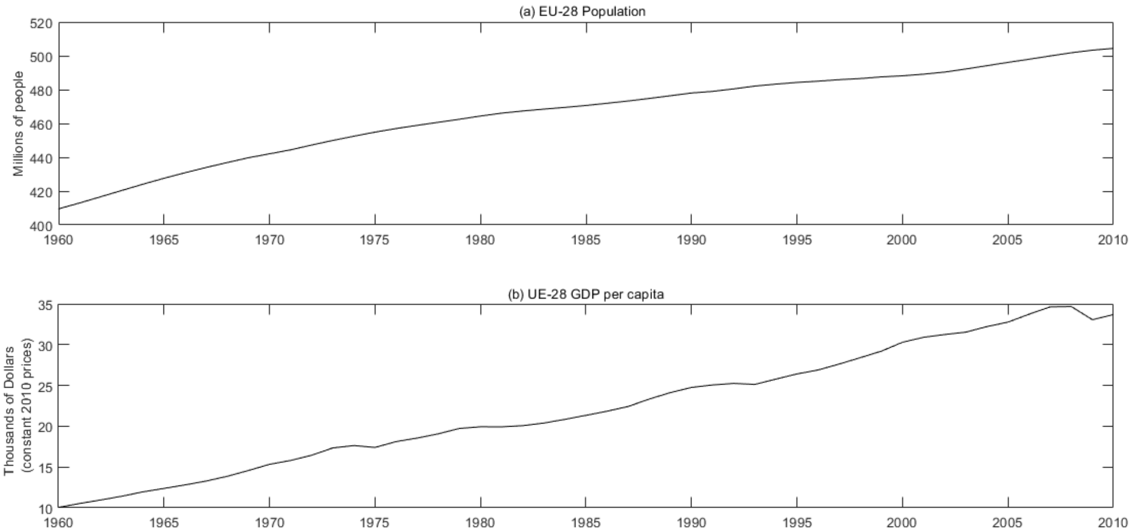


Table 8: Descriptive statistics of population and GDP *per* capita, 1960-2010

	Mean	Standard deviation	Maximum	Minimum
POPULATION (in millions of people)	466.19	25.83	504.42	409.49
GDP (GDP *per* capita in thousands of constant 2010 US$)	22.25	7.35	34.67	10.05

First step: Testing for a long-run relationship

The first step of the empirical analysis strategy is to verify whether there is a long-run equilibrium between the variables population GDP *per* capita. The study of this relationship cannot be done by estimating a simple regression model. Specifically, a high degree of fit and statistically significant parameters can be observed even when there is no underlying causal relationship between the variables. This problem, known as spurious regressions, is common when working with economic time series that are non-stationary (i.e., they have a linear or exponential time-trend).

We then first analyze if the two series are stationary. Table 9 reports the results of the Philips-Perron and the Augmented Dickey-Fuller (ADF) tests. In the case of the GDP *per* capita, the statistics suggest that the series is non-stationary and integrated of order one

(I(1)).[15] In contrast, both tests indicate that the population variable is stationary, i.e., integrated of order zero (I(0)).

Table 9: Results of the Philips-Perron and the Augmented Dickey-Fuller (ADF) root tests.

Variables	**Level**		**First Differences**	
	P-P test	**ADF test**	**P-P test**	**ADF test**
GDPpc	-1.76	-4.21	-5.16***	-6.16***
POP	-10.26***	-7.85***	-	-

*Note: The null hypothesis is that there is a unit root. The tests are based on assuming a constant and a trend. The ADF test is calculated allowing for breaking points. The symbols *, ** and *** mean that the null hypothesis is rejected at respectively 10, 5 and 1 percent level of significance.*

In order to deal with the non-stationary of our series whilst examining the relationship between the two aggregates, we employ the ARDL bounds testing approach. This approach has already been applied to study the long run relationship between population and economic development in the last years (Furuoka, 2009; Savas, 2010).

The ARDL bounds testing approach to cointegration, explained in Pesaran and Pesaran (1997) and Pesaran et al. (2001), is a tool that allows us to explore the existence of a long-run relationship when the variables are not of the same other, as it happens in our case. With small or finite sample data, this approach is more efficient than the ones obtained with traditional cointegration techniques, and provides long-run estimates that are "super-consistent" (Pesaran and Shin, 1999).

The ARDL approach to cointegration involves estimating the unrestricted Error Correction Model (ECM):

$$\Delta GDP_t = \alpha_0 + \sum_{j=1}^{p-1} \alpha_j \cdot \Delta GDP_{t-j} + \sum_{j=0}^{q-1} \beta_j \cdot \Delta POP_{t-j} + \delta \cdot T + \vartheta \cdot GDP_{t-1} + \theta \cdot POP_{t-1} + \varepsilon_t \quad (1)$$

where the symbol Δ represents the first-difference operator, *T* is the trend and δ its associated parameter. α_0 is the constant, ϑ and θ are the long-run parameters whilst α_j and β_{ij} are the short-run parameters and ε_t the disturbance term. The optimal combination of the number of lags (*p* and *q*) is chosen by using the Akaike information criterion (AIC).

[15] The order of integration refers to how many times the series has to be differentiated to become stationary. In other terms, how many times it has to be computed the first difference with respect to the value of the previous year to make the series stationary.

The ARDL bounds testing approach is based on different specifications of the ECM: (i) No intercept and no trend ($\alpha_0 = \delta = 0$), (ii) Intercept and no trend ($\delta = 0$), (iii) Intercept and trend ($\alpha_0 \neq 0$ and $\delta \neq 0$). For each specification, the approach involves estimating two statistical tests. The first one is an F-statistic that tests the joint null hypothesis for a set of parameters. The second is an individual t-statistic that checks the nullity of the parameter associated with the lagged dependent variable. Pesaran *et al.* (2001) proposed critical value bounds for each one of these test and for each scenario. The rules of decision are as it follows:

- If the statistic falls above the respective critical upper bound, then we have evidence of a long-run relationship.
- If the statistic is below the respective critical lower bound, then we cannot confirm the existence of a long-run relationship between variables.
- If the statistic remains between the upper and lower critical bounds, then the inference is inconclusive.

Table 10 shows the statistics and the critical values for all the different specifications defined in Pesaran et al. (2001). Except for one specification (no intercept and no trend), all statistics corresponding to each specification are above the critical upper bound. *This result implies that there exists a long run relationship between the population and economic development for EU28 countries over the period 1960-2010.*

Table 10:: Bounds testing approach to cointegration: GDP *per* capita and Population in EU28.

Specification	Test Value	Lower Bound Critical Value	Upper Bound Critical Value	Long-run Relationship
No intercept and no trend	$F_I = 0.17$	3.15	4.11	**NO**
	$\lvert t_I \rvert$=0.15	1.95	2.60	**NO**
Restricted intercept and no trend	$F_{II} = 8.34$	3.62	4.16	**YES**
Unrestricted intercept and no trend	$F_{III} = 12.18$	4.94	5.73	**YES**
	$\lvert t_{III} \rvert$=4.89	2.86	3.22	**YES**
Unrestricted intercept and restricted trend	$F_{IV} = 7.93$	4.68	5.15	**YES**
Unrestricted intercept and unrestricted trend	$F_V = 11.81$	6.56	7.30	**YES**
	$\lvert t_V \rvert$==-4.79	3.41	3.69	**YES**

Note: Lower and upper critical values for F and t statistics are at 5 percent level of significance. They were taken from Pesaran et al. (2001).

Second step: studying of causal relationships through the Granger causality test.

In the second step, we explore the direction of the relationship between population and GDP *per* capita direction. For that end, we use the Granger causality test proposed by Granger (1969). The procedure is based on the estimation of the following equations:

$$\Delta GDP_t = \alpha_0 + \sum_{j=1}^{p} \alpha_j \cdot \Delta GDP_{t-j} + \sum_{j=1}^{p} \beta_j \cdot \Delta POP_{t-j} + \varepsilon_t \quad (2)$$

$$\Delta POP_t = \alpha_0 + \sum_{j=1}^{p} \beta_j \cdot \Delta POPGDP_{t-j} + \sum_{j=1}^{p} \alpha_j \cdot \Delta GDP_{t-j} + \varepsilon_t \quad (3)$$

where ε_t is assumed to be a white-noise error term. Equation (2) reflects the fact that economic growth (ΔGDP_t) can be expressed as a function of its own past ($\sum_{j=1}^{p} \alpha_j \cdot \Delta GDP_{t-j}$) as well as the past values of population growth ($\sum_{j=1}^{p} \beta_j \cdot \Delta POP_{t-j}$). Equation (3) mirrors the same but for the case of population growth (ΔPOP_t). The rule of decision is that population growth (economic growth) has a casual effect on economic growth (population growth) in the sense of Granger if some of the parameters β_j (α_j) are statistically different from zero.

The AIC was used to determine the optimal lag length (p) of equations (2) and (3), assuming that the maximum lag length is set at 10. The optimal lag length was p=1 for both equations. Table 11 presents the results of the statistical hypothesis testing. According to the results, the null hypothesis that population growth does not (Granger) cause population growth can be rejected at the 10 percent level of statistical significance (p-value=0.08). In the same vein, we can reject the null hypothesis that economic growth does not (Granger) causes population growth (p-value=0.06). Therefore, we have evidence of the existence of a bi-directional relationship between population and economic development.

Table 11: Results of the Granger Causality Test

NULL HYPOTHESIS $\left(H_0 : \beta_1 = \beta_2 = ... = \beta_p = 0\right)$	Optimal Lag Length	F-Statistic	p-value
Population growth does not Granger Cause economic growth	1	1.80	0.08
Economic growth does not Granger Cause population growth	1	1.94	0.06

Note: The lag length is based on the Akaike Information criterion. The residuals show no serial autocorrelation.

Third step: Exploring the sign of the causal relationship.

In the third step, we investigate the sign and magnitude of the causal relationships found in the previous step. Following Pesaran *et al.* (2001), the long-run equations between economic development (population) with population growth (economic development) can be expressed as follows:

$$POP_t = \beta_0 + \beta_1 \cdot GDP_t + E_t^{POP} \quad (4)$$

$$GDP_t = \alpha_0 + \alpha_1 \cdot POP_t + E_t^{GDP} \quad (5)$$

where β_1 (α_1) measures the effect on the long-run of GDP *per* capita (population) on population (GDP *per* capita). Since both variables are expressed in logarithm, these parameters can be interpreted as elasticities. The vector error correction models (ECM) associated with each one of the long-run equations are given by:

$$\Delta POP_t = \gamma_0 + \sum_{i=1}^{p} \gamma_i \cdot \Delta POP_{t-i} + \sum_{i=0}^{p} \mu_i \cdot \Delta GDP_{t-i} + \phi \cdot E_{t-1}^{POP} + \omega_t \quad (6)$$

$$\Delta GDP_t = \theta_0 + \sum_{i=1}^{p} \theta_i \cdot \Delta GDP_{t-i} + \sum_{i=0}^{p} \vartheta_i \cdot \Delta POP_{t-i} + \varphi \cdot E_{t-1}^{GDP} + \omega_t \quad (7)$$

with (6) and (7) representing the short-run dynamics and E_t^{POP} and E_t^{GDP} corresponding to the error correction terms defined as

$E_t^{POP} = POP_t - \hat{\beta}_0 - \hat{\beta}_1 \cdot GDP_t \quad (8)$ and $E_t^{GDP} = GDP_t - \hat{\alpha}_0 - \hat{\alpha}_1 \cdot POP_t \quad (9)$

The estimation of the ECMs allows to explore the short-run impacts as measured by μ_j and ϑ_i. In addition, the parameter ϕ and φ determine the speed of the adjustment back from any deviation from the long-run equilibrium (Frankel et al., 2004). The estimated parameters β_1 and α_1 approximate the magnitudes and signs of the causal long-run bidirectional relationships found between population and economic development, provided that: (i) the estimated coefficients $\hat{\beta}_1$ and $\hat{\alpha}_1$ are statistically different from zero,

(ii) the residuals of the ECMs do no exhibit any problems of autocorrelation, heteroscedasticity or misspecification, and (iii) the estimated coefficient of the lagged error correction terms $\hat{\phi}$ and $\hat{\varphi}$ are negative and statistically significant. These conditions are necessary to corroborate the existence of a long-run relationship that was found earlier with the ARDL bounds testing approach (Kremers et al., 1992; Granger et al., 2000).

Table 12 shows the estimates of the long-run coefficients $\hat{\beta}_1$ and $\hat{\alpha}_1$. Both estimates are statistically significant and have positive signs. This implies that there is a positive bi-directional long-run relationship: population increases economic growth, and vice versa. $\hat{\beta}_1$ is equal to 0.16, meaning that an increase of one percent in GDP *per* capita leads to an increase of 0.16 percent in the level of population. The estimated value for $\hat{\alpha}_1$implies that an increase of 1 percent in population causes an increase by 6.25 percent in GDP *per* capita.

Table 12: Estimation of the long-run parameters

POP $\xrightarrow{\hat{\alpha}_1}$ GDP	GDP $\xrightarrow{\hat{\beta}_1}$ POP
6.25 (0.00)	0.16 (0.00)

Note: p-values are represented in brackets.

Table 13 displays the results of the estimated ECMs when population growth is the dependent variable. The modeling procedure is based on a general-to-specific approach (Hendry, 1995). [16] Starting with Table 13, the short-run dynamic of population growth only depends on its own most recent past (ΔPOP_{t-1}), and not on the economic conditions. The estimated coefficient of the lagged error correction term $\hat{\phi}$ is statistically significant and equal to -0.066. Therefore, it meets the requirements of being negative and statistically significant. This finding corroborates the earlier finding of a long-run relationship using the bounds testing approach. Additionally, $\hat{\phi}$ can be interpreted as an approximation of the speed of adjustment to the long-run equilibrium after a shock. To be more specific, the deviation from the long-run equilibrium induced by a shock is corrected by nearly 7 percent over the following year. In the lower part of Table 13, we show a battery of diagnostic tests. The LM test indicates absence of serial correlation in

[16] In the general-to-specific modeling approach, the maximum number of lags was set at p=10. We assume the stepwise backwards process in which the cutting level of significance was fixed at 10 percent.

the residuals, and the White test and B-P-G tests do not detect any problem of heteroscedasticity. The statistics associated with the Ramsey's RESET suggests that the estimation is not suffering from any misspecification problem. All these findings imply that $\hat{\beta}_1$ can be used as a proxy of the effect of the economic development on population in the long-run.

Table 13: Results of the Error Correction Model when population is the dependent variable.

RELATIONSHIP				
GDP ⟶ POP				
Short-run variables	**Estimated Coefficients**		**t-statistic**	**p-value**
Constant	0.002		3.95	0.00
ΔPOP_{t-1}	0.709		11.19	0.00
E_{t-1}^{POP}	-0.066		-4.84	0.00
Trend	-0.001		-3.56	0.00
DIAGNOSTIC CHECKING				
			Value	**p-value**
	Adjusted-R^2		0.95	-
Autocorrelation	LM test	LM(1)	0.001	0.98
		LM(2)	0.20	0.82
Heteroskedasticity	White Test		1.56	0.20
	B-P-G Test		1.25	0.30
Misspecification	Ramsey's RESET Test		0.60	0.55

Note: LM test stands for the Breusch-Godfrey serial correlation and B-P-G for the Breusch-Pagan-Godfrey test and Ramsey's RESET test.

Table 14 shows the results of the estimated ECM when GDP *per* capita is the dependent variable. The findings are the same as those reported in Table 13. More precisely, the short-run dynamic of GDP only depends on its own past values (ΔGDP_{t-1} and ΔGDP_{t-6}). The estimated coefficient of the lagged error correction term $\hat{\varphi}$ is negative and statistically significant at 10 percent. The speed of adjustment to the long-run dynamic is estimated to be 0.11. The diagnosis tests show neither problem of autocorrelation nor of heteroscedasticity and, according to the results of the Ramsey test, the estimation is not suffering from a misspecification problem. Hence $\hat{\alpha}_1$ should be a reliable estimate of the effect of population on economic development in the long-run.

Table 14: Results of the Error Correction Model when GDP *per* capita is the dependent variable.

RELATIONSHIP POP → GDP			
Short-run Variable	**Estimated Coefficients**	**t-statistic**	**p-value**
ΔGDP_{t-1}	0.46	3.82	0.00
ΔGDP_{t-6}	0.39	3.41	0.00
E_{t-1}^{GDP}	-0.11	-1.74	0.09

DIAGNOSTIC CHECKING				
			Value	**p-value**
	Adjusted-R^2		0.23	-
Autocorrelation	LM test	LM(1)	0.14	0.71
		LM(2)	0.40	0.67
Heteroskedasticity	White Test		1.10	0.36
	B-P-G Test		2.01	0.13
Misspecification	Ramsey Test		0.25	0.80

Note: LM test stands for the Breusch-Godfrey serial correlation and B-P-G for the Breusch-Pagan-Godfrey test and Ramsey's RESET test.

3.4 Conclusion and future research avenues

The relationship between population and economic development has led to an intense theoretical debate about the direction and the sign of such relationship. While according to the Malthusian view, population growth has a negative effect on economic growth, the Revisionist stream argues that population should promote economic development. At the same time, there is no consensus on the influence of economic development on population growth.

Empirical research does not help to shed any conclusive evidence. The mixed results mirror the lack of consensus in the theoretical literature. Some empirical studies conclude in favor of bi-directional relationships whilst others conclude that there is only a unique direction from economic development to population growth, or from population growth to economic development.

In this second part of the report, we briefly revised the main theoretical explanations on the association between population and economic development, and provided a literature review of the most salient and recent articles published on this topic. Additionally, we contribute to the existing empirical literature by analyzing the relationship between population and economic development for the EU28 countries over the period 1960-2010. Our results support the existence of a bidirectional and positive relationship between population and GDP *per* capita.

The findings are still preliminary. Alternative econometric methods, such as the "Pooled Mean Group Estimation of Dynamic" explained in Pesaran et al. (1999), should be considered in the future. We also need to examine whether the bi-directional relationship is robust over different time periods and across groups of EU countries.

4 Discussion and conclusive remarks

The objectives of this report on population dynamics were threefold. The first objective was to offer a literature review of the main drivers of population growth. The second objective was to carry out an empirical analysis aimed at unveiling the determinants of population growth in EU sub-regions (NUTS3 level) over the period 2000-2010. Spatial econometrics is employed to account for spatial dependence among neighbouring regions. The third objective was to review the existing evidence on the long-run relationship between economic and population growth and, then, to proceed with an empirical assessment of the relationship between these two aggregates for Europe over the period 1960-2010. Time-series econometric tools have been used for this analysis. The literature review helped determining which are the main factors affecting population dynamics from the theoretical point whilst the two separates empirical analyses provided new empirical evidence. The main findings of both the litterature reviews and empirical analyses were discussed, along with their implications and future extensions.

In a nutshell, the results of the first empirical analysis confirm that economic conditions play a key role in shaping population growth. Not only do the economic characteristics of the own region matter, but the economic features of neighbouring regions are also found to be important drivers of population dynamics. *Ceteris paribus*, capital or second-tier metro regions have experienced larger population growth than the other EU28 regions between 2000 and 2010, while regional population growth has been found to be negatively affected by the distance to the those capitals or second tiers metro regions. Furthermore, urban regions have experienced lower population growth in comparison to regions with a larger share of rural population. Finally, as expected, ageing and denser regions have experienced less population growth.

Interesting results emerged when we tested for the presence of heterogeneous effects across different categories of regions. Whilst, overall, the empirical results for sub-groups of regions support the previous findings, some interesting differences are worth being outlined. The presence of a coast was not a significant determinant of regional population growth on the full sample, while it turned to be a relevant element when the analysis is limited to the 25 percent poorest regions, the EU13 regions and the regions having experienced a negative population growth over the 2000-2010 period. Educational attainment has a negative effect on the full sample, while it positively affects population growth in the poorest and EU13 regions.

The second empirical analysis supports the existence of a bidirectional and positive relationship between population and economic growth.

Both studies are subject to limitations that we should tackle in the future. In particular, the first analysis will be expanded in order to cover a longer time period. In addition, the robustness of the empirical results to different weighting matrix (e.g. contiguity matrix, distance-based matrices and nearest-neighbour matrices) will be tested and all NUTS3 regions belonging to the same metro region will be aggregated into one single region. Finally, an alternative population density measures will be considered.

The second empirical analysis will also be extended in order to shed some light on the possible presence of heterogeneous effects across EU Member States and over different time periods. This would help assessing whether the evidence of a positive and bidirectional link between population growth and economic is (or is not) consistent across Europe and over time.

References

Barreira A. P., Ramalho J. J. S., Panagopoulos T. and Guimarães, M. H. (2017). Factors Driving the Population Growth and Decline of Portuguese Cities, *Growth and Change*, 48, 4, 853-868.

Baum-Snow N. (2007). Did highways cause suburbanization?, *Quarterly Journal of Economics*, 122, 775-805.

Beeson P. E., DeJong D. N. and Troesken W. (2001). Population growth in U.S. counties, 1840–1990, Regional Science and Urban Economics, 31, 669–699

Boarnet M.G., Chalermpong S., Geho E. (2005). Specification issues in models of population and employment growth, *Papers in Regional Science*, 84, 1, 21–46.

Brander J. A., and Dowrick S. (1994). The role of fertility and population in economic growth, *Journal of Population Economics*, 12, 1, 91–116.

Brueckner M. and H. Schwandt (2015). Income and Population Growth, *The Economic Journal*, 125, 589, 1653–76.

Burchfield M., Overman H., Puga D., and Turner M. (2006). Causes of sprawl: A portrait from space, *The Quarterly Journal of Economics*, 587–633.

Casson M. (2013). The determinants of local population growth: A study of Oxfordshire in the nineteenth century, Explorations in Economic History, 50, 28-45.

Chang T., Chu H., Deale F., Gupta R. and S. M. Miller (2017). The Relationship between Population Growth and Standard-of-Living Growth Over 1870-2013: Evidence from a Bootstrapped Panel Granger Causality Test, *Empirica-Journal of Applied Economics and Economic Policy*, 44, 1, 175-201

Cheshire P. C. and Magrini S. (2006). Population growth in European cities: Weather matters but only nationally, *Regional Studies*, 40, 23–37

Chi G. and Marcouiller D.W. (2013). Natural amenities and their effects on migration along the urban-rural continuum, *Annals of Regional Science*, 50, 3, 861-883.

Chi G. and Ventura S. J. (2011a). An integrated framework of population change: Influential factors, spatial dynamics, and temporal variation. Growth and Change, 42(4), 549–570.

Chi G. and Ventura S. J. (2011b). Population Change and Its Driving Factors in Rural, Suburban, and Urban Areas of Wisconsin, USA, 1970–2000. International Journal of Population Research, 2011, 1–14.

Clark D. E. and Murphy C. A. (1996). Countywide Employment and Population Growth: An Analysis of the 1980s, Journal of Regional Science, 36, 235–256.

Cullen J. B. and Levitt S. D. (1999). Crime, Urban Flight, and the Consequences for Cities, *Review of Economics and Statistics*, 81, 159–69.

da Silva D. F. C., Elhorst, J. P., and da Mota Silveira Neto, R. (2017). Urban and rural population growth in a spatial panel of municipalities. Regional Studies, 51, 6, 894–908.

da Mata D., Deichmann U., Henderson J. V., Lall S. V. and Wang H. G. (2007). Determinants of city growth in Brazil, *Journal of Urban Economics*, 62, 252–272.

Dall'Erba S. and Le Gallo J. (2008). Regional convergence and the impact of structural funds over 1989–1999: a spatial econometric analysis. *Papers in Regional Science*, 87, 219–244.

Darrat A. F. and Al-Yousif Y. K. (1999). On the Long-Run Relationship between Population and Economic Growth: Some Time Series Evidence for Developing Countries, *Eastern Economic Journal*, 25(3), 301-313.

Dawson P. J. and Richard Tiffin R. (1998). Is there a long-run relationship between population growth and living standards? The case of India, *Journal of Development Studies*, 34, 5, 149-156.

Duranton G (2016) Determinants of city growth in Colombia, Papers in Regional Science, 95, 1, 101-132.

Duranton G. and Turner M. A. (2012). Urban growth and transportation, *Review of Economic Studies*, 79, 1407–1440

European Commission (2017a), My Region, My Europe, Our Future. Seventh Report on economic, social and territorial cohesion, Publications Office of the European Union: Luxembourg.

European Commission (2017b) ‘White paper on the Future of Europe Reflections and scenarios for the EU‑27 by 2025’, COM(2017) 2025 of 1 March 2017.

European Commission (2017c) ‘Reflection paper on the future of EU finances’, COM(2017) 358 of 28 June 2017.

European Commission (2017d), Competitiveness in low-income and low-growth regions — The lagging regions report, Commission Staff Working Document (2017) 132 final of 10 April 2017.

Eurostat (2011) Eurostat Regional Yearbook 2011. Eurostat Statistical Books. Eurostat, European Commission, Luxembourg: Publications Office of the European Union.

Ellen I. G. and O’Regan K. (2010). Crime and urban flight revisited: The effect of the 1990s drop in crime on cities, *Journal of Urban Economics*, 68, 3, 247–259.

Faria J. R., León-Ledesma M. A. and Sachsida, A. (2006). Population and income: Is there a puzzle?, *Journal of Development Studies*, 42, 6, 909–917.

Frankel J., Schmukler S. L. and Servén L. (2004). Global transmission of interest rates: monetary independence and currency regime, *Journal of International Money and Finance*, 23, 701-733.

Fukuda (2012) Population growth and local public finance in Japanese cities, *Applied Economics*, 44, 15, 1941-1949.

Furuoka F. (2009). Population Growth and Economic Development: New Empirical Evidence from Thailand, *Economics Bulletin*, 29, 1, 1–13.

Furuoka F. (2013). Population and economic development in Indonesia: A revisit with new data and methods, *Acta Oeconomica*, 63, 4, 451-467.

Garcia-López, M. A., Holl, A. and 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.

Garza-Rodriguez J., Andrade-Velasco C. I., Martinez-Silva K. D., Renteria-Rodriguez F. D. and Vallejo-Castillo P. A. (2016). The relationship between population growth and economic growth in Mexico, Economics Bulletin, 36, 1, 97-107.

Gibbons, Stephen and Henry G. Overman (2012). Mostly Pointless Spatial Econometrics?, *Journal of Regional Science*, 52, 2, 172–191.

Glaeser E. L. and Saiz A. (2003). The Rise of the Skilled City." NBER Working, Paper No. 10191.

Glaeser E.L. and J.M. Shapiro (2003). Urban Growth in the 1990s: Is City Living Back?, Journal of Regional Science, 43, 1, 139—165.

Glaeser E. L., Kallal H., Scheinkman J. A., Shleifer A. (1992). Growth in cities, *Journal of Political Economy*, 100, 1126–1152.

Glaeser E. L., Scheinkman J. A., Shleifer A. (1995). Economic growth in a cross-section of cities, *Journal of Monetary Economics*, 36, 117–143

Glaeser E. L., Kolko J., Saiz A. (2001). Consumer city, *Journal of Economic Geography*, 1, 27–50

Glaeser E. L. and Saiz A. (2004). The rise of the skilled city, *Brookings-Wharton Papers on Urban Affairs, 1, 4*7–105

Glaeser E. L. and Kohlhase J. (2004). Cities, regions and the decline of transport costs, *Papers in Regional Science*, 83, 1, 197-228.

González-Val R. and Lanaspa L. (2016). Patterns in US Urban Growth, 1790–2000, *Regional Studies*, 1-21, 2014

González-Val R. (2015). Cross-sectional growth in US cities from 1990 to 2000, *Journal of Geographical Systems*, 17, 83–106.

Granger, C.W.J. (1969). Investigating causal relations by econometrics models and cross spectral methods, *Econometrica*, 37, 424-438.

Granger, C.W.J., Huang B., Yang, C. (2000). A bivariate causality between stock prices and exchange rates: evidence from recent Asian flu, The Quarterly Review of Economics and Finance, 40, 337–354.

Hasan S. M. (2010). The long-run relationship between population and per capita income growth in China, *Journal of Policy Modeling*, 2010, 32, 3, 355-372.

Hendry D. F. (1995). Dynamic Econometrics: Advanced Text in Econometrics. Oxford, UK: Oxford University Press.

Huang, T. H. and Xie Z. (2013). Population and economic growth: a simultaneous equation perspective, *Applied Economics*, 45, 3820-3826.

Huang T. L., Orazem P. F. and Wohlgemuth D. (2002). Rural population growth, 1950-1990: The roles of human capital, industry structure, and government policy. *American Journal of Agricultural Economics*, 84, 3, 615–627.

Iammarino S., Rodriguez-Pose A. and Storper, M. (2017). Why Regional Development Matters for Europe's Economic Future, Working Papers 07/2017, Directorate General for Regional and Urban Policy, European Commission.

Jacobs-Crisioni, C., and Koomen, E. (2017). Population growth, accessibility spillovers and persistent borders: Historical growth in West-European municipalities, *Journal of Transport Geography*, 62, 80–91.

Jung W. S. and Quddus M. (1986). Economic Development and Population Growth. International Evidence, *Revista de Análisis Económico*, 1, 1, 33-46.

Kapuria-Foreman, V. (1995). Population a nd Growth Causality in Developing Countries, *Journal of Developing Areas*, 29, 531-540.

Kelley A. C. and Schmidt R. M. (1994). Population and Income Change: Recent Evidence, World Bank Discussion Paper 249. Washington, DC: World Bank.

Kelley A.C. (1988). Economic Consequences of Population Change in the Third World, *Journal of Economic Literature*, 27, 1685-1728

Kremers, J. J, Ericson, N. R. and Dolado, J. J. (1992). The power of cointegration tests, *Oxford Bulletin of Economics and Statistics*, 54, 325–47

Lee B. S., Chun S. E. and Kim S. Y. (2007). The effects of regional characteristics on population growth in Korean cities, counties and wards, *Journal of Asian Economics*, 18, 3, 490-508.

Lehmijoki U. and Palokangas T. (2006). Political instability, gender discrimination, and population growth in developing countries. Journal of Population Economics, 19, 2, 431–446.

Lewis W. and Stanley B. (2016). An Integrated Spatial Model of Population Change in South. *The Review of Regional Studies*, 46, 127–142.

Lutz J. M. (2001). Determinants of population growth in urban centres in the Republic of Ireland. Urban Studies, 38(8), 1329–1340.

Lutz W. and Quiang R. (2002). Determinants of human population growth, Philosophical Transactions of the Royal Society B, 357, 1197–1210.

Martori J. C., Apparicio P. and Ngui A. N. (2016). Understanding Immigrant Population Growth Within Urban Areas: A Spatial Econometric Approach, *Journal of International Migration and Integration*, 17, 1, 215–234.

Nakibullah, A. (1998). Population Growth and Development: the case of Bangladesh, *Applied Economic Letters*, 5, 231-234.

Oueslati W., Alvanides S. and Garrod G., (2015). Determinants of urban sprawl in European cities, *Urban Studies*, 52, 1594–1614.

Partridge M., Bollman R. D., Olfert, M. R. and Alasia A. (2007). Riding the wave of urban growth in the countryside: spread, backwash, or stagnation?, *Land Economics*, 83, 2, 128–152.

Partridge M. D., Rickman D. S., Ali K. and Olfert M. R. (2008). Lost in space: Population dynamics in the American hinterlands and small cities, *Journal of Economic Geography*, 8, 727–757

Pegou-Sibe, J., Chiatchoua, C. and Megne M. N. (2016). The Long Run Relationship between Population Growth and Economic Growth: a Panel Data Analysis of 30 of the most Populated Countries of the World, *Análisis Económico*, 31, 77, 205-218.

Pesaran M. H., Shin Y. and Smith R. P. (1999). Pooled Mean Group Estimation of Dynamic Heterogeneous Panels, *Journal of the American Statistical Association*, 94, 446, 621-634

Pesaran, M. H. and Pesaran, B. (1997). Working with Microfit 4.0: Interactive Econometric Analysis, Oxford: Oxford University Press.

Pesaran, M. H. and Shin, Y. (1999), "An Autoregressive Distributed Lag Modelling Approach to Cointegration Analysis", in S. Strom and P. Diamond (Eds.), Econometrics and Economic Theory in the 20th Century: The Ragnar Frisch Centennial Symposium, Cambridge: Cambridge University Press.

Pesaran, M., Shin, Y. and Smith, R. (2001). Bounds testing approaches to the analysis of level relationships, *Journal of Applied Econometrics*, 16, 289-326.

Pirotte A. and Madre J. L. (2011). Determinants of urban sprawl in France: An analysis using a hierarchical Bayes approach on panel data, *Urban Studies*, 48, 13, 2865–2886.

Rappaport J. (2007). Moving to nice weather, *Regional Science and Urban Economics*, 37, 375–398.

Rees P., Van der Gaag N., de Beer J. and Heins, F. (2012). European regional populations: current trends, future pathways, and policy options. *European Journal of Population/Revue Européenne de Démographie*, 28, 4, 385-416.

Rickman, D. S., and Wang, H. (2017). US regional population growth 2000–2010: Natural amenities or urban agglomeration?, *Papers in Regional Science*, 96, 69–90.

Savaş, B. (2008). The Relationship between Population and Economic Growth: Empirical Evidence From the Central Asian Economies, *Journal of Central Asian & Caucasian Studies*, 3(6), 161-183.

Sohn J. (2012). Does city location determine urban population growth? the case of small and medium cities in Korea, *Tijdschrift Voor Economische En Sociale Geografie*, 103, 3, 276–292.

Thirlwall A. P. (1994). Growth and Development, Basingstoke: Macmillan Press.

Thornton J. (2001). Population Growth and Economic Growth: Long-Run Evidence from Latin America, *Southern Economic Journal*, 68, 2, 464–468.

Trendle B. (2009). The determinants of population and employment growth in small Queensland regions, Economics Analysis & Policy, 39, 2, 295–307.

Vega, S. H., and Elhorst, J. P. (2015). The SLX model, Journal of Regional Science, 55, 3, 339–363.

Veneri P. and Ruiz V. (2016). Urban-to-rural population growth linkages: Evidence from OECD TL3 regions, Journal of Regional Science, 56, 1, 3–24.

Winters J. V. (2011). Human Capital and Population Growth in Nonmetropolitan U.S. Counties: The Importance of College Student Migration, *Economic Development Quarterly*, 25, 4, 353–365.

Appendix A – Summary of the papers (in alphabetical order)

- **Barreira et al. (2017)** focus on Portuguese cities dynamics in the period 1991-2011. The empirical analysis uses as covariates several factors such as demographic, employment, housing, and climate variables. More precisely: employment shares in primary secondary and tertiary sectors, the maximum temperature, the housing stock - distinguishing for the age of houses, vacant houses – distinguishing for the proportion of old, middle and new houses vacant, unemployment rate and the share of population in working age. The estimated model is a random effects panel model, with population growth as dependent variable in growth rate, while all economic and housing variables in shares, taken at the beginning of each 10-year period (1991 and 2001). Their findings suggest that the main drivers explaining city population growth are the variables: unemployment rate, shares of employment in the secondary and tertiary sectors, middle-aged vacant houses and maximum temperature.

- **Baum-Snow (2011)** demonstrates that the construction of new limited access highways has contributed to accelerate the process of suburbanization in US metropolitan areas, declining the population of the central city, over the period 1950-1990. Highways played a relevant role in the spatial distribution of the US metropolitan areas over the considered period. The author bases the analysis on cross section regressions, panel data and on the use of instrumental variables.

- **Beeson et al (2001)** use county level census data to determine the main factors that explain location and growth of the US population. The authors analyze the effect of natural and produced characteristics of the county in 1840 to explain the population growth over the period 1840-1990. The results show that access to transportation networks, either natural (access to ocean) or produced (railroad), was an important source of growth over the period. Other factors were found to exert a positive impact: industry mix (share of employment in commerce and manufacturing), educational infrastructures and good weather. On the other hand, mountainous terrain has had a negative influence.

- **Billari and Kohler (2004)** present some descriptive analysis of the correlations between fertility and key determinants of fertility decisions such as the timing of parental home leaving, marriage and divorce patterns as well as the participation of women on the labour market. The analysis is cross-country and correlates national fertility rates with its determinants at two different periods, i.e. during respectively the mid-seventies and late-nineties. Billari and Kohler (2004) documents that in Europe, marriage was positively associated with fertility in 1975 but then in 1999 the association turns to be negative and not significantly different from zero. Similarly, divorce and fertility were negatively correlated in the seventies but then the relationship became not significant as of 1999. The study also documents the time changing cross-country correlation between fertility and women's labor force participation. This correlation was negative in the mid-seventies and then turns positive in the mid-nineties. Finally the paper compares fertility patterns across cohorts and discuss how late union formation or first birth child bearing has changed across cohorts and could explain the decrease fertility observed over time in some countries. This paper is based on cross-country correlations and as such do not control for other country characteristics. This is however an interesting studies to have a broad overview of the determinants of fertility in Europe.

- **Bloom et al. (2009)** examine the effect of fertility on women labour market participation across a sample of 97 countries and over a 40 years period spanning from 1960 to 2000. In order to account for the potential reverse causality, fertility is instrumented with changes in national abortion legislation over time. The underlying assumptions are that abortion is a method to end pregnancy and that change over time in the national legislation on abortion is exogenous, i.e. not correlated with other social factors that also correlate with female labour market participation. The first stage estimates suggest that moving from very restrictive legal constraints on abortion to less restrictive ones with abortion being available on demand, reduce the total fertility rate by 0.4 children per women. The second-stage estimate shows that higher fertility impacts negatively on women labour market participation, with each birth reducing the total labour supply by around 1.9 year per woman.

- **Boarnet et al. (2005)** examine the sensitivity of the population-employment growth models in a spatial econometric framework when different weight matrices are assumed. They pay special attention to the dynamic stability of the estimated lag parameters of the model. The empirical study focuses on the employment and population growth of the county of Orange (California, US), from 1980 to 1990. The main results suggest that changes in employment in surrounding labor market areas has positive effect on population growth, but it is not robust to different weight matrices. More robust is the negative effect of the initial level of population (i.e., the most populated counties in 1980 had a lower rate of growth). Additionally, it was also robust the finding that counties with larger proportion of black and Hispanic population in 1980n had higher rates of growth. Land use variables such as single family residential, multi-family residential and agriculture showed a robust positive effect on population growth.

- **Brander and Dowrick (1994)** aim at investigating the effects of population growth and on fertility rates on and economic growth. They focus on a panel of 107 countries in the period 1960-1985 and differentiate countries in sub-samples depending on their level of development. They use an econometric OLS model corrected for heteroscedasticity, as well as a panel fixed effect and a panel random effect model. Lastly they adopt a IV approach to correct for the potential endogeneity of birth rates. The main finding is that fertility, namely birth rates variation, negatively affects economic growth, namely per capita income growth.

- **Brueckner and Schwandt (2015)** provide a macroeconomic analysis aimed at assessing whether populations grow as countries become richer. They analyse a panel of 139 countries in the period 1960–2007, and found, through instrumental variable estimator, that a 1 percentage point increase in GDP per capita growth over a 10-year period increases countries' population growth by around 0.1 percentage points.

- **Burchfield et al (2006)** study urban sprawling development across US municipal areas between 1976 and 1992. In the first part of the paper, comparing high altitude photographs from 1976 with satellite images from 1992, the authors can examine residential and commercial land development over this period. The authors find that residential development do not tend to be more scattered in the nineties than in was in the mid-sixties whilst commercial development patterns vary substantially across metropolitan areas. In the second part of the paper, the authors examine the determinants of sprawl and more specifically test the implications of the monocentric model in terms of urban development intensity. They find that sprawl, as measured by the average value of the sprawl index over the 1976-1992 period, is positively related to employment dispersion (as opposed to centralized employment), the importance of cars relatively to public transports, land characteristics, in particular the value of undeveloped plots of lands. On the other hand, they do not report any effect of consumer amenities on urban sprawl. When it comes to geographical variables, mountains are preventing sprawls while rugged terrain and temperate climate have the opposite effect.

- **Chang et al. (2017)** study the relationship between annual population growth and annual per capita GDP growth for 21 countries over the 1870-2013 period. The purpose of the paper is to identify the direction of the causal relationship between the two aggregates. The authors estimate a system of equations which allows for dependencies across countries but also for country heterogeneity in the relationship between standard-of-living growth and population growth countries. Using the panel causality test approach, the authors show that there are country and time variations in the direction of the causal relationship between population and GDP per capita growth. For instance, the authors report a one-way granger causality from population growth to GDP per capita in Finland or France, whilst Germany or Canada show a one-way causality running from GDP per capital growth to population growth and in Austria and Italia there seems to be a two-way causality. Because of the long time series, the authors could not include any additional control variables, beyond lag values of the both aggregate.

- **Cheshire and Magrini (2006)** study the main determinants of population growth in EU-12. The analysis is based on cross-sectional ordinary least squares model and on a spatial lag model that accounts for spatial dependences. The sample is composed by 121 Functional Urban Regions (FUR), and the period under analysis goes from 1980 to 2000. The authors consider as determinants of population growth a set of variables that capture the industrial structure of the FUR, the gain of the European integration, geographical position and weather variables. The most relevant finding of this study is that it provides strong evidence on the importance of weather conditions to explain differences in the dynamic of population growth.

- **Chi and Marcouillier (2013)** study the effect of natural amenities in affecting regional population growth. More precisely, the authors test if the natural amenity impacts vary with the characteristics of the region, i.e. whether the region is urban, suburban, peri urban or rural. The underlying assumption is that regional areas have specific characteristics that make the natural amenities more or less important for migration decisions. Urban areas offer other cultural opportunities that might be affect more than the natural amenities. On the contrary, migrants to rural areas likely put more weights on natural amenities. Using data at the "minor civil division" level of the US Lake State of Wisconsin the author examine the determinants of in migration between 1995 and 2000. Natural amenities are measured with a set of variables related to the presence of forest, water, wetlands, public lands, lake/river/coast, view sheds and gold courses. The authors control for a large set of characteristics of the area related to the demographic, accessibility and livability characteristics of the municipalities. These characteristics measured through 32 variables are regrouped with principal component analysis into six main indices. OLS and Spatial lag models were employed. In addition, a spatial regime specification was used in order to be able to estimate coefficient and spatial lag effects separately for each type of geographical area. The results suggest that natural amenities have the largest importance on in-migration to rural areas adjacent to metro areas while no effect is observed for migration into urban areas.

- **Chi and Ventura (2011a)** develop an integrated framework for understanding population change. The designed framework is composed of five explanatory components: (i) demographic characteristics, (ii) socioeconomic conditions, (iii) transportation accessibility, (iv) natural amenities and (v) land development. These influential factors of population change conditioned by spatial dynamics and temporal variation. The authors also propose a practical procedure for analysing the complexity between population change and influential factors.

- Chi and Ventura (2011b) applied the integrated framework defined in Chi and Ventura (2011a) to analyse empirically population change from 1970 to 2000 at the municipality level (MCD) in Wisconsin, USA. The authors translate empirically the theoretical framework describe in the former paper. The *demographic characteristics* are captured through the following variables: population density, percentage of young and old population, percentage of blacks and Hispanics. The *socioeconomic conditions* contain unemployment rate, income, percentage of population with high school education, percentage of population with Bachelor's degree, percentage of college population, percentage of housing units using public water, percentage of seasonal housing in the municipality, real estate value, county seat status, percentage of workers in retail industry and percent workers in agriculture industry. The t*ransportation accessibility* is captured through the following variables: proximity to central cities, proximity to airports, proximity to major highways, highway density, and public transportation system. The *natural amenities* group is composed by the variables percentage of forest coverage, percentage of water coverage, percentage of wetland coverage, percentage of public land coverage, lengths of lakeshore/riverbank/ coastline, golf courses, and view shed. Finally, the *land use and development* is captured by the variables water, wetland, slope, tax exempted lands, and built-up lands. The authors use principal factor analysis to reduce data dimensions and to generate composite indicators for the four groups of dimensions (demographic, livability, accessibility and natural amenities), whereas for land use and development a "ModelBuilder" function in ArcGIS has been used to derive the developability index. OLS regression on population change in the separate periods (1970-1980; 1980-1990; 1990-2000) have been estimated and, as spatial dependence was depicted, also spatial lag models for population change in the same periods, as well as spatial lag models for different urban areas (rural sub-urban and urban). The authors show that the factors have varying effects on population change across both space and time and they do confirm a significant effect of spatial spillovers from neighbouring municipalities. Moreover, transportation accessibility is not found to affect population change directly but it acts as a facilitator of population flows. Accessibility itself does not promote population change but it seems to strengthen the spatial lag effects of population change, thus having and indirect effect.

- **Clark and Murphy (1996)** study the determinants of change in population and employment density in US counties between 1989 and 1981.Whilst most of the papers study population and employment growth patterns separately, this paper uses two-stage least squares in order to model the endogenous relationship between employment and population growth. The authors not only include classical controls such as the demographic, economic and climate characteristics of the counties, as well as location controls but also measures of federal and local government spending and taxation variables. The results suggest that population density changes are largely driven by natural amenities. Employment and education are also positively associated with population density variations, whilst the opposite is found for lagged population density and local taxes.

- **Cullen and Levitt (1999)** explore the relationship between population growth and city crime rates for US cities and metropolitan areas employing data over the period 1976-1993. They use correlation, OLS regression, panel data and instrumental variables as techniques of analyses assuming different periodicities. The main results confirms the strong negative effect of crime on urban growth, regardless of the time-scale of the data (i.e., annual, 5 years difference or 10 years difference), the level of aggregation (i.e., city level or household level) or the different control variables that are included in the analysis. The authors discover that almost all of the negative effect crime on city's population growth is due to out-migrants rather than a decrease in new arrivals. Regarding the control variables, the authors find some evidence that the percent of workers employed in manufacturing, a high percentage of high-school graduates, warm temperatures and low precipitations are associated with a faster city population growth. The unemployment in metropolitan areas are linked with a decline in population, and there are mixing results and not robust results for the income and age variables.
- **da Mata et al. (2007)** use a database of 123 agglomerations over the period 1970-2000 to explore the main determinants of Brazilian city growth. The econometric procedure is based on linear models estimated by OLS and GMM-IV. Their empirical findings indicate that the increase in labor force quality, the reduction in intercity-transport costs, the reduction of rural income opportunities and increases in market potential for goods are all of them significant determinants that have positive impacts on population growth. On the other hand, the initial level of population and the local crime and violence are relevant determinants that lower population growth.

- **da Silva et al. (2016)** study the determinants of population growth in Brazilian minimum comparable areas over the period 1970-2010. The determinants of population growth over each decade are function of previous population growth rate and a set of variable capturing the economic dynamism as well as the amenities of the local area. To account for spatial dependences, a general nesting spatial (GNS) specification is employed. This is all the more important that at such a small disaggregated level of analysis, the characteristics of the neighboring municipalities (in terms of pollution air, criminality, etc.) is likely to affect population growth patterns of neighboring municipalities. The empirical results suggest that education, GDP per capita, and the share of the manufacturing sector relative to the service sector are positively associated with population growth. On contrary, the share of employment in agriculture, the agricultural population and the density affect negatively population growth. These results are in line with the assumption that productivity is higher in municipalities with larger potential markets and with an important part of manufacturing activities. Population growth is also substantially dependent upon the characteristics of neighboring municipalities.

- **Darrat and Al-Yousif (1999)** test the intertemporal relationship between population growth and economic development in a sample of twenty developing countries over the period 1950-1996. To that end, they verify that there is a long term equilibrium between both variables by using the Johansen cointegration test. Once they verify the existence of equilibrium, they construct an error-correction model to capture the short- and the long-run dynamics of the relationship. The authors find that population stimulates economic growth in more than half of the countries, which validates the revisionist view. For the case of China, they also find that that economic growth has a negative impact on population growth, indicating that households tend to prefer quality to quantity of children. The latter finding corroborates the demographic transition theory.
- **Dawson and Tiffin (1999)** aim at investigating the long-run relationship between population growth and economic growth, namely measured as GDP pro capita. They focus on India in the period 1950-1993. Interestingly enough, they conclude that

population growth is not causing economic growth and that economic growth does not cause population growth, at least in a developing country as India. This non significant effect is achieved by applying time series co-integration analysis, finding that population in India is trend stationary ad economic growth has one unit root, implying there cannot be a long run relationship between the two. This is found by applying Dickey-Fuller test, Granger causality test and the Johansen procedure.

- **Duranton (2016**) using OLS regressions and the method of instrumental variables, explores which are the main drivers of population growth in Colombian cities for the period 1993-2010. The author also analyzes metropolitan areas and assumes longer periods of time (1938-2010 and 1870-2010) to determine the factors of the long-run growth. The factors analyzed are grouped in the following determinants: (i) demographics (initial population, fertility birth rates and migration), (ii) local labor market (wages and squared wages), (iii) human capital (share of university educated workers and its squared), (iv) labor demand (e.g., industrial diversity and industrial specialization), and (v) urban amenities and cost of living such as climate (e.g., temperature and precipitations), geographic characteristics (e.g., altitude), institutional characteristics (e.g., violence and criminality), natural resources (e.g., extraction of oil or minerals) and road connectivity (road index). The main findings reveal that the key drivers to explain population growth are the demographic factors, wages (included the quadratic form), human capital and industrial structure. The author also finds important the role of some amenities such as roads and connectivity, climate factors (temperature and precipitations), criminality and production of oil.

- **Duranton and Turner (2012)** provide evidence that leads to a better understanding of the role of roads in the growth of US cities between 1983 and 2003. The analysis rely on an instrumental variables estimation to explain employment growth and population growth. The authors structure the study in three parts. In the first one, they develop a simple structural model that describes the joint evolution of highways and employment in the cities. In the second part, they explain the instrumental variables used to identify key parameters of the structural model. They also provide out-of-sample estimates to validate the estimation and the main assumptions of the structural model. Finally, in the last part of the study, they use the estimates to assess hypothetical transportation policies. Their findings provide a basis for estimating the impact of changes in transportation infrastructures. For instance, the main result stressed by the authors is that an increase of one percent in a city's roadway leads to a 1.5 percent increase of employment and 1.3 percent larger population 20 years later.

- **Ellen and O'Regan (2010)** focus specifically on how changes in crime affects population growth in cities. By focusing on 100 US central cities in the period 1980-2000 and using both panel fixed effects and instrumental variables estimations the paper finds a confirmation that lowering crime rates do not attract new residents as a result, but it helps citizens to remain in the cities.

- **Faria et al. (2006)** examine the relationship between population and per-capita income by using a panel of 125 countries over the time period that spans from 1950 to 2000. The authors model the rate of growth of population in function of lags of the log per capita GDP and its squared values. The lagging of these variables is made to control for the possible endogeneity of the explanatory variable. They also include regional dummies in some of the regressions to account for any time invariant effect of the different world regions. They use six different estimators: pooled OLS regressions, the fixed effect estimator, the random effect estimator, the maximimun likelihood estimator and the population averages estimators. The results are not too robust between the different estimation methods considered in the study. Nevertheless, the general result shows that there seems to be an inverted U-shape relation between population growth and per capita income: as income increases, population also increases but at a decreasing rate. At a certain threshold level, the increase in income leads to a decline in the level of population. This finding underlines the existence of a non-linear relationship between population and income, and supports the stylized fact known in Demography as the demographic transition. However, the authors also put a note of cautious since they find differences in the relationship across the different sub-groups of countries and periods.

- **Fukuda (2012)** empirically assesses which are the main determinants of city growth for the period 2000 to 2005 in Japan. The author employs a database of 321 mid-and small size Japanese cities. The analysis is based on estimating different OLS regression models and spatial econometric models (specifically, the author assumes a spatial autoregressive model and a spatial error model). The results show that the number of local government officers and the unemployment rate exerts a negative effect on population growth, while the ratio of graduates of junior colleges, universities and graduate schools, the ratio of sewerage, the monthly average of the lowest daily temperature and the number of interchanges have a significant positive effect. The author also includes four different Fiscal Indicators that measure the flexibility of the financial structure, which is a proxy for better policies. The significance and the sign of the Fiscal Indicators supports that better policies increase the population growth of Japanese cities.

- **Furuoka (2009)** finds out the association between population growth and economic growth for Thailand over the period 1961-2003. The author uses the bounds testing approach to cointegration to investigate if these variables are statistically related in the long-run, and the Granger causality test to check if one variable causes the other. The findings reveal that there is a long-run relationship between economic growth and population growth. Specifically, the Granger causality test allows the author to claim that the growth of population in Thailand had a positive impact on its economic development. However, the reverse causality (i.e., economic growth promotes more population) is not statistically verified.

- **Furuoka (2013)** explores the relationship between population growth and economic development for the case of Indonesia during the period 1960-2007. The author uses the traditional Johansen cointegration test and a non-linear cointegration test (Breitung cointegration rank test) to detect the existence of a long-run equilibrium between the two variables. The Granger causality test indicates that there exists a unidirectional causality effect from Indonesia's population growth to its economic development.

- **García-López et al. (2014)** use data on 77 metropolitan areas in Spain, over the period 1960-2011, to estimate the effect of highways on population growth. The authors employ OLS regression to estimate the impact and, in order to avoid the potential endogeneity problem of highway provision, they also apply the technique of instrumental variables. The results show that highways contributed to a decline in central cities population growth, whilst they benefit the expansion of the population of suburban areas. Moreover, being closer to a highway favors municipal density growth.

- **Garza-Rrodríguez, et al. (2016)** perform cointergration econometric analysis to assess the long run relationship between economic growth and population growth. They build the economic analysis on data for the period 1960-2014 for Mexico. The main result is in contrast with the previously described studies for Latin America and India as in the short run, it was found that economic growth has a negative effect on population growth, in the long run, population has a positive effect on economic growth and economic growth has a positive effect on population growth. Interestingly enough, the article does not explain why the findings diverge from the one by Thornton 2001, who also included Mexico as a country to focus on.
- **Glaeser and Kohlhase (2004)** document the decline in the United States and over the last century of the transportation cost for goods with as consequences a negative association between population growth in counties and the share of employment in the agricultural or the manufacturing sectors. Over the same period, on the contrary, the transportation cost for people remained high and might have even increased because of the traffic congestion. This empirical fact, coupled with the change in the transportation modes (from railways to cars, airplane) have made more likely that cities and, more broadly dense areas, will host services. Given that highly educated people tend to concentrate more in services and might benefit more from interactions, the authors show that densely populated areas are also more likely to have a high share of educated individuals. In addition, whilst the economic characteristics of the areas (natural resources) are of less importance, its natural or cultural amenities are likely to enter into the location decisions of individuals. The empirical evidence is based on scatter plots.

- **Glaeser and Saiz (2004)** investigate the main determinants of population growth using a panel of US metropolitan areas and cities for the period 1970-2000. Even though they include in the analysis different possible determinants such as climate variables and sectoral variables, they mainly focus on the impact of human capital. The results clearly suggest that human capital increases population growth. Other interesting results reveal that warm, dry places grow much faster than cold, wet places. Additionally, cities specialized in the manufacturing industry slow their growth down.

- **Glaeser and Shapiro (2003)** using different OLS specifications, examine the factors that explain the population growth of US cities and metropolitan areas in the decades of the 80s and 90s. They discovered three basic stylized facts in the determination of the rates of population growth. The first one is that car cities grow in comparison with those cities where public transport is important. Second, low density cities grow faster than high density cities. Third, cities with warm and dry weather tend to grow more and, finally, places with more skilled workers grow faster than cities with weaker human capital.

- **Glaeser et al. (1995)** examine the urban growth patterns in US cities and metropolitan areas (standard metropolitan statistical areas) between 1960 and 1990 with respect to various economic, geographic, demographic, educational and social and political characteristics in 1960.The analysis is based on the estimation by OLS of different linear models. They conclude that income and population growth are positively related to initial schooling, and negatively related to initial unemployment and initial share of employment in manufacturing.

- **Glaeser et al. (2001)** discuss the importance of urban density in facilitating consumption. They argue that the future of cities strongly depends on whether cities are attractive for consumers to live. They carry out a simple multivariate regression of county population growth in the US for the period that spans from 1977 to 1995. The findings from this regression show that climate variables (temperate climate and dryness) are strong predictors of local growth. They find that proximity to the coast is also a significant explanatory variable. The presence of live performance venues and restaurants, as proxies for different forms of consumer amenities, were significant to explain population growth at the county level. On the other hand, amenities appealing to less educated workers (bowling alleys and movie theaters) had a significant negative impact. However, the authors do not find a significant association between art museums and population growth.

- **González-Val (2015)** analyzes the growth in 1152 US cities from a cross-sectional approach over the period from 1990 to 2000. The author includes sprawl, human capital, productive structure, physical and socio-economic variables as potential determinants of growth. The analysis is founded on two different approaches: the traditional linear growth model and quantile regression. The possible existence of spatial dependence is considered in both approaches (spatial lag error model and spatial autoregressive lag model for the linear model, and spatial quantile regression for the other case). In general, the findings reveal that cities with specialized economies, denser in terms of population and with initial higher unemployment rates grew less. The weather and human capital variables were also found significant to explain city growth. The author also provides some additional interesting results. The first one is that this study corroborates the persistence of population growth as previous level of growth shows a positive effect (persistence effect). Second, the results reveal the existence of spatial spillover effects, as well as a non-linear behavior.

- **González-Val and Lanaspa (2016)** examine the growth of cities in the United States during the period 1790-2000. By using a panel unit root test analysis, these authors find evidence that US city growth rates are independent of the initial levels of population (i.e., they do not find signs of mean reversion). However, the results obtained from a cluster procedure reveal that there is a convergence in city growth rates within clusters, which means that there can be local mean-reverting behaviors. Finally, the calculation of Markov transition matrices shows high mobility in the distribution of cities.

- **Hasan (2010)** explores both the short-run and the long-run relationships between population and per capita income in mainland China over the period of 1952-1998. The author follows a time series approach where the existence of a long-run equilibrium between the variables is analyzed. In a further step, the author estimates different vector autoregressive (VAR) models. The empirical results show that there is a long run negative relationship flowing from per capita income to population; i.e., growing income lowers population growth. This finding supports the commonly accepted idea that, as income increases, families tend to substitute from quantity to quality of children. At the same time, the author also finds empirical evidence that population exerts a positive long run influence on the levels of per capita income.

- **Huang and Xie (2013)** focus on the link between economic growth and population growth by estimating the simultaneity of this causal relationship. They model simultaneous Autoregressive Distributed Lags (ADL) panel models, to estimate the simultaneous effects of economic growth on population growth and vice versa. The analysis is based on a panel of 90 countries over the period 1980–2007. To overcome identification problems, in a GMM approach they exploit the lagged values of respectively population growth and economic growth as instruments. In modeling the determinants of population growth two additional variables are included: saving rate and the gross secondary school enrolment ratio. Interestingly enough, they find that in the short run population growth negatively affects economic growth, while this effect is not found in the long run. On the opposite direction, economic growth does not seem to substantially determine population growth neither in the short nor in the long run. As for the control variables, saving rates has found to negatively affect population growth but only in middle income countries, while schooling displays no significant effect on population growth (whereas it does affect positively economic growth).

- **Huang et al. (2002)** focus on US rural counties and aimed at assessing whether to establish which factors cause rural counties to grow or decline over time, by focusing the 1950–90 period. The authors follow a strict human capital perspective, assuming human capital stock to be a good predictor of out-migration choices of individuals in working age. The empirical analysis is conducted through Instrumental Variable regression on counties having population growth rate of individuals aged 20-64 as dependent variable. As explanatory variables: the rural median income, the median of school years completed, the percentage of population with high school degree, the distance to a city, and Herfindhal index of Employment, a set of local and state government expenditures (e.g. welfare, education, debt, highway and debt), and a set of demographic controls (proportions of farm, black, young and old), as well as dummy variables for each decade, measures of average county rainfall, January and July temperature, and a dummy variable for Shannon County, South Dakota, which had no county government.

- **Jacobs-Crisioni and Koomen (2017)** study how accessibility affects population growth in municipalities locates in Western mainland Europe in the period 1961-2011. The focus of the study lies in the following countries: Belgium, France, Germany, Italy, Luxembourg, the Netherlands, Switzerland, Portugal, Spain, and Austria. Accessibility is the core variable of the study and it is constructed as domestic accessibility and foreign accessibility. These measures are a function of population in destination municipalities and distance decayed travel times to the destination municipality. Population density changes at the municipality level are the dependent variables. The empirical analysis is based on a fixed effect panel model having a time-invariant coefficient for each municipality. Results find support of a positive and significant role for domestic accessibility on population density growth. However, the study also suggests that national borders negatively affect density growth, thus stressing for the presence of a significant border effect. The determinants of population growth over each decade are function of previous population growth rate and a set of variable capturing the economic dynamism as well as the amenities of the local area. To account for spatial dependences, a general nesting spatial (GNS) specification is employed. This is all the more important that at such a small disaggregated level of analysis, the characteristics of the neighboring municipalities (in terms of pollution air, criminality, etc.) is likely to affect population growth patterns of neighboring municipalities. The empirical results suggest that education, GDP per capita, and the share of the manufacturing sector relative to the service sector are positively associated with population growth. On contrary, the share of employment in agriculture, the agricultural population and the density affect negatively population growth. These results are in line with the assumption that productivity is higher in municipalities with larger potential markets and with an important part of manufacturing activities. Population growth is also substantially dependent upon the characteristics of neighboring municipalities.

- **Jung and Quddus (1986)** is one of the first studies that adopts a time series framework to analyze the relationship between population growth and economic growth. The authors use the Granger causality test to check the existence of a causal link, as well as the direction and the sign of such link, for 44 countries assuming a sample period that ranges from 1952 to 1980 in the best of the cases. The authors do not find a consistent and robust evidence of causality between these two variables. Not many countries of the sample present a significant relationship of causality between these variables; but the sign of the relationship is not clear given that some countries show a positive causal effect while others have a negative effect. In summary, the authors conclude that the causal direction from economic development to population, and vice versa, is ambiguous and not robust.

- **Kapuria-Foreman (1995)** studies the relationship between population growth and economic growth in fifteen low- and middle-income developing countries. The findings so not support the idea that population growth exerts a negative impact on economic growth. The author conducts a Granger causality test to check if population growth causes economic growth or vice versa. The main conclusion pf the study is that the rate of growth of population affects and is affected by the rate growth of per capita income.

- **Lee et al. (2007)** search for the main determinants of population growth in Korean cities during the period from 1980 to 2000. The findings show that the regional education level and the share of manufacturing employment are the main determinants of the Korean cities' population growth. However, the sub-sample analyses reveal that the importance of the manufacturing is declining, while the effect of education remains strong. The authors also find a negative effect on population growth of the distance to the nearest metropolitan area. Finally, their results support the convergence hypothesis that establishes that cities with large population tend experience slower population growth.

- **Lehmijoki and Palokangas (2006)** aims at assessing the interdependence of political instability, population growth, and gender discrimination in the case of developing countries. The article combines a theoretical model with empirical evidence. The first is predicting countries with high risk of internal conflicts to generate high population growth rates. The second is conducted on a sample of 69 low- and middle-income countries with data for the period 1989–1999, upon which OLS estimates are conducted exploiting different variables to proxy for internal instability. Findings support the theoretical predictions: high risk of internal instability can explain high population growth.

- **Lewis and Stanley (2016)** using a spatially lagged explanatory variables (SLX) model, study the population change of counties in South Carolina (US) over the period 1998-2012. In general, they find that local births, income, percentage of black population and poverty slow down population growth. They evidence that population growth in one county depends on local characteristics in neighboring counties (i.e., they verify the existence of significant spatial dependences among counties). Specifically, they find that population declines with increases in income and single female pregnancy rates in surrounding countries. Additionally, they also provide evidences of significant differences between rural and urban counties. Income and black population share have a negative impact both for rural and for urban counties. While retiree shares negatively affect population changes in rural counties, single female pregnancy do so in urban counties. They also find important differences for the effects of surrounding counties. Specifically, they discover that local urban counties are influenced by changes in neighboring rural and urban counties. On the other hand, local rural counties are only affected by changes in surrounding rural counties.

- **Lutz (2001)** aims at unveiling which features of cities and towns in Ireland have been associated to subsequent population growths. The paper focuses on irish urban places, defined as having a population greater than 1500 in the period 1966-1991 and it exploits census data for the years 1966, 1971, 1981, 1986 and 1991. Almost 20 independent variables reflecting socio-economic and geographical characteristics have been included. The proximity to the capital (Dublin) or the closest urban centers (limerick or Cork) has been included. In contrast to similar literature using the distance in kilometers, this has transformed the proximity in a categorical variable being one if the urban place it less than 10 km far from the three above mentioned largest urban places, 2 if it is around 11-20 and so on up to the latest category 6 encompassing the over 160 km from the city. This variable is the one of main interest and it displays the most interesting effect, as population in urban centers is found to be driven by proximity to Dublin. The method is based on simple multivariate regression analysis conducted separately on the growth of population between each sub period in the sample, leading to 4 set of estimations. Overall, results seem to be very much specific on the period and country chosen, failing to look generalizable.

- **Lutz and Qiang (2002)** study the most influencing factors that determine population growth and fertility rates for a set of 187 countries. The authors estimate 14 cross-sectional models in lags of 5 years for the whole sample, as well as for a subset of the 147 less developed countries. They focus on the effect of population density on population growth and fertility, but they also include other socio-economic variables such as the proportion of females in the labor force, the female literacy rate, the proportion of urban population, the GDP per capita and an index of food production. The empirical results reflect that the patterns and even the signs of the estimated coefficients of the influential factors of population growth were not consistent over time, and were statistically insignificant in general. Population density showed a positive and significant effect only for the population growth for the years 1960 and 1965; and GDP per capita and the index of food production were negative and significant for the year 1980. With respect to the most influential factor to explain fertility rates, the authors find that the variable female literacy rate and population density seem to be the most influential factor. In general, both factors have a negative and significant effect on fertility rates over the different considered periods.

- **Martori et al. (2014)** study the determinants of no EU immigrant population growth within the Barcelona metropolitan area during the 2001-2008 periods. The analysis is carried out at the census tract level and spatial lag as well as a spatial error models are employed to account for spatial effects. The labor market conditions, the presence and diversity of the already existing immigration population, the housing market, the population density and the distance to the business center have been found in the literature to be associated with immigrant population growth, The empirical results for the Barcelona metropolitan area suggest that non EU-Immigrants tend to localize in areas with a large mixed of immigrants and low income, far from the business center. The authors also use the generalized methods of moments with the spatial lag and the spatial errors model and report similar findings. Other similar studies covering different metropolitan areas could be considered for the literature review (Arapoglou, 2012, Musterd and Deurlo, 2002, Zorlu and Mulder, 2008, Aslund, 2005).Moreover, transportation accessibility is not found to affect population change directly but it acts as a facilitator of population flows. Accessibility itself does not promote population change but it seems to strengthen the spatial lag effects of population change, thus having and indirect effect.

- **Nakibullah (2010)** investigates if the dynamic of the population growth in Bangladesh is endogenous (i.e., if it depends on economic development). Using the Granger causality test, the findings clearly indicate that the real GDP per capita Granger cause population growth in Bangladesh over the period object of study (i.e., from 1959 to 1990). This means that population growth is endogenous in the development process of Bangladesh. Based on this finding, the author recommend implementing political measures tend to increase human capital because of these measures raise growth per capita and reduce fertility rates.

- **Osili and Long (2008)** investigate the causal relationship between education and fertility. Specifically, they study the impact of the large-scale and nationwide program called Universal Primary Education that was implemented in Nigeria in 1976. The authors used data of the 1999 Nigerian Demographic Health Survey and applied two different econometric approaches: instrumental variables and differences-in-differences. The results of the study indicate that the implemented education program had a significant impact on both female education and fertility decisions. To be more precise, there was an increase in the levels of female education that caused a reduction in fertility. This empirical evidence supports that female schooling increases the opportunity cost of child-bearing and the knowledge about contraceptive methods. Finally, female schooling reduces the rates of child mortality.

- **Oueslati et al. (2015)** explore the characteristics of urban cities associated with urban sprawl in Europe. The authors build on the monocentric model and rely on an Hauman-Tylor instrumental variable approach in a panel framework to account for time varying unobservable factors. The urban audit dataset for the time periods 1990. 2000 and 2006 is combined with data on land use for a sample of 237 European cities. In line with the monocentric model, the authors report a positive association between urban sprawl, economic development (GDP per capita) and population, whilst the cost of agricultural lands and transport cost decrease urban sprawl. Furthermore, urban sprawls are more likely in cites which are well connected to the rest of the world through airport and when air pollution is high. On the contrary, cultural amenities in cities affect negatively on sprawl. The authors also show that income growth is the most important reason of urban expansion over the period studied.

- **Partridge et al. (2007)** study employment growth dynamics of US non Metropolitan County over the period 1990-2004. The purpose of the paper is to examine if there are spatial heterogeneities in the determinants of employment growth. They control for the initial period population density as well as for a large set of natural amenities (climate and topographic variables), economic (unemployment rate, household income, agriculture share, industry mix employment growth rate), demographic (population age shares, population shares by educational level and ethnicity, immigration flows, and distance (5 metrics of distance relative to urban centers) related characteristics of the counties. The authors compare the empirical results obtained with a classical OLS and the ones derived from the use of geographically weighted regressions. The authors argue that some determinants might have opposite effects across counties but average to zero on the full sample. The empirical results confirm that natural amenities and human capital related variables exhibit spatial heterogeneity.

- **Partridge et al. (2008)** examine the effect of proximity to urban agglomerations on population growth in hinterlands American counties and small urban areas over the period 1950-2000. The analysis is also performed over sub-periods to test the stability of the results over time and for different samples of counties, the samples differing according to the definition of hinterlands and small urban areas considered. GMM-based cross-sectional estimations are employed. The aim of the study is to measure the effects of distance to the nearest urban area and to higher tiers of the urban hierarchy. The underlying assumption is that the type of services, goods as well as the labor market potential differ with the type of urban areas, with the top tier cities offering the most whilst the lowest tier only provide basic services. This suggests that the effect of distance could be specific to the urban hierarchy. After controlling for characteristics of the counties (climate, indicators for being close to the sea and great lakes) state-fixed effects and a measure of potential market, the study finds that distance to the nearest urban agglomeration is still inversely related to population changes in rural counties. Furthermore, there is an additional penalty associated with the incremental distances to reach the next higher tiered metro areas. The effect of distance also seems to have become more important after 1970.

- **Pegou-Sibe et al. (2016)** analyze the relationship between population growth and per capita income for 30 of the most populated countries over 1960 to 2013. The findings support a long-run relationship between the two variables, indicating that there is significant positive relationship. The causality test reflects that there is a bi-directional causal effect.

- **Pirotte and Madre (2011)** examine the determinants of urban sprawl in France at the municipal level for four metropolitan areas over a time period spanning from 1985 until 1998. The time period covered allows examining if sprawl determinants differ during periods of rapid economic growth with respect to periods of recession and recovery. Separate estimates for the four metropolitan areas (and within each metropolitan area for conurbation and outer suburbs) are carried out and random coefficient models are employed. The authors report a positive association between urban sprawl and income growth as well as with socio-economic segregation and population density. The income effect is larger for period of rapid economic growth than during recovery. The authors show that the factors influencing urban sprawl varies across metropolitan areas and over time.

- **Rappaport (2007)** examines the effect of weather on the US population growth using county-level census from 1970 to 2000. To do so, the author regresses the population growth on different weather variables, as well as on their quadratic values. After controlling for the impact of other determinants such as the county sectorial structure, the coastal proximity, topographical variables, initial population density and surrounding population, the author's findings strongly support the idea that US residents have been moving to places with nice weather. In fact, the influence of the weather factors is extremely robust to using alternative weather measures and different techniques of estimation (i.e., OLS and GLS).

- **Rickman and Wang (2017)** focus on US counties in years 2000-2010 and propose a Spatial equilibrium model, more precisely a Spatial hedonic growth model, to assess the role of natural amenities and urban agglomeration economies on population growth. Their results suggests the presence of a different effects of household amenity demand, natural amenity, attractiveness of cities, elasticity of land supply and productivity according to whether the area is rural or urban. For metropolitan areas the authors indeed find support that the attractiveness of the city drives population growth more than economic productivity growth. This in other terms implies that "rather than agglomeration economies producing jobs that cause in-migration, it is the attractiveness of cities to households that appeared to spur population growth" (page 89).

- **Savas (2008)** analyzes the relationship between population and economic development in five countries of Central Asia (i.e., Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan and Uzbekistan) between the years 1989 and 2007. The author uses the ARDL bounds testing approach to cointegration to find a long-run equilibrium between population and economic development, and the Granger causality test to determine the direction of the relationship. The researcher finds that these countries are in the demographic phase characterized by high birth rates and low death rates (i.e., in the second phase of the demographic transition called "post-Malthusian regime"). At this stage, it seems that the relationship between the two variables is significantly positive. Specifically, there seems to be a bi-directional causality between the two variables for the countries over the period under analysis.

- **Singh (1994)** using cross-country models estimated by ordinary least squares, explores the effect that investment in the education of women, the greater participation of women in the labor force, the access to health services and the use of birth control measures services have on fertility and child mortality. The analysis is focused for the less developed countries, and for the decade of the 1980's. The author finds that women's education, women's labor force participation and the use of contraceptive measures significantly reduce fertility rates, slowing down population growth. For the case of child mortality, the results reveal that women's education, women's labor participation and access to health services reduce children´s mortality. All these results lead the author to recommend not reducing public spending in education and health services in the less developed countries.

- **Skirbekk et al. (2014)** compare population growth patterns observed during the transition period in Europe and the rest of the word. Europe has been characterized by lower population growth than in other geographical areas. The authors discuss the main reasons behind those different growth patterns, namely the spread of education, cultural aspects such as late marriage and minimal births out of wedlock, the effect of industrial revolution on the cost of children and European specific events such as the black death and its effect on the participation of women on the labor market. The authors also simulate what would have been the population growth in Asa, Africa, Latin America or Australia and New Zealand, had the population growth followed the same pattern of the one observed in Europe. Clearly the population growth observed in these regional would have been much lower under this counterfactual scenario. This suggests that population changes are largely driven by specific cultural or economic factors and not that much by universal population growth trajectories. The authors conclude that rural areas suffer from "brain drain", as increases in the education levels of rural areas lead to a reduction in rural population growth. Moreover, the economic returns of educational attainments are higher in urban than in rural labour markets.

Furthermore, rural economies with a more diversified economy have higher population growth rates than rural are strongly specialized.

- **Thornton (2001)** following the same methodology of Dawson and Tiffin (1999), focuses on the long-run relationship between economic growth and population growth in seven Latin American countries. Different time period are covered depended on the country specific data availability (e.g. Argentina, Brazil, Chile and Venezuela cover the period 1990-1994, Colombia 1925-1994, Mexico 1921-1994 and Peru 1913-1994. Interestingly enough their study find the same absence of a causality already described for the India case, with population trend being stationary while GDP per capita having one unit root, leading them to conclude that population growth neither causes nor is caused by GDP per capita growth.

- **Trendle (2009)** investigates the main determinants of population and employment changes for the 125 Local Government Areas of Queensland (Australia) over the period 1996-2006. The author also studies the existence of bi-directional causality between population and employment. The analysis is based on two-stage least square regressions in which the author also takes into account spatial dependences by estimating a spatial autoregressive lag model. The results show that there is some evidence of spatial dependence, as well as simultaneity (i.e., employment changes cause population changes and vice versa). The author also finds that higher levels of education, employment and social advantages (approximated by an index) have a positive and significant influence on population changes. Conversely, higher initial levels of population leads to a reduction in population growth. A surprising finding is the estimated effect for the variable income. A priori, the author expected a positive influence of the variable income since high-income areas would attract more in-migrants. However, a negative and significant effect is estimated for this variable. One possible explanation to this finding provided by the author is that areas with higher incomes also have higher housing prices (higher living costs in general), which disincentive in-migration.

- **Veneri and Ruiz (2016)** study how population growth in urban areas is affected by their near urban centers, drawing on the evidence that rural regions that are located far from important populated centers are depopulating and economic declining. By applying a semiparametric approach – namely a penalized spline geo-additive model and by focusing on OECD small regions (TL3) in the period 2000-2008 they do find rural regions to benefit from growth in urban places. All in all the presence of positive population growth spillovers from urban or intermediate regions to rural ones is confirmed, and it is found to be stronger for EU than for US. Interestingly enough they do also find that this effect declines with distance, measured as the km distance of the closest urban region from the centroid of the urban region to the centroid of the other region.

- **Wei et al. (2015)** find out which the main determinant factors that explain population growth in China over the period 1949-2012. They estimate a time series regression model in which population growth is explained by the following factors: sex ratio, employment rate, fiscal expenditures degree of urbanization and a dummy to collect natural disasters. The results show that only the variables employment rate, with a positive effect, and fiscal expenditures, with a negative effect, are statistically significant to explain population growth in China over the period considered in the study.

- **Winters (2011)** studies the importance of student migration in the relationship between human capital and population growth for nonmetropolitan counties in the United States over the period 1995-2000. Many students who move to an area for higher education might decide to stay in the area after their schooling is complete because of networks with local employers established during the studies, friendships or a taste for the local amenities. The ability to retain ex-students causes both population growth and an increase in the average level of human capital. To test this assumption, the authors estimate in/out/net migration growth equations with one of the explanatory variables being the share of the adult population with a college degree in the county. These migration equations are also estimated separately by type of migrants, i.e. higher education students and non-students. The author controls for the demographic, economic characteristics of the nonmetropolitan counties. Distance related variables and several variables measuring the natural amenities of the counties are also included as explanatory variables. The author employs spatial error models. The results suggest that the share of adults with a college education is positively associated with in-migration of higher students, whilst such an effect cannot be observed for in-migration of non-students. This effect is even larger when the sample is limited to college town counties.

Appendix B – Tables with the reviewed papers (in alphabetical order)

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Barreira et al., (2017)	Cities	1991-2011	Portugal	· Employment shares in primary secondary and tertiary sectors · Maximum temperature · Housing stock- age of houses · Vacant houses – proportion of old, middle and new houses vacant · Unemployment rate · Share of population in working age	Determinants of cities population changes	Random effects panel model. Dependent variable in growth rate, all economic and housing variables in shares, taken at the beginning of each 10-year period (1991 and 2001).
Baum-Snow (2007)	Census Defined Places	1950-1990	USA	· Highways (number of rays, 1990 rays times fraction of ray miles). · Income · Income Gini coefficient · Population of the metropolitan area	Impact of highways on central cities population. Analysis of the degree of suburbanization.	Instrumental variables, cross-section OLS regression, Panel Data regression.

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Beeson et al. (2001)	Counties	1840-1990	USA	- Natural Factors in 1840:Climate (precipitation, heating-degree days, cooling-degree days). Mineral resources (dummies for the production coal, iron and other minerals). Access to natural transportation networks (ocean, rivers and lakes). Mountainous terrain Land area - Man Produced Factors in 1840:Industry mix (employment shares in different sectors). Educational infrastructures (library, books per capita, literacy rate, college). Access to man-made transportation system (railroads and canals).	Explanation of the main drivers of population growth.	Linear regression model estimated by OLS.
Billari and Kohler (2004)	Country	Two data points: 1975 and 1996 (or 1999)	Varying number of European countries	No control variables	Correlation between fertility and fertility - related behaviors	Analysis based on bivariate statistics
Bloom et al. (2009)	Country	1960-2000	97 countries	- Fertility rates. - Percentage of urban population. - capital stock/working age population - Infant mortality rates. - Male and female education.	Effect of fertility on women labor market participation	Panel data

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Boarnet et al. (2005)	Census Defined Places	1980-1990	Orange County (USA)	- Population growth in 1980 - Proportion of Hispanics - Proportion of Black - Changes in employment in surrounding areas. - Stock housing in 1940 and 1960 - Land use in 1980: single family residential, multi-family residential; agriculture use, general office use, retail stores and commercial services, public facilities, light industrial, wholesaling and warehousing and vacant.	Sensitivity analysis to different specifications in models of population and employment growth.	Spatial lagged explanatory SLX model, but only for the variables initial level of employment and changes in employment.
Brander and Dowrick (1994)	Country	1960-1985	107 countries	- Fertility (number of births) - GDP	Effects of fertility growth on economic growth	Robust OLS pooled, Panel Fixed and Random effect
Brueckner and Schandt (2015)	Country	1960-2007	139 Countries	- Economic variables: GDP per capita growth, Ratio female to male employment ratio, world income, trade, productivity. - Demographic variables: Fertility rate, infant mortality, under-five mortality, share of the population aged 1-14, 15-64 and +65, old age dependence ratio, ratio female to male population ratio, immigration	Effects on population growth of shocks to national income	Panel data and Instrumental Variables.
Burchfield et al (2006)	Metropolitan areas	1976-1992	USA	- Consumer amenities such as live performance venues per capital, restaurants/bar per capita - Number of street car passengers per capita in 1902, road intensity in the urban fringe-Past population growth - Standard deviation in the past pollution growth rates - Geographical information: mountain (range of elevation), terrain ruggedness index, mean annual cooling and heating degree days - Measure of political geography.	The purpose is to examine (i) residential and commercial land development between 1976 and 1992 and (ii) the determinants of urban sprawls	Bootstrap panel causality test

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Chang et al. (2016)	Country	1870-2013	21 countries	No explanatory variables	Identify the direction of the relationship between population growth and standard-of-living.	Granger Causality test with panel data and bootstrapping
Cheshire and Magrini (2006)	Functional Urban Regions (FUR)	1980-200	European Union-12 FUR	- Sectoral structure (industrial employment, agricultural employment, coalfield, Port size). - Gain of European integration Geographical position (west-country, south-country, west-EU, south-EU). - National ex-FUR population growth - Weather variables (wet day frequency ratio, frost frequency ratio, maximum temperature)	Explanation of the main drivers of population growth, with special attention to the weather variables.	OLS linear regression model and Spatial autoregressive model

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method

Chi and Ventura (2011)	Municipality level (MCD)	1970 to 2000	Wisconsin, USA	- Demographic characteristics: population density, percentage of young and old population, percentage of blacks and Hispanics; - Socioeconomic conditions: unemployment rate, income, percentage of population with high school education, percentage of population with Bachelor's degree, percentage of college population, percentage of housing units using public water, percentage of seasonal housing in the municipality, real estate value, county seat status, percentage of workers in retail industry and percent workers in agriculture industry; - Transportation accessibility: proximity to central cities, proximity to airports, proximity to major highways, highway density, and public transportation system - Natural amenities: percentage of forest coverage, percentage of water coverage, percentage of wetland coverage, percentage of public land coverage, lengths of lakeshore/riverbank/ coastline, golf courses, and view shed; - Land use and development: water, wetland, slope, tax exempted lands, and built-up lands.	Determinants of population growth at MCD level	▪ - Principal factor analysis and spatial overlay method ▪ - OLS and Spatial Lag econometric model

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Clark and Murphy (1996)	County	1989-1981	USA	- Population equation: Current employment density, lagged population density, fiscal policy variables, local demographic and neighborhood and economic characteristics (accessibility, education, poverty), natural amenities related variables. - Employment equation: current population density, lagged employment density, fiscal policy variables, local business conditions, controls for location, local demographic, neighborhood and economic characteristics (accessibility, education, poverty), natural amenities related variables	The purpose is to simultaneously estimate determinants of employment and population density changes	Two-stage least squares estimation
Cullen and Levitt (1999)	City and Metropolitan areas	1976-1993	USA	- City crime rates - Climate variables: Average July temperature; average January temperature, average yearly precipitations. - Education variables: high-school graduate. - Age variables: aged 0-17, aged 18-24, aged 25-44 and aged 45-64. - Share of workers in manufacturing - Income - Unemployment rate - Percentage of black people	Exploration of the effect of crime on population growth. Other determinants are considered as variables of control.	Linear regression model estimated by weighted OLS, instrumental variables and panel data.

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
da Mata et al. (2007)	City	1970-2000	Brazil	- Rural income opportunities - Market potential - Quality of labor force (schooling years). - Intercity-transport costs - State capital dummy - Initial level of population - Industrial composition - Local crime and violence	Explanation of the main drivers of population growth.	Linear regression model estimated by OLS and IV-GMM.
da Silva et al. (2017)	Minimum comparable areas	1970–2010	Brazil	- Literacy rate - GDP - Percentage of employes in Agriculture - Share of Manufacturing/Service employees - Employment rate - Birth rate - Mean age - Amenity related variables - Density - Homicide rate - Share of households supplied by water company - Share of households supplied by sewer company	Determinants of population growth	OLS and Spatial Error Model, Dynamic Spatial Durbin Models, spatial Durbin error model, Spatial Autoregressive Model.
Darrat and Al-Yousif (1999)	Country	1950-1996	20 developing countries	- Population growth - Per capita income growth	Relationship between Population and per capita income	Time series approach based on the Granges causality test.
Dawson and Tiffin (1998)	Country	1950-1993	India	- per capita income - Population	Relationship between Population and per capita income	Time series approach based on Johansen cointegration test

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Duranton (2016)	City	1993-2010	Colombia	- Demographic factors (raw growth rates and flows of migration) - Local labor market (wages) - Human capital (education) - Labor demand (industrial structure) - Geographic characteristics - Institutional characteristics (violence, criminality) - Natural resources - Road connectivity	Explanation of the main drivers of population growth	OLS linear regression model and method of instrumental variables
Duranton and Turner (2012)	City	1983-2003	US Metropolitan Statistical Areas	- Variables focused on transportation infrastructures: Kilometers of interstate highways in 1983, Kilometers of 1947 planned interstate highway, Kilometers of railroads in 1898, Routes of major expeditions of exploration between 1528 and 1850. - Climate variables: heating and cooling days, - Physical geographic characteristics: share of aquifers, elevation range and index of terrain ruggedness. - Socio-demographic characteristics: share of poor, share of adults with at least college education, income, segregation index, share of manufacturing employment and historical levels of population and employment.	Explanation of the main drivers of employment growth and population growth, with special attention to the effect of transportation infrastructures (highways).	Linear regression model estimated by OLS, and Instrumental Variables.

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Faria et al. (2006)	Country	1950-2000	125 countries	- Gross Domestic Product - Squared Gross Domestic Product - Regional dummies	Relationship between Population and per capita income	- Pooled OLS regression - Panel Data.
Ellen and O'Regan, 2010	City	1980-2000 (census data)	100 US central cities	- City's population: percentage of black people, the percentage of foreign born. - City median family income - Criminality. - Unemployment rate. - State per capita income. - Age distribution of the state population and its change in time.	overall city growth affected by changes in crime	Pooled OLS, Panel with city fixed effects, panel with city and time fixed effects, IV using severity of the criminal justice system as instrument for crime
Fukuda (2012)	City	2000-2005	321 Mid and small size Japanese cities	- Ratio of graduates of junior colleges, universities and graduate schools - Number of local government officers - Unemployment rates - Ratio of sewerage - Monthly average of the lowest daily temperature. - 4 Fiscal Indicators (ordinary balance ratio, ratio of outstanding borrowing, debt service payment ratio and financial index power).	Explanation of the main drivers of population growth, with special attention to fiscal indicators.	OLS linear regression model and spatial econometric models (spatial autoregressive model and spatial error model)

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Furuoka (2009)	Country	1961-2003	Thailand	- per capita income - Population	Relationship between Population and per capita income	Time series approach based on an the ARDL bounds testing approach to cointegration and Granger causality test
Furuoka (2013)	Country	1960-2007	Indonesia	- per capita income - Population	Relationship between Population and per capita income	Time series approach based on Johansen cointegration test, Breitung cointegration rank test and Granger causality test.
García-López et al. (2014)	City	1960-2001	77 Spanish metropolitan areas	- Transport infrastructures: highway rays in 2006, old Roman roads, 1760 Bourbon roads. - Land area - Geographical variables: Distance to coast, altitude, terrain ruggedness - Initial population 1960 - Age structure in the central city - Historic structures (cathedral and universities constructed before 1700).	Estimation of the impact of highway infrastructures on population growth.	Linear regression model estimated by OLS, and instrumental variables technique.
Garza-Rodríguez et al. (2016)	Country	1960-2014	Mexico	Population growth and GDP pc growth	long-run relationship between economic growth and population growth	Structural break cointegration analysis
Glaeser and Kohlhase (2004)	County	1920-2000, 1980-2000,	USA	No explanatory variables	The focus is on the implications of declining transport costs for goods over the last century whilst the transportation cost for people remained high and might have even increased because of the traffic congestion.	Scatter plots/ bivariate statistics

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Glaeser and Saiz (2004)	Cities and Metropolitan Statistical Areas	1970-2000	USA	- Weather variables: Heating-degree days, Annual precipitations. - Sectoral variables: Share of workers in manufacturing, share of workers in professional service, share of workers in trade, - Unemployment rate - Human capital: share of population with a college degree, share of high school dropouts, and number of colleges per capita in 1940. - Initial level of population.	Explanation of the main drivers of employment growth and population growth, with special attention to the effect of human capital.	Panel data model
Glaeser and Shapiro (2003)	Cities and Metropolitan Areas	1980s and 1990s	USA	- Region dummies - Median age of residents - Land area growth - Manufacturing share - · Density-related variables - Initial levels of population - Population per square mile - Percent driving alone to work - Percent taking public transportation - · Weather variables - Mean daily January temperature - Mean daily July temperature - Average annual precipitation - · Human capital variables: - Share of population with college or high school diplomas - Per capita income - Percent in poverty	Explanation of the main drivers of population growth in cities.	Linear regression model estimated by OLS.

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Glaeser et al. (1995)	Cities and Metropolitan Areas	1960-1990	USA	- Geographic variables: regional dummies. - Demographic variables: Initial population, past population growth. - Migrants - Socio-economic variables: initial level of income, Sectoral Composition (manufacturing share), unemployment, Inequality and Education. - Socio-political variables: racial composition and segregation, size and nature of government.	Explanation of the main drivers of population growth.	Linear regression models estimated by OLS.
Glaeser et al. (2001)	County	1977-1995	USA	- Climate variables: Temperate climate (inverse of average deviation from 70 degrees), Dryness (inverse of average precipitations). - Geographic variables: Distance to the coast. - Consumer city amenities: Live performance venues per capita, Restaurants per capita, Art museums per capita, Movie theaters per capita, Bowling alleys per capita, Restaurants per capita.	Explanation of the main drivers of population growth in cities.	Linear regression model estimated by OLS.

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
González-Val (2015)	City	1990-2000	USA	- Urban Sprawl Variables in 1990: Land area growth, Population per square mile, Median travel time to work. - Human Capital Variables in 1990: % high-school graduates or higher education. - Productive Structure Variables in 1990: Unemployment rate, Urban Diversity Index. - Weather Variables in 1990: Temperature Index, Percentage of Water Area, Annual Precipitations. - Control Variables in 1990: Initial per capita income, City population growth 1980-1990, Geographical dummy variables.	Explanation of the main drivers of population growth.	Linear regression model estimated by OLS, Spatial Error Model, Spatial Autoregressive Model and Spatial Quantile Regression.
González-Val and Lanaspa (2016)	City	1790-2000	USA	Previous size of population (testing of mean reversion)	Study of the path of growth of US cities (testing mean reversion in city sizes, mobility in the distribution of cities and cluster analysis).	Panel unit root tests Markov transition matrices Cluster analysis
Chi and Marcouillier (2013)	Minor civil division level	1995-2000	Lake State of Wisconsin, USA	Natural amenities: presence of forest, water, wetlands, public lands, lake/river/coast, viewsheds and gold courses. Demographic, accessibility and livability indices based on 32 variables.	The purpose of the paper is to study the effect of natural amenities on regional population growth.	OLS and Spatial lag models were employed with spatial regional specifications
Hasan (2010)	Country	1952-1998	China	- Per capita income - Real per capita capital stock - Saving ratio - Technological progress	Relationship between Population and per capita income	Time series approach based on cointegration and vector error correction modeling technique.

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Huang and Xie (2013)	Country	1980–2007	90 countries	- Population growth - annual growth rate of GDP - saving rate - secondary school enrolment ratio	Simultaneous effect of economic growth and population growth	Panel GMM
Huang, et al. (2002)	306 Southern and Midwestern rural counties	1950-1990	USA	- Rural median income, - School years completed, - Percentage of population with high school degree, - Distance to a city, - Herfindhal index of Employment, - Local and state government expenditures (e.g. welfare, education, debt, highway and debt). - demographic controls (proportions of farm, black, young and old)	Brain Drain from rural to urban counties	IV regression and ANOVA analysis
Jacobs-Crisioni and Koomen (2017)	Municipalities	1961-2011	Belgium, France, Germany, Italy, Luxembourg, the Netherlands, Switzerland, Portugal, Spain, and Austria.	- Accessibility, domestic (composite indicator) - Accessibility, foreign (composite indicator)	Unveil the connections between population growth and transport accessibility	Panel fixed effect
Jung and Quddus (1986)	Country	1952-1980	44 (19 of them developed countries)	- GNP per capita - Crude birth rate	Relationship between Population and per capita income	Time series approach based on the Granges causality test.
Kapuria-Foreman (1995)	Country		15 low- and middle-income developing countries	- Population growth - Per capita income growth	Relationship between Population and per capita income	Time series approach based on the Granges causality test.

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method

Lee et al. (2007)	City	1980-2000 (and sub-sample analyses every 5 years)	Korean cities	- Average years of education - Average age of residents - Square of average age - Per capita local burden - Unemployment rate - Manufacturing employment share - Employment share of the first industry - Distance from the nearest metropolitan area - Land-use regulations - Metropolitan, city and regional dummies.	Explanation of the main drivers of population growth	OLS linear regression model
Lehmijoki and Palokangas (2006)	Country	1989-1999	69 low- and middle-income countries	- GDP per capita - Adult female literacy rate - GDP share of international trade - Population density - Military expenditure share of GDP -Population growth 20 years ago - Regional dummies	Exploration of the political instability on population growth.	OLS regression model
Lewis and Stanley (2016)	County	1998-2012	South Carolina (US)	Information about local County and neighboring Counties of the following variables: - Manufacturing shares - per capita income - Birth rate - Black population share - Poverty rate - Density of recreational activities - Share of retirees - Share of single female pregnancy	Determination of the main factors that explain population growth in the counties of South Carolina (US). Comparison between urban and rural counties.	Spatial lagged explanatory variables (SLX) model

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Lutz (2001)	Urban	1966-1991	Ireland	- Proximity to the capital or largest urban centres (Dublin, Cork, Limerick) - Proportion of males - Percentage of employees by economic sectors - Percentage of young (<14) population - Percentage of students in the population	Assess which features of cities in Ireland have been associated to subsequent population growth	Multivariate regression on the growth in population: by 10 year regressions
Lutz and Qiang (2002)	Country	1960-1990	187 countries 143 Less Developed Countries	- Population density - Female labor force participation rate - Female literacy rate - Population urban - GDP per capita - Food production index	Main determinants of population growth (special focus on the effect of population density)	Cross-sectional regressions in 5 years steps from 1960 to 1990.
Martori et al, 2016	Metropolitan area	2001-2008	Barcelona, Spain	Indicators of immigrant diversity, unemployment rate, educational level, income, characteristics of the dwellings, density, and distance to the business centre. Values of the explanatory variables at the initial period.	The purpose is to study the determinants of no EU immigrant population growth within the Barcelona metropolitan area during the 2001-2008 periods.	Spatial lag and spatial error models employ. GMM of these two specifications also used. Decomposition of the effect into direct and indirect effects.
Nakibullah (2010)	Country	1959-1990	Bangladesh	· Gross Domestic Product per capita	Relationship between Population and per capita income	Time series approach based on the Granger causality test
Osili and Long (2008)	Country	1999	Nigeria	- Levels of Education - Fertility rates	Impact of female schooling on fertility rates.	Instrumental variables and differences-in-differences approach

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Oueslati et al. (2015)	Cities	1990, 2000 and 2006	237 European cities	- GDP per capita - Initial population - highway density - ratio of agricultural - added value on agricultural land - crime - rain - temperature - elevation - cinema per habitants - concentration of N02 - Airports	The paper tests the monocentric model and finds that sprawl is driven by GDP per capita and population while transportation costs and agricultural land values limit sprawls. Panel specifications and a Hausman Taylor estimator are used.	Nonlinear and linear cointegration tests, Vector Error Correction Model
Partridge et al (2007)	Rural Counties	1990-2004	USA	- Initial population density - Natural amenities: climate variables, percent water area, topography measure. - Demographic variables: 1990 population shares of four education categories, the percent of population that immigrated between 1985 and 1990, six 1990 population age shares, and five race and ethnicity population shares. - Economic variables: initial unemployment rate, median household income, goods-producing and agriculture shares, industry mix employment growth rate. - 5 metrics to measure the distance to urban center related.	The purpose is to examine if there are spatial heterogeneities in the determinants of employment growth.	Linear regression model Weighted linear regression model Spatial regression model

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Partridge et al. (2008)	Counties	1950-2000 and sub-periods	USA	- Climate variables - Potential market proxy - State fixed effect - Distance to the nearest urban areas and to the higher tier cities	The purpose is to study the effect of distance to urban areas in rural counties and to test if the effect of distance go beyond the potential market effect associated with urban areas.	GMM-based cross-sectional estimations
Pegou-Sibe et al. (2016)	Country	1960-2013	The 30 most populated countries	- per capita income - Population	Relationship between Population and per capita income	Time series approach based on an Error Correction Model and Granger causality test
Pirotte and Madre (2011)	Metropolitan areas	1985-1998	France	Indicators of income growth, income inequality, population density over different time periods	The purpose is to examine the determinants of urban sprawls following the monocentric model. Panel specifications and random coefficient models are employed.	Instrumental variables Instruments based on natural experiments
Rappaport (2007)	County	1970-2000	USA	- Sectoral structure (industrial employment, agricultural employment, mining employment…) - Coastal proximity and topography. - Initial population density and surrounding total population. - Market potential to local firms. - Census divisions. - Weather variables and their quadratic forms (January daily maximum temperature, July daily heat index, July real humidity, annual precipitations, annual precipitation days).	Explanation of the main drivers of population growth, with special attention to the weather variables	Linear regression model estimated by OLS and GLS.

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Rickman and Wang (2017)	Rural-Urban General Equilibrium Model	2000-2010	USA	- Household amenity demand - Natural amenity - Attractiveness of cities - Productivity - Wages	Focus on the role of natural amenities and urban agglomeration economies on population growth	Spatial equilibrium model: Spatial hedonic growth model
Savas (2008)	Country	1989-2007	Kazakhstan, Kyrgyzstan, Tajikistan, Turkmenistan and Uzbekistan	- per capita income - Population	Relationship between Population and per capita income	Time series approach based on an the ARDL bounds testing approach to cointegration Granger causality test
Singh (1994)	Country	1980s	Less developed countries	- Women's education - Women's labor force participation - Health and family planning services - Birth control measures	Effects on fertility and mortality rates.	Cross-sectional regression model
Skirbekk et al. (2014)	Major Geographical Areas	1820-2100	Kuwait, Indonesia and China, France, Germany and Sweden	No explanatory variables	Global and regional population growth during the transition period, had the rest of the world experienced the same population growth patterns than Europe	Simulations
Thornton (2001)	Country	1900-1994	Argentina, Brazil, Chile, Colombia, Mexico, Peru and Venezuela	- per capita income - Population	Relationship between Population and per capita income	Time series approach based on Johansen cointegration test

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Trendle (2009)	Local Government Areas	1996-2006	Queensland (Australia)	· Determinants for population growth: ▪ Previous level of population ▪ Employment ▪ Levels of income ▪ Levels of Education (proportion of the population aged 15 and over with a post-school qualification). ▪ Proportion of employees in health related jobs. ▪ Index of social advantage of the region · Determinants for employment: ▪ Previous level of employment ▪ Levels of Population ▪ Levels of income ▪ Unemployment rate ▪ Index of social advantage of the region ▪ Index of regional specialization (Herfindhal index of industrial concentration).	Investigation of the main determinants of population and employment changes; as well as the study of the causal relationships between population and employment.	Two Stage Least Squares and Spatial Econometric Model (Spatial Autoregressive Model).

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Veneri and Ruiz (2016)	NUTS3	2000-2008	OECD regions	- Distance from each rural region to the closest urban centre, Density Elderly dependency rate (population older than 64 / population aged 15-64). - GDP rural - GDP urban or intermediate - Share of employment in industrial activities - Unemployment rate rural regions - Growth rate of GDP – urban and intermediate - Total population urban and intermediate regions - Growth rate population urban and intermediate regions - Urban region dummy	Rural regions	penalized spline geo-additive model, cross section
Wei et al. (2015)	Country	1949-2012	China	- Sex ratio - Degree of urbanization - Employment rate (as proxy of income). - Government Policies (fiscal expenditures). - Natural disasters	Main determinants of population growth	Time series regression model estimated by ordinary less squares.

Author	Level	Period	Region/country	Relevant explanatory variables	Focus	Method
Winters (2011)	County	1995-2000	USA	- Economic and Demographic variables: county population, per capita income, the share eof employment in manufacturing. - Natural Amenities: mean January temperature, mean January sun hours, mean July temperature, mean July relative humidity, area coverage, topography score). - Distance variables (population to the nearest MA, distance to the nearest MA, distance to the nearest MA with a population of at least 250000, 500000 and 1.5 million.	The purpose is to examine the importance of student migration in the relationship between human capital and population growth	Spatial error model

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications
You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-AE-18-010-EN-N

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub
ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

Joint Research Centre

EU Science Hub

Publications Office

doi:10.2760/513062

ISBN 978-92-79-93404-9