Integrative Mechanisms for Addressing Spatial Justice and Territorial Inequalities in Europe

D 3.3. Report on Economic Growth and Spatial Inequalities

Version 1

Authors: Alfredo Cartone, Domenica Panzera, Paolo Postiglione, Panagiotis Artelaris, Ana Viñuela

Grant Agreement No.: 726950

Programme call: H2020-SC6-REV-INEQUAL-2016-2017

Type of action: RIA – Research & Innovation Action

Project Start Date: 01-01-2017

Duration: 60 months

Deliverable Lead Beneficiary: Ud'A

Dissemination Level: PU

Contact of responsible author: postigli@unich.it

This project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 726950.

Dissemination level:

- PU = Public
- CO = Confidential, only for members of the consortium (including the Commission Services)

Change control

VERSION	DATE	AUTHOR	ORGANISATION	DESCRIPTION / COMMENTS

Bibliographic Information

Alfredo Cartone, Domenica Panzera, Paolo Postiglione, Panagiotis Artelaris, Ana Viñuela (2020). *Deliverable 3.3 - Report on Economic Growth and Spatial Inequalities*. IMAJINE WP3 Territorial Inequalities and Economic Growth

This report can be downloaded from our website http://imajine-project.eu/

January 2020

Acronyms and Abbreviations

AIC	Akaike Information Criterion
ARDECO	Annual Regional Database of the European Commission's Directorate General for Urban and Regional Policy
EU	European Union
GDP	Gross Domestic Product
GS2SLS	Generalised Spatial Two Stage Least Squares
GWPCA	Geographically Weighted Principal Component Analysis
MCMC	Markov Chain Monte Carlo
NEG	New Economic Geography
NUTS	Nomenclature of Territorial Units for Statistics
OECD	Organization for Economic Cooperation and Development
OLS	Ordinary Least Squares
PCA	Principal Component Analysis
PPS	Purchasing Power Standard
R&D	Research and Development
SDM	Spatial Durbin Model
US	United States
WP	Work Package

Index

List of Figures

List of Tables

NON-TECHNICAL SUMMARY

The rise in the number of people living in poverty or at risk of social exclusion has contributed to broadened economic and social disparities between and within the European Union (EU) Member States, putting at risk the principle of cohesion (European Commission, 2014). The 2008 global crisis partially stalled economic growth and convergence with effects on EU inequality. Furthermore, within-country levels of inequality have increased in many EU Member States, particularly in Central and Eastern European countries (Beckfield, 2019).

Since 1986, with the *Single European Act*, cohesion become a foundational idea of the European Community, which was called upon to support reduction of disparities among regions and overcome the backwardness of less-developed regions. In 2008, a broader idea of cohesion was introduced with the *Lisbon Treaty*, including in the social, economic, and territorial dimensions. In this context, addressing regional disparities (i.e. different levels of socio-economic development at the regional level) assumes relevance for achieving social, economic, and territorial cohesion. Moreover, regional inequality is not only relevant by itself, but also because of its potential implications for economic growth. In fact, regional inequality may impact the economic convergence process by retarding or promoting growth in richer and poorer regions (De Dominicis, 2014).

Given the evidence provided in the WP2 of the current project, we acknowledge how the analysis of regional disparities at a detailed geographical level can define strategies that meet local peculiarities. Hence, a deeper understanding on what is going on within the regions (e.g. NUTS 2)[1] could be required for more effective actions. With respect to the key concepts of spatial justice and territorial cohesion expressed in WP1, a strong consideration of geographical approaches informs this report, as well other research from WP3 and WP2

Different approaches can be used to analyse inequality. However, it is evident that regional inequality requires addressing issues linked to their specific location (Rey and Smith, 2013; Panzera and Postiglione, 2020). This implies accounting for features that can suggest effects of significant interconnections. In addition, taking into account

[1] Especially in deliverable D2.4, the Authors offer a wide evidence of within NUTS2 disparities. Hence, they conclude the need for more detailed analyses, which appears in line with the aim of supporting territorial justice (see deliverable D1.3 for a summary of key concepts within the IMAJINE Project).

heterogeneity across regions sheds light on structural differences among the drivers of growth and inequality.

Inequality is also a multidimensional challenge. Composite indicators may provide extremely important evidence that enlarges the picture usually provided by analyses on regional GDP or other economic variables.

All those issues are addressed in this report, which provides an empirical analysis of inequality and growth at various geographical scales over a long period. The empirical findings derived from the analysis are not only meant to increase knowledge of well-known regional disparities, but also to address potential links, relationships, and structural differences that drive the magnitude of regional inequalities. In a nutshell, policymakers should be informed on how interdependencies, interconnections, and structural differences affect development. Those issues could be at the basis of efficient and effective policies to tackle disparities and increase awareness of EU actions.

Considering space in the analysis of regional inequality

Usually, spatial information is not considered when measuring inequality. Data are used to calculate synthetic indices, such as the Gini index or entropy-based measures, which provide a measure of inequality that fails in capturing the geographical position of data. This condition implies that geographical location do not impact the resulting inequality measure when GDP values are constant. Looking at maps included into WP2 (deliverable D2.4), it can be seen that significant patterns of inequalities over the spatial dimension emerges at first glance. Therefore, we want to explore how geographical position could influence inequalities.

Because of the spatial dependence, GDP could assume similar values in neighbouring regions, or it could exhibit differences, presenting evidence of a negative spatial association. Including the spatial dimension when measuring regional inequality facilitates the understanding of inequality dynamics over territories. An empirical analysis is offered in Section 6, mainly focused on the GDP per worker observed in NUTS 3 EU regions from 1995 to 2019.

Considering the impact of inequality on regional growth

The primary literature on this topic is reviewed in Section 2. A theoretical model that relates inequality and growth is defined and presented in Section 4. In Section 6, the model is used in an empirical analysis focused on EU NUTS 2 regions. The proposed model investigates how intra-regional inequality impacts growth in various EU regions. The analysis of regional inequality and regional development contribute to the debate over the effects of inequality on economic growth (see, deliverable D1.3). This aspect is also considered with respect to two groups of regions, i.e. less-developed regions and the group composed by transition and more-developed regions. These groups are identified according to the EU Cohesion policy for the 2014-2020 funding period.

Dimensions of inequality

Composite indicators may be used to explore disparities and development at very local level. However, the definition of a composite indicator leads to choices of whether or not to assign different levels of importance to the considered dimensions, selection of methodologies for data treatment, and careful interpretation. When defined, composite indicators should be used to provide a detailed mapping of socio-economic conditions. An exploratory composite indicator of local economic development is proposed in Section 5, using local data from deliverable D2.3 (IMAJINE WP2). The selected geographical scale is consistent with the ideas emphasized in deliverable D2.4. The empirical analysis is presented in Section 6 and allows to complement previous evidence in terms of multidimensionality and fine geographical scale.

1. INTRODUCTION

1.1 Spatial justice, spatial inequality, and local development: what do we address?

Deliverable 3.3 aims at offering empirical evidence related to the themes of inequalities and economic growth at different geographical scales. Regional units, at both the NUTS 2 and the NUTS 3 levels are considered, as well as local geographical units (i.e. municipalities). In this Section, we focus on some theoretical backgrounds that help us to set a more precise perimeter for our research, emphasizing links with other IMAJINE WPs. In fact, many of the general ideas that inspire the project and this report have been previously debated by other WPs (WP1 & WP2). However, some choices are needed in order to measure inequality in a way that can lead to more mindful Cohesion policies.

In the deliverable D1.3 a range of definitions of territorial inequalities and spatial justice are offered. In the same report, it is noted that:

> "*a common meaning attached to the concept of spatial justice is <<spatially even or equal accessibility to certain services and opportunities>>. In some cases, this definition is formulated in a way that <<people should not be disadvantaged by their place of residence>>.*"

This result involves a key issue in the definition and measurement of spatial inequality. In fact, linking location (in a geographical sense) to a condition of disadvantage opens two relevant questions:

1) What is the effect of spatial location on inequality?
2) How much of the regions' inequality is due to their spatial location?

Those questions have received little attention in the literature and may offer supporting evidence in the field of spatial justice (Rey and Janikas, 2005). The first one is connected to the detection of a null, positive, rather than negative effect of space on inequality. The second is tightly connected to quantifying this effect.
When considering regional units of analysis, we mainly focus on traditional measures of inequality that are opportunely modified to account for the spatial interactions and the geographical position of units (see Section 3). Great relevance is attributed to the distinction between within regions inequality and between region inequality (see Section 4.2), and a methodology to assess the impact of inequality on regional growth is presented (see Section 4.1).

Additional key concepts need to be recalled when approaching measurement of inequality at local level. On the one hand, researchers may be interested in measuring subjective inequality, by turning to impressions and opinion about self-perceived conditions. On the other side, one may refer to a universally accepted definition of needs to decide if a human being is deprived or not. This debate is well-known either in the field of deprivation (Townsend, 1987) and human well-being (Sen, 1980). This point is also faced earlier in the project, where a distinction between the two concepts is shown with a discussion on measurement. In fact, there is a body of literature emphasizing the importance of subjective well-being as reviewed in deliverable D2.4. Furthermore, inequalities across territories within the EU are analysed from the public opinion's perspective in WP4. In our report, we mainly follow the evidence from previous deliverables, developing an empirical setting that uses the WP2 indicators.

Scientific literature in many fields of social sciences does not converge on the use of income as a generally accepted indicator for well-being. However, data about income are often considered starting points for measuring inequality (Silber, 2012). Empirical researches on inequality often trade-off between data availability and representativeness of material needs (Bartiaux et al. 2019). In this report, we face limitations connected to data availability at a detailed geographical scale. Hence, in some applications throughout this report, we will mainly focus on income distribution to address what effect has space on inequality. The same remark is applied for the study of the influence of inequality on economic growth when considering lower geographical scales.

Other Authors have highlighted the importance of considering different dimensions at the same time (among others, Stiglitz et al. 2009). The debate around composite indicators can be overviewed by different angles: from a measurement and validation perspective (Decancq and Lugo, 2013; Nardo and Saisana, 2008), from a policy perspective (Munda and Nardo, 2009), and from an epistemological perspective (Saltelli, 2007). Despite different orientations, we notice how composite indicators are nowadays tools to build "narratives" (Kuc-Czarnecka et al. 2020) that can support larger awareness. Therefore, major attention is devoted to the definitions of multivariate local indicators as well as to the definitions of indicators that include geographical characteristics.

From a methodological perspective, a key concept that inspires us is embedded in the term "spatial". In this report, "spatial" is referred to attributes of data that rise from the presence of geographical reference. Haining (2003, p.1) refers to "spatial" as:

> "*each item of data has a geographical reference, so we know where each case occurs on a map.*"

In the statistical as in the econometric literature, the term "spatial" has gained progressive relevance. However, when we ground into the spatial econometric literature (Paelinck, 1978), spatial dynamics may often appear mechanical (Corrado and Fingleton, 2012). Only mentioning bias and efficiency in data treatment would not help the general reader to address relevance of spatial data treatment for policies. Tobler's (1970) first law of geography represents a crucial point for spatial analysis, which also roots in a deeper understanding of social phenomena. In fact, besides being a milestone (Arbia et al. 1996) for quantitative analysis, the contribution of Tobler led to a new way to look at local effects. As Tobler (2004) noted on the success of its first law:

> "*I am a great believer in simplicity, when this is possible. For example, the point in science is to achieve as many results as possible with the fewest hypotheses. So, in order to simplify the problem of depicting the growth of population in the Detroit region, I tried to eliminate complicating factors. This is when I invoked <<the first law of geography: everything is related to everything else but near things are more related than distant things.>> Doing this allowed me to concentrate on local effects*"

Therefore, to point out the effects of geographical attributes on inequality, we will use spatial techniques to shed more light on local effects.

Spatial techniques offer us a way to interpret many geographical issues linked to inequality, but they cannot be interpreted in a simplistic way. Spatial interactions descend from a cascade of features that go far beyond the mere physical sense (i.e. roads, transportation, commuting flows). Similarities and geographical differences happen because proximity leads to cultural similarities as well (Anselin, 1988). Following this line, "spatial" should not be used as a crude attribute that emerges from data. A more useful assessment of spatial analyses in social sciences comes from interpreting them as "regularities" (Tobler, 2004) that lead to a better knowledge of territorial phenomena."

1.2 Regional growth and inequality

The issue of regional economic growth and inequality has always been a lively topic for scholars and policymakers (Myrdal, 1957; Borts and Stein, 1964; Williamson, 1965). However, in the last few years, the discussion concerning these topics has attracted remarkable attention, leading to renewed interest stimulated, among others, by i) recent increases in income inequality in several advanced countries, ii) the explosion of interest in analysing the determinants of economic growth, iii) the availability of comparable national and regional data, and iv) the fact that both theoretical and empirical studies have produced controversial and ambiguous results.

This widespread interest has also been stimulated because regional inequality and economic growth have been dramatically reshaped, and they are becoming more challenging for the EU and its Member States. For instance, the EU regions have been experiencing a period of extraordinary change and transformation as a result of globalization (i.e. economic integration), which has marginalized several regions in the rich world. The recent economic crisis and recession have led to increased inequality (Brakman et al. 2015; Capello et al. 2015), and recent political changes (e.g. the surge of populism and anti-establishment extremism) can be linked to regional inequality and economic growth (Artelaris and Tsirbas, 2018; Dijkstra et al. 2020).

Although there has been a long-lasting discussion on the relationship between inequality and growth, and the theoretical and empirical literature is well-developed and extensive, the links and interconnections are not fully understood. The empirical evidence is far from clear cut, and much more research is needed in this area and relatively little research has focused on understanding how inequality affects economic growth in the EU. However, assessing the impact of inequality is extremely relevant for promoting economic growth, local development, and territorial cohesion.

1.3 The spatial dimension

Spatial effects have been introduced in the econometric literature (Anselin, 1988). Spatial dependence can be defined as a property of geographical data for which near observations tend to assume similar values (positive dependence) or dissimilar values (negative dependence). Spatial dependence is often caused by significant interconnections between neighbourhoods, but it is not limited to that. On the other side, spatial heterogeneity defines structural differences that leads to “instability” of statistical relationships over space.

The most popular inequality measures include the class of generalized entropy inequality measures (Shorrocks, 1980) and the Gini index of inequality (Gini, 1912). All these measures are invariant with respect to location, i.e. they do not account for the geographical position of data, and thus, they are insensitive to any spatial permutation. This condition of inequality measures, known as anonymity (Bickenbach and Bode, 2008), has been addressed in some recent contributions, which have focused primarily on assessing the contribution of spatial dependence to regional inequality (Arbia, 2001; Arbia and Piras, 2009; Rey and Smith, 2013; Márquez et al. 2019; Panzera and Postiglione, 2020).

Spatial interconnections could be usefully incorporated in the analysis of the link between inequality and growth. As suggested in the research lines that concludes deliverable D2.4, a study of regional disparities and economic growth should consider both the tendency of similar values to cluster together in space and local differences. In fact, the factors that drive economic growth may follow dynamics that are linked to local structure of sectors, long-lasting effects of specialization, and similarities (or differences) in the social context.

Further, focusing on a multidimensional concept of inequality, the different aspects that significantly contribute to balanced economic development are synthetized through the use of composite indicators. After selecting some relevant dimensions, their aggregation into a single indicator can be obtained by different methodologies. In this report, an easy-to-implement approach is carried on to account for the spatial nature of data.

The treatment of spatial effects is not only relevant from a methodological point of view. Considering the territorial dimension, the interactions among regions, and their differences, is relevant from a policy perspective. In fact, it helps developing interdependent scenarios. These considerations assume a special relevance with regard to EU regions, where disparities are made more evident by EU enlargement and recent economic crises.

All the aforementioned issues are addressed in this report. This report includes seven sections. In Section 2, we review the literature on the relationship between inequality and growth. Section 3 concerns spatial approaches to the inequality measurement In Section 4, we define a spatial model that relates inequality and regional growth. Section 5 is devoted to the definition of composite indicators that synthetize different dimensions of disparities at the local level. In Section 6, we present empirical studies conducted at different spatial scales. Finally, conclusions are provided in Section 7.

2. LITERATURE REVIEW

2.1 Regional economic growth and inequality: Theoretical background

Growth theories provide a rich framework for explaining the relationship between economic growth and income inequality. However, they provide different predictions of this link. Neoclassical growth theory is a usual theoretical starting point that has gained widespread acceptance by regional and other economists. The work of Solow

(1956) and Borts and Stein (1964) suggest that regional differences disappear with growth (see Barro and Sala-i-Martin, 1995, for a review). Under the assumption of diminishing returns to capital, homogeneous regions that present, for instance, similar preferences and technology will tend to show reduced regional differences, and per capita income will converge to a common level (convergence towards a single steady state). Conversely, if economies are heterogeneous, convergence will occur only in a conditional sense (conditional convergence) because economies will grow toward different steady-state positions.

In contrast, a number of growth theories consider the emergence of economic inequality among economies the most possible outcome, and obviously, regional convergence does not occur. For instance, most models of endogenous growth theories (see Aghion and Howitt, 1997, for a review), initiated by the pioneering work of Romer (1986) and Lucas (1988), tend to agree that economic growth is a spatially cumulative process that is likely to increase inequality. Romer and Lucas, for example, predict diverging growth patterns primarily because of the lack of diminishing returns to capital, which is related to the endogenous character of technology.

Similarly, a positive link is also predicted by the growth theory of cumulative causation developed by Myrdal (1957) and Kaldor (1970) for explaining how different economies exhibit different performances. In contrast to the aforementioned theories, these models have a medium-term view and are often described as ‘soft’ development theories due to a lack of applied mathematical rigour and empiricism (Plummer and Taylor, 2001). The main interest of these less influential and more regionally oriented theories (rather than countries) is the interregional interactions derived from the growth process in an economy and their effect on national development. The core is the assumption of circular and cumulative causation, resulting in a self-sustaining and unbalanced process of economic growth across economies. As a result, the most advanced regions may grow faster than the rest. These theories are closed related to concepts such as agglomeration economies, growth poles, learning regions, and competitive advantage.

New Economic Geography (NEG), which is the most recent theoretical development in this field, has much in common with the growth theory of cumulative causation because it stresses that economic growth tends to be an unbalanced process favouring the initially advantaged economies (Krugman, 1991; Fujita et al. 1999). As Cheshire and Malecki (2004) point out, NEG has made an important contribution to the formal theory of regional growth since Borts and Stein’s contribution. The main assumptions of this theory include, among others, increasing returns to scale, monopolistic competition, labour and capital mobility, and non-zero transportation costs. The spatial distribution of economic activity can be explained by agglomeration (or centripetal) and dispersion (or centrifugal) forces. The former includes backward and forward linkages of firms, externalities, and scaled economies, while the latter includes negative externalities, transport costs, and intensification of competition.

In the middle of this theoretical spectrum, some recent growth models (Azariadis and Drazen, 1990; Galor, 1996) generate multiple steady-state (locally stable) equilibria and convergence clubs (see Azariadis, 1996). Club convergence implies convergence to a common level only for economies that are both identical in their structural characteristics and similar in their initial conditions. In other words, these models transcend the ‘all or nothing’ logic behind conventional growth models and maintain that convergence may occur for different groups of (regional) economies. Multiple equilibria and convergence clubs can emerge on account of differences in, among others, human or physical capital, income distribution, capital or market imperfections, local complementarities, and externalities.

From a more micro perspective, researchers have tried to examine whether inequality retards or accelerates economic growth in several theoretical studies (for review, see Aghion et al. 1999; De Dominicis et al. 2008). These research efforts are part of the explosion of interest in the determinants of economic growth fuelled, among others, by the development of endogenous growth theory in the 1990s. This literature offers several alternative explanations for the possible effects of income inequality on economic growth and arguments them in both directions (positive and negative link). Early reports predicted a positive relationship between income inequality and growth based on the Kaldorian hypothesis, i.e. greater income inequality promotes saving, capital accumulation, and hence growth (Bandyopadhyay and Tang, 2011; Weide and Milanovic, 2018). However, a large body of literature in the early 1990s suggested a negative link between the two variables providing alternative theoretical channels. In the following years, other theoretical models showed positive effects again. Interestingly, the main arguments for the positive or negative link are similar and associated with i) the political economy/institutional mechanisms, ii) human capital endowments, and iii) the effects of social and political dissatisfaction (Perugini and Martino, 2008). Although most of the arguments are focused at the national level, some of them are particularly relevant at the regional level (Royuela et al. 2019).

A further aspect concerns the dynamics of regional inequality across different economic cycles. Regional inequality is expected to rise in periods of economic expansion and decrease in periods of contraction (i.e. pro-cyclicality), or vice versa (i.e. anti-cyclicality). A pro-cyclical behaviour can be explained, for example, by the fact that expansion cycles begin at the poles of economic activity, where the interaction between agglomeration effects and market size provides a lead over other regions. On the contrary, during a recession, these poles are more exposed to supply and demand contractions, thus they are more likely to be negatively affected than the rest of the regions, resulting in decreasing regional inequality (Berry, 1988). An alternative explanation is proposed by Rodríguez-Pose and Fratesi (2007) who suggest, based on an examination of the southern European countries, that shelter economies (which are typically less developed, more isolated, and more dependent on factors such as public investment and employment) are unable to catch up with the more advanced economies in periods of expansion. This occurs primarily because they are less exposed to changes in market conditions. Not surprisingly, the opposite is true for the contraction period. Other scholars such as Pekkala (2000) propose an anti-cyclical pattern because the mobility of labour is higher and regional policies are more efficient in periods of expansion, resulting in a more even distribution of wealth across territories.

The main characteristic of the aforementioned theoretical contributions is that they examine the causation from inequality to growth. As mentioned above, this relationship is uncertain and still under debate.

A frequently neglected contribution is the work of Williamson (1965), which is partly motivated by the famous inverted-U hypothesis. This hypothesis was originally advanced by Kuznets (1955), who related in a systematic way the level of economic development with the level of personal income inequality, arguing a rising trend in the early stages of development but a declining trend in later stages. Also influenced by the works of Myrdal (1957) and Hirschman (1958), Williamson (1965) modified the inverted-U (or so-called bell-shaped) hypothesis relating the national development process to regional disparities. As Nijkamp (1990) notes ‘an important analytical framework, the so-called Williamson hypothesis, has to a large extent been neglected in the scientific literature’. Cheshire and Malecki (2004) stress that ‘memories seem increasingly short and Williamson’s work has been all but forgotten’.

Williamson identified four main factors that govern the evolution of regional inequality: labour migration, capital migration, interregional linkages, and central government policy. In the earlier stages of development, production factors, such as labour and capital, are concentrated in relatively few growth poles. In parallel, there is an absence of interregional linkages that minimizes spread effects. Finally, the central government's policy is focused on strengthening aggregate national growth, supporting growth poles, and increasing regional inequality. As a result, the earlier stages of development are associated with rapid income growth in the core regions and increasing regional disparities (i.e. divergence). However, this situation is unlikely to persist indefinitely. In the latter stages, both capital and labour migration become less selective as national labour and capital markets become more sophisticated. Furthermore, interregional linkages are reinforced, strengthening spread effects, whereas the central government pursues a redistributive policy where resources are transferred from richer to poorer regions. As a result, the latter stages of development are associated with decreasing regional disparities and convergence. In other words, Williamson argues that less advanced countries are characterized by increasing inequality and more advanced countries by decreasing inequality, meaning regional inequality tends to be higher in middle-income countries.

More recently, a few scholars have proposed a pattern of increase-decrease-increase ('augmented inverted U'). Williamson's factors may lead towards reduction, but at the same time, these factors operate within a rapidly changing environment that may possibly offset their influence at higher income levels (Easterlin, 1958). The operation of 'net backwash' effects might be stronger than the operation of 'net spread' effects in advanced economies. According to Amos (1988), the main reasons driving increased regional inequality in advanced economies may include increases in suburbanization and the movement toward a service-based economy. Fan and Casetti (1994) consider the main elements of each phase to be flows of production factors, particularly capital and labour. More specifically, they suggest that regional divergence and polarization are expected in the first phase, especially because of the existence of a 'leading sector', resulting in formation of a core region and its self-sustaining growth. Growth and prosperity of the core region can be explained in part by circular cumulative causation and the core-periphery model. The second phase of regional dynamics is characterized by slower growth and decline of core regions and new growth in the former periphery mainly because the advantages of agglomeration in the core regions are counteracted by a new set of forces that favour new locations of growth and agglomeration in the periphery. Finally, in the third phase, regional divergence is expected mainly because of spatial restructuring, which triggers new directions of capital flows. In essence, capital moves back to selected locations within the traditional core regions where the new leading sectors thrive.

Finally, another possible interpretation is associated with the fact that regional economies change in a number of ways during development because they become larger, deeper, more diversified, and with different structures and levels of intraregional and interregional integration. As regional economies change, it is reasonable to assume that the balance of forces leading them to convergence or divergence also changes (Petrakos et. al. 2011; Artelaris and Petrakos, 2016). More specifically, less advanced regions are more likely to be characterized by a productive system where resource-intensive activities dominate, and where markets are relatively shallow or fragmented.

2.2 Regional economic growth and inequality: empirical findings

Since theories deliver different messages about the growth-inequality nexus, theoretical developments have been accompanied by a growing number of empirical studies. However, these studies have also provided mixed evidence. Although several of them report a positive relationship, others support the existence of a negative link.

The empirical studies have been based on both cross-sectional and panel approaches. Cross-sectional analyses exhibit weaknesses mainly related to uncertainty in the model. This uncertainty is, to great extent, caused by the absence of a generally accepted formal theory of economic growth, parameter heterogeneity, outliers, endogeneity, measurement errors, and error correlation. Cross-sectional specifications and fixed-effects panel models can be considered complements rather than substitutes; the former capture the way in which persistent cross-sectional differences in inequality affect long-run growth rates, while the latter capture how time-series changes within a region affect changes in its growth rate over time (Royuela and García, 2015).

Spatial econometrics specification, for both the cross-sectional and panel data models, has also been used in recent years as it has widely acknowledged the significant role of spatial effects in the economic growth literature; regression models might be incorrectly defined due to ignored spatial autocorrelation/dependence (Rey and Montouri, 1999; Ertur and Koch, 2007) caused by several factors, such as trade between regions, labour and capital mobility, and technology and knowledge diffusion. This may lead to serious bias and/or inefficiency in the estimates of coefficients. As a result, spatial econometric techniques should be applied using cross-sectional or panel data methods.

Because there are several theoretical ways to approach the growth-inequality nexus, there are also several empirical ways to approach this link. Some researchers have indirectly examined this topic by looking at the issue of economic convergence/divergence, which can also be used to test the validity of the primary growth theories. The dominant approach in this literature follows the work of Baumol (1986), Barro (1991), and Barro and Sala-i-Martin (1992). At the core of the approach is the concept of unconditional β-convergence: convergence can occur in an absolute sense if economies are homogeneous (unconditional β-convergence) because they will converge towards the same steady state.

Since these seminal cross-country studies, the convergence/divergence issue at regional levels has been extensively debated as regions are assumed to consist of more homogenous economies. However, results from several studies do not confirm the convergence hypothesis supporting the existence of selective tendencies, convergence clubs, and asymmetric shocks, which might lead to greater spatially inequality (see Magrini, 2004; Petrakos and Artelaris, 2009, for reviews). For further information on the convergence issue, please refer to Deliverable D 3.2.

Some researchers have directly examined the growth-inequality nexus using different growth models and by focusing on the effects of inequality on economic growth. Initially, these approaches were used to explain differences in growth among countries, but they were later applied at the regional level. In this kind of analysis, the dependent variable is usually represented by the average growth in per capita GDP over a period throughout a broad range of economies.

This literature has been mushrooming in recent years, with puzzling results greatly attributed to differences in data, the time span of the data, sample coverage, and estimation methods (for literature surveys, see, for instance, Aghion et al. 1999; De Dominicis et. al 2008). In general terms, the first wave of empirical studies in the early 1990s found a negative effect of high inequality on subsequent growth, using primarily cross-sectional data (Persson and Tabellini, 1994; Alesina and Rodrik, 1994; Perotti 1996). The main problem of these studies is associated with the quality and comparability of the inequality data (resulting in measurement errors) and with severe econometric problems, such as omitted-variable bias. A new dataset by Deininger and Squire (1996) alleviated some of these problems in terms of reliability of the data, and they offered a new opportunity for a second wave of empirical studies to produce more reliable results. The majority of these studies found a positive relationship (see, for instance, Benabou, 1996; Partridge, 1997; Li and Zou, 1998) using panel data methods. Forbes (2000), using a panel of mostly rich countries, found that higher inequality was positively associated with growth and concluded that the relationship is negative in the long run while it is positive in the short or medium run.

Mixed evidence was found, or previous findings were reconciled, by investigating potential nonlinear effects in other empirical studies in the literature. In essence, different effects were found in these studies, where rich and poor countries were considered. For instance, Barro (2000) concludes, using panel data, that higher inequality tends to retard growth in poor countries and encourage growth in richer places. Voitchovsky (2005) examined inequality between the poor and the rich, and a negative link was found for the poor while a positive link was found for the rich. In a similar but different study, Weide and Milanovic (2018) examined the link at different points of the income distribution using micro-census data from United States (US); the reported evidence indicates that high levels of inequality reduce income growth for the poor but increase growth for the rich. In general terms, studies focusing on the short- and/or medium-run tend to find a positive relationship. In contrast, studies where longer periods were analysed tend to find a negative relationship (Atems and Jones, 2015). This is probably because the theoretical models/arguments mentioned above are likely to apply only in the long run (Knowles, 2005).

Surprisingly, although the literature is extensive at the national level, it is scarce at the regional level due to the limited number of available data (for an excellent discussion, see Royuela et al. 2019). Most of the existing studies are limited to American states and yielded mixed evidence (see, for instance, Partridge, 1997, 2005; Panizza, 2002; Frank, 2008). For instance, Partridge (1997) found a positive link between income inequality and growth using US data while Panizza (2002) found no linkage using panel data methods. In a more recent study, Partridge (2005) examined the sensitivity of the link to different econometric methods, highlighting the fact that cross-sectional methods tend to reveal a positive relationship while the link is less clear when using panel techniques. One of the few empirical studies for the EU regions is conducted by Perugini and Martino (2008). Their results imply that more regional inequality can generate higher regional growth in the short and medium-term. However, the significance and strength of the link decreases in spatial econometric models.

As mentioned above, there is a group of empirical studies wherein investigators have tried to examine short-term dynamics, i.e. whether regional inequality is expected to rise in periods of economic expansion and decrease in periods of contraction (pro-cyclicality) or vice-versa (anti-cyclicality). Most studies of this type have provided evidence in favour of pro-cyclicality. For instance, Chatterji and Dewhurst (1996), who examined both regions and counties of Great Britain, Terrasi (1999), who investigated the Italian regions, Petrakos and Saratsis (2000), who used Greek regional data (NUTS 3), Azzoni (2001) for the case of Brazil, Petrakos et al. (2005) for EU member states, and Rodríguez-Pose and Fratesi (2007), who examined southern European countries, have all concluded that regional inequality decreases during recessions and increases in periods of boom of the national economy. However, results from a few studies suggest that regional disparity is a phenomenon with anti-cyclical behaviour that increases in periods of recessions or shocks and decreases during economic booms (see for instance Dunford, 1993, for the EU level).

More recently, economic crisis provides an important opportunity to examine the short-term dynamics between inequality and growth. For instance, Brakman et al. (2015) and Capello et al. (2015) for the EU, Monastiriotis (2011) and Petrakos and

Psycharis (2016) for Greek regions, and Đokić et al. (2016) for Croatia have found evidence of anti-cyclicality during the recent period of economic crisis. In a more recent study, Royuela et al. (2019) found a negative association between inequality and economic growth since the start of the economic crisis by examining combined household survey data and macroeconomic databases covering over 200 comparable regions in 15 OECD countries. However, results from other recent studies reveal the opposite pattern; for instance, OECD (2014), Palaskas et al. (2015), and Artelaris (2017), for the case of Greece as well as Davies (2011), Christopherson et al. (2013), Donald et al. (2014), European Commission (2014) for the EU have presented evidence that the less advanced and/or urbanized regions are relatively more resilient during a period of crisis.

Finally, the direction of causality of the inequality-growth nexus is uncertain and still under debate. As a result, it is important to review the literature that examines reverse causation, i.e. how causation goes from growth to inequality (inverted- U hypothesis). The evidence presented in the literature is mixed because some studies confirm the inverted-U hypothesis while others do not. In the seminal study of Williamson (1965), regional data from several countries appeared to confirm this hypothesis. Williamson also explored the relationship using county data from the USA. However, in this case, as Williamson suggested, a linear rather than nonlinear (negative) relation is expected because US states were highly developed. The econometric results confirmed this negative relationship and Williamson concluded that the level of within-state disparities of income is greater in less-developed states or 'the states with lowest income per capita are typically those with the greatest inter-county inequality' (Williamson, 1965).

Surprisingly, only a few empirical studies have examined this link in the following years, primarily because spatial collection of data for less advanced countries is difficult and/or historical data for single countries is scarce. In a recent cross-country study, Lessmann (2014) used unique panel data of spatial inequality in 56 countries at different stages of economic development to confirm the existence of an inverted-U, suggesting that spatial inequality increases again at very high levels of economic development. For single countries, Wang and Ge (2004) conducted an econometric analysis for less-developed regions of China and confirmed the inverted U-hypothesis. In addition, Janikas and Rey (2005) used spatial data from the USA in a more recent study; they found that regional inequality increases in States with higher income levels, implying there is a positive relationship between inequality and development.

A few studies have been conducted using EU data. Davies and Hallet (2002), for instance, examined the U-hypothesis for specific EU countries. As the authors pointed out, beyond a particular level of per capita GDP, regional inequality may decrease, resulting in a stabilization of spatial imbalances. Petrakos et al. (2005) provided evidence that relatively advanced countries are characterized by a negative relation between the level of regional inequality and the level of development. Finally, more recently, Artelaris and Petrakos (2016) examined the relationship between intraregional spatial inequality and regional income level in the EU-27 regions. Their results did not support the inverted-U hypothesis but revealed the presence of a U pattern, implying a negative relationship between intraregional spatial inequality and regional income at lower income levels and a positive relationship at higher income levels.

3. INEQUALITY AND THE ROLE OF SPACE

3.1 Inequality measures and the geographical position of data

The issue of inequality has inspired a long tradition of theoretical and empirical research, and this issue has become more relevant in the political agenda. In the EU, inequality is a crucial point that influences regional policies. In fact, after its enlargement, the EU had to face with large inequality between regions, and the financial and economic crises exacerbated existing disparities.

At the European level, reducing disparities between countries, regions, and social groups has inspired the EU Cohesion Policy. This policy provides support to less-developed regions, helping them to improve productivity and to create better living conditions. The amount of inequality, its mechanisms, and consequences have been subjects of many studies. For a review of the main contributions in the literature, see, among others, Heshmati (2006) and Cowell (2011). Analysing regional inequality generally implies considering the differences across regional GDP rather than the income differences between individuals or households within a regional economy (Rey and Janikas, 2005).

The study of regional inequality requires considering additional issues that are related to the nature of georeferenced data. In fact, data on regional GDP are collected with reference to geographical units that cannot be considered as independent entities. In fact, geographical proximity among regional units could cause their economic behaviour to become similar. In the literature, similarity of observations of a given phenomenon, that are collected from nearby locations, is known as spatial dependence (Anselin, 1988). Spatial dependence, and hence the similarity of observations from neighbouring regions, is supported by the First Law of Geography, which states that '*everything is related to everything else, but near things are more related than distant things*' (Tobler, 1970).

According to these considerations, the spatial position of data could impact on regional inequality. In particular, the geographical position of data could determine similarities or dissimilarities across regional GDP. In other terms, because of spatial dependence, regional GDPs might assume similar values when observed in neighbouring regions or might exhibit differences, giving evidence of a negative spatial association. Most of the indices used to measure regional inequality discard the geographical position of data, and thus, the potentially implied spatial dependence among observations. This indicates that very different geographical distributions of GDP values provide the same inequality measure. This property of inequality measures is called 'anonymity' (Bickenbach and Bode, 2008). Because of the anonymity condition, inequality measures, such as the Gini index and entropy-based measures, fail to capture differences in the geographical position of data. This implies that inequality measures cannot account for relevant interconnections and dependencies. An illustration of the anonymity condition is shown in Figure 1.

Figure 1(a) shows the real distribution of regional GDP across the 1343 NUTS 3 EU regions belonging to the 28 EU Member States (the 5 overseas department of France are excluded from the analysis).

The values depicted in the map show GDP per worker for the year 2019. Darker colours indicate higher GDP per worker, and vice versa for lighter colours. Source of data is Cambridge Econometrics database that is now available free of charge from the new Annual Regional Database of the European Commission's Directorate General for Regional and Urban Policy (ARDECO, https://ec.europa.eu/knowledge4policy/territorial/ardeco-database_en) maintained by the European Commission's Joint Research Centre.

A measure of inequality, such as the Gini index, G, is reported, along with a measure of spatial dependence (i.e. Moran's I). The Gini index G[2] can take values from 0 to 1. Values near 0 suggest equality, while values near 1 indicate strong inequality.

[2] Let x_i be an observation of the random variable X observed on unit i, where $i = 1,2,\dots,N$. The Gini index G can be expressed as follows:

$$G = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}|x_i - x_j|}{2N\mu_x}$$

Moran's I[3] is a measure that accounts for the similarity (dissimilarity) of values across space. Positive values of I indicate the observations of the variable under consideration exhibit positive spatial autocorrelation, i.e. that is similar values tend to cluster together in space. Negative values of I indicate the presence of negative spatial autocorrelation, i.e. high (low) values of the variable are surrounded by low (high) values of the same variable. Calculating Moran's I requires assuming a proximity criterion among regional units. In the current application, for each regional unit, we identify the neighbouring regions based on the k nearest neighbours' criterion[4], with $k = 7$.

---

where $\mu_x = \sum_i x_i /N$ is the mean value of the observations x_i.

[3] Moran's I is computed for values x_i of X as follows:

$$I = \frac{N}{\sum_i \sum_j w_{ij}} \frac{\sum_i \sum_j w_{ij}(x_i - \mu_x)(x_j - \mu_x)}{\sum_i (x_i - \mu_x)^2}$$

where w_{ij} is the element of the matrix $\mathbf{W}$ expressing the proximity relationships between geographical units.

[4] Consider the set $N_k(i)$ that contains the k closest units to i. For each k, the k- nearest neighbour matrix $\mathbf{W}$ has $w_{ij} = 1$ if $j \in N_k(i)$ and $w_{ij} = 0$ otherwise.

Figure 1 – Inequality and spatial dependence of GDP per worker across NUTS 3 EU regions (2019): (a) real spatial distribution of regional GDPs and (b) spatial distribution obtained by randomly assigning GDP values to different regions. Source: Own Elaboration on ARDECO database.

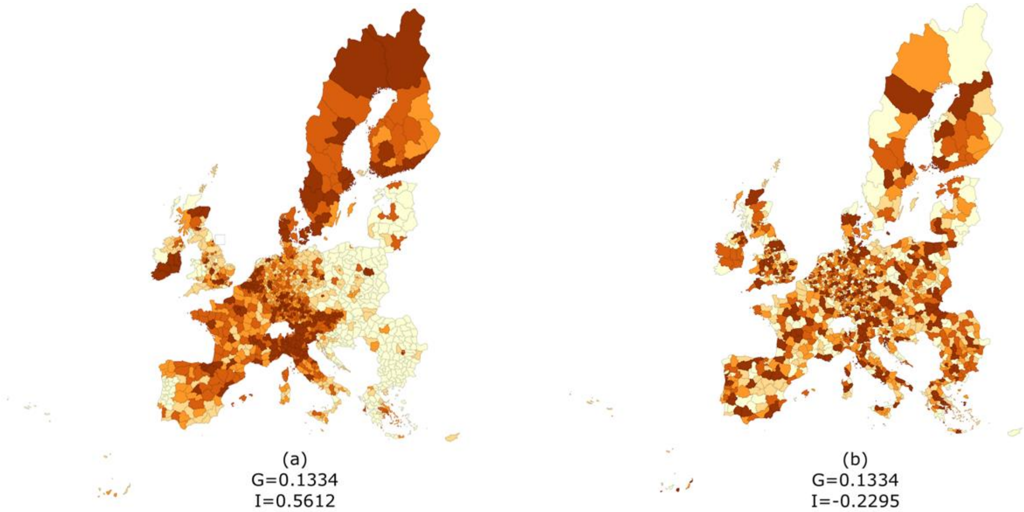


Figure 1 (b) shows the geographical distribution that is obtained by randomly assigning the values of GDP per worker across the considered EU NUTS 3 regions. This geographical distribution is constructed for comparison with the real distribution in (a). The indices G and I are calculated for this distribution and reported in the figure.

Some differences between the geographical distributions in (a) and (b) emerge. In particular, the distribution in (a) appears more clustered with respect to the distribution in (b) that appears more dispersed. In fact, the distribution in (a) is characterized by a high level of positive spatial association, as measured by Moran's I. More scattered values are reported in the distribution in (b) that is characterized by negative spatial association (i.e. a negative value of Moran's I). However, both the distributions in Figure 1 are characterized by the same level of inequality, as measured by the Gini index. This result confirms that this inequality measure is not sensitive to spatial permutations of the data.

3.2 Measurement of inequality and spatial dependence

Over the past decades, some contributions in the literature have highlighted the importance of combining measures of inequality with measures of spatial dependence in order to account for the spatial position of data in assessing inequality.

Focusing on spatial concentration, Arbia (2001) identified two different aspects of this phenomenon: such as a-spatial variability, which is invariant to permutations, and polarization, which refers to the geographical position of observations. To summarize these different aspects of spatial concentration, the author suggested combining measures of a-spatial concentration, such as the Gini index, with measures of spatial autocorrelation (like Moran's I) or measures of local association that measure polarization. The author presented an empirical example using Italian employment data in manufacturing industries to show the different aspects of concentration that emerge from considering these measures. A number of different approaches for combining complementary information derived from these measures have also been proposed.

Arbia and Piras (2009) introduced a new class of statistics that integrates the ideas of a-spatial concentration and polarization. This class of statistics is defined as the linear correlation coefficient between a random variable X across N spatial units, and a random variable X^*, corresponding to the permutation of the N values of X that maximizes a measure of positive spatial association. Formally, we have:

$$\lambda = \frac{\sum_{i=1}^{N}(X_i-\mu)(X_i^*-\mu)}{\sum_{i=1}^{N}(X_i-\mu)^2} \tag{1}$$

with μ denoting the mean of X_i. This class of measures accounts for both a-spatial concentration (the variance in the denominator) and spatial correlation (the numerator). The authors discussed the properties of this measure and an approximate sampling theory. They also identified some possible extensions of the proposed statistic, such as its use for comparing the concentration of the same variable measured over two different time periods or in two different countries, and its extension to other measures of inequality. The proposed measure has been used with data on the growth rate of output in the service sector for NUTS 2 EU regions.

Rey and Smith (2013) considered the Gini index G in relative mean difference form and rewrote the sum of all pairwise differences as the sum of absolute differences between pairs of neighbouring observations and absolute differences between pairs of non-neighbouring observations. Formally, we have:

$$G = \frac{\sum_{i=1}^{N}\sum_{j=1}^{N} w_{ij}|x_i-x_j|}{2N^2\mu} + \frac{\sum_{i=1}^{N}\sum_{j=1}^{N}(1-w_{ij})|x_i-x_j|}{2N^2\mu} \tag{2}$$

where w_{ij} denotes the generic element of a binary spatial weight matrix expressing the proximity relationship between locations i and j. This decomposition allows the identification of a neighbouring component (i.e. the first term on the right side) and a non-neighbour component (i.e. the second term on the right side) of the Gini index, and reveals as this index nests a measure of spatial autocorrelation. In fact, when the degree of positive spatial autocorrelation increases, the second term should increase relative to the first because the association between observations in space would be in effect. The result is opposite in the presence of negative spatial autocorrelation (Rey and Smith, 2013). To illustrate their measure, the authors calculated the index G, and its neighbouring and non-neighbour components, for US per capita income. They found that overall inequality is primarily explained by its spatial component.

The contribution of Márquez et al. (2019) is focused on an entropy-based measure, such as the Theil index T. Just like the Gini index G, this statistic is insensitive to the spatial position of data. Following the idea that spatial dependence is important in modelling regional inequality, the authors identified a decomposition of the Theil index T into its spatial and non-spatial components. Specifically, to illustrate the influence of spatial association on inequality, the authors identified a neighbourhood Theil index that provides a measure of inequality that only considers information from neighbouring regions. By subtracting the Neighbourhood Theil from the conventional Theil, a Specific Theil index that accounts for non-spatial inequality was defined. These measures are applied to data related to the NUTS 3 EU regions. The empirical evidence showed that regional inequality is primarily determined by neighbourhood inequality.

A further contribution focused on considering the impact of spatial dependence on an inequality measurement was recently proposed by Panzera and Postiglione (2020). The authors introduced a new index, γ, based on the Gini correlation measure (Schechtman and Yitzhaki, 1987). This new measure considers both inequality and spatial autocorrelation. Specifically, this measure is defined as the Gini correlation between the variable $\boldsymbol{Y}$, expressing GDP per capita observed on N regional units, and the same variable spatially lagged $\mathbf{WY}$, with $\mathbf{W}$ denoting the spatial weight matrix that summarizes the proximity relationships between regional units. The value assumed by the variable $\mathbf{WY}$ expresses, for each region, the average GDP per capita observed in neighbouring regions. Following the definition of the Gini correlation (Schechtman and Yitzhaki, 1987), the index γ is defined as follows:

$$\gamma = \frac{Cov(Y, R_{WY}/N)}{Cov\ (Y, R_Y/N)} \tag{3}$$

where R_Y and R_{WY} denote the ranks of $\mathbf{Y}$ and $\mathbf{WY}$, respectively. These ranks, assigned as 1 to the lower value of the variables and N to the higher value, are divided by the number of geographical units N and represent empirical estimates of the cumulative distribution functions of $\mathbf{Y}$ and $\mathbf{WY}$, respectively. In equation (3), the numerator corresponds to the Gini covariance between $\mathbf{Y}$ and $\mathbf{WY}$ and provides a measure of spatial autocorrelation, while the denominator expresses the variability in regional GDP per capita. In this sense, the measure in (3) permits the relative contribution from spatial dependence to regional inequality to be quantified.

Taking advantage of the correspondence between the Gini index G and twice the covariance between the variable and the rank of the variable divided by its mean (Lerman and Yitzhaki, 1984; Schechtman and Yitzhaki, 1987), the index in (3) can be rewritten as:

$$\gamma = G_S/G \tag{4}$$

where

$$G_s = \frac{2Cov(Y, R_{WY}/N)}{\mu_Y} \tag{5}$$

is a spatial Gini that varies between $-G$ and G (Dawkins, 2004), and

$$G = \frac{2Cov(Y, R_Y/N)}{\mu_Y} \tag{6}$$

is the Gini index G, expressed in terms of a covariance. As an additional result, the authors derived the following decomposition of G:

$$G = G_s + G_{ns} \tag{7}$$

where G_{ns} is a non-spatial Gini index that captures the component of inequality that is not due to a specified pattern of spatial dependence. The component G_{ns} is such that $0 \leq G_{ns} \leq 2G$.

The index defined in (3) and (4) ranges from -1 to 1. When the ranking of $\mathbf{WY}$ is identical to the original rank of $\mathbf{Y}$, $G_s = G$ and $\gamma = 1$, thus indicating overall inequality is completely explained by the specified spatial dependence. As the ranking of the regional GDPs (i.e. $\mathbf{Y}$) becomes more dissimilar to the ranking of average GDPs in neighbouring regions (i.e. $\mathbf{WY}$), the spatial component of inequality decreases and eventually reaches its minimum value of $-G$ when the average GDPs in neighbouring regions are ranked as opposite with respect to the original order of regional GDPs. In this case, the non-spatial component of inequality reaches its maximum value of $2G$ and $\gamma = -1$. When $\mathbf{Y}$ and $\mathbf{WY}$ are uncorrelated, we have $G_s = 0$, $G = G_{ns}$, and $\gamma = 0$, indicating overall inequality is completely explained by its non-spatial component.

All the different contributions proposed in the literature highlight the opportunity to consider spatial dependence when analysing regional inequality. Considering this aspect is especially relevant when assessing regional inequality, in order to guide policies at subnational or national levels. An in-depth understanding of inequality is also required when one intends to consider how regional inequality impacts regional growth. Some proposals to assess this impact and, in general, the relationship between inequality and economic growth, are presented in the following Sections.

4. INEQUALITY AND ECONOMIC GROWTH

4.1 Model specification

As discussed in Section 2, the link between inequality and growth has been widely analysed theoretically and empirically, revealing mixed results. Some recent studies found different results for more-developed and less-developed regions within the EU. A positive relationship between intraregional inequality and the level of regional income was identified, but only for more developed regions (Artelaris and Petrakos, 2016). This result is primarily consistent with evidence obtained at country level (Barro, 2000; Voitchovsky, 2005).

Analysing the differences in growth, and the complex relationship between inequality and growth, could offer useful insights from a policy perspective. Such an analysis would facilitate the assessment of the effectiveness of the EU Cohesion Policy, which is concerned with removing potential barriers to growth and to promote the process of development in various regions within the EU.

The spatial nature of data should also be considered to properly analyse the relationship between inequality and growth. In fact, spatial interactions among regional units could impact regional disparities, and the link between inequality and growth can be influenced by spatial effects. Some contributions in this direction have been proposed in the literature (De Dominicis, 2014).

In this section, we describe a spatial cross-sectional model that relates economic growth with intraregional inequality and a number of other control variables. The reverse causality relationship from growth to inequality is not investigated in our analysis as our objective is to assess the role of regional disparities in driving economic growth in EU regions. For a recent contribution on assessing the impact of regional income levels on inequality using spatial econometrics tools see Artelaris and Petrakos (2016). In our specification, the inequality variable is measured at the beginning of the period under investigation to mitigate the concern about feedback effect of GDP dynamics on regional inequality (Cingano, 2014).We focus on intraregional inequality, which is the within-group component of inequality, since it provides evidence of the contribution from each region to overall disparities observed in the EU at the regional level.

The considered specification is defined as a spatially augmented Solow growth model (Solow, 1956). A spatially augmented version of the Solow growth model has already been proposed by Ertur and Koch (2007), who extended this model by explicitly accounting for technological interdependence among geographical units. Interdependence is assumed to spread out through proximity relationships, and the resulting empirical model corresponds to the spatial Durbin model (SDM). According to this specification, the GDP per worker growth rate for a given geographical unit is expressed as a function of the initial GDP per worker, the investment in physical capital and the real working population growth in the same unit. The formulated assumption determines the introduction, as additional explanatory variables in the model, of the spatial lag of the independent variables, as well as of the spatially lagged dependent variable. This implies that the GDP per worker growth rate in a region depends on the aforementioned independent variables in the region, on the same variables in neighbouring regions, and on growth in GDP per worker in neighbouring regions. This spatially augmented version of the Solow model can be expressed as follows:

$$g_i = \beta_0 + \beta_1 \ln y_{it-T} + \beta_2 \ln s_i^k + \beta_3 \ln(n_i + l + k) + \rho \sum_{j=1}^{N} w_{ij} g_j + \rho_1 \sum_{j=1}^{N} w_{ij} \ln y_{jt-T} + \rho_2 \sum_{j=1}^{N} w_{ij} \ln s_j^k + \rho_3 \sum_{j=1}^{N} w_{ij} \ln\left(n_j + l + k\right) + \varepsilon_i \qquad (8)$$

where $g_i = \frac{1}{T} \ln\left(\frac{y_{it}}{y_{it-T}}\right)$ is the GDP per worker growth rate, with T denoting the number of time periods under investigation, and y_{it} and y_{it-T} expressing the GDP per worker at the final and initial periods, respectively. s_i^k is the fraction of output in region i invested in physical capital, n_i indicates the growth rate of the working population, l is the growth rate of technology, k is the depreciation rate of capital, and $\varepsilon_i \sim N(0, \sigma^2)$ are error terms. The model includes the aforementioned variables and their spatial lags that are identified by introducing the proximity matrix $\mathbf{W}$. For each variable observed in region i, the spatial lag corresponds to the average of the observations for the same variable collected from neighbouring regions. The proximity relationship between regional units can be defined using contiguity criteria, physical distances, and economic and social distances. Combinations of these criteria are also possible (for some useful indications on this topic see Conley and Topa, 2002). The specification in (8) is known in the literature as the SDM.

The annual speed of convergence implied by the parameter β_1 is given by $\lambda = -\ln(1 + \beta_1 T)/T$, as in the classical Solow model (Elhorst et al. 2010). In contrast to Ertur and Koch (2007), we estimate our model only under the condition that the speed of convergence is identical for all economies. Possible differences in the speed of convergence among regions are accounted for in our empirical analysis through the identification of clusters of regions with different levels of development (see Section 6.2).

Under the assumption of identical speed of convergence, the GDP per worker steady state can be expressed, in matrix notation, as (Elhorst, 2001; Elhorst et al. 2010**)** $\mathbf{y}^* = (\mathbf{I} + \frac{\rho_1}{\beta_1}\mathbf{W})^{-1}\left\{-\frac{\beta_o}{\beta_1} - \frac{\beta_2}{\beta_1}\mathbf{s} - \frac{\beta_3}{\beta_1}\mathbf{n} - \frac{\rho_2}{\beta_1}\mathbf{Ws} - \frac{\rho_3}{\beta_1}\mathbf{n}\right\}$ where $\mathbf{y}^*$**,** $\mathbf{s}$ and $\mathbf{n}$ are the vector of the steady state values for GDP per worker, and the vectors of the observations related to $\ln s_i^k$ and $\ln(n_i + l + k)$, respectively. The matrix $\mathbf{W}$ is defined as above, and $\mathbf{I}$ express an $N \times N$ identity matrix. The coefficients $\frac{\rho_2}{\beta_1}$ and $\frac{\rho_3}{\beta_1}$, if significantly different from zero, indicate that the steady state position of a particular economy is related to the variables $\ln s_i^k$ and $\ln(n_i + l + k)$ in its neighbouring economies (Elhorst et al. 2010).

In order to investigate the impact of inequality on growth, we extend the model developed by Ertur and Koch (2007) by expressing the investment in physical capital (s_i^k) as a function of an inequality measure ($Ineq_i$). This assumption is consistent with De Dominicis (2014) and is motivated by the idea that higher inequality stimulates higher investment in physical capital.

This assumption found support in some early and recent contributions. Specifically, following standard economic theory, saving and investment rates can be assumed identical, and a high initial inequality of income favours high saving because rich people have a higher propensity to save compared to poor people (Kaldor, 1956). Given this argument, more inequality could be associated with faster growth. If the effect of inequality is explained by the savings gap between richer and poorer economies, a potentially negative effect on growth could be associated with redistribution of resources, which corresponds to lower inequality (Thorbecke and Charumilind, 2002). Following this line, Galor (2000) argues that high income inequality has a positive impact on saving rates, and Koo and Song (2016) reported a positive correlation between inequality and aggregate savings.

We use such a model to investigate the relationship between inequality and growth by analysing how growth of GDP in a region is influenced by inequality in a given region and by inequality in neighbouring regions.

The proposed spatially augmented model for analysing the impact of inequality on growth can be expressed as follows:

$$g_i = \beta_0 + \beta_1 \ln y_{it-T} + \beta_2 \ln Ineq_i + \beta_3 \ln(n_i + l + k) + \rho \sum_{j=1}^{N} w_{ij} g_j + \rho_1 \sum_{j=1}^{N} w_{ij} \ln y_{jt-T} + \rho_2 \sum_{j=1}^{N} w_{ij} \ln Ineq_j + \rho_3 \sum_{j=1}^{N} w_{ij} \ln\left(n_j + l + k\right) + \varepsilon_i \qquad (9)$$

where $\ln Ineq_i$ represents an inequality measure that replaces the measure of investment in physical capital, according to the aforementioned assumption. All other variables have the same definitions as in equation (8).

The model in (9) allows one to test the conditional β-convergence hypothesis, which occurs when the parameter associated with the initial GDP per worker (i.e. β_1) is negative and statistically significant. This specification affirms that inequality in a given region has a positive impact on growth in the same region and, in contrast, is negatively influenced by inequality in neighbouring regions. This model is empirically tested on regions at the NUTS 2 level (see Section 6).

Our model differs from that proposed by De Dominicis (2014) regarding the definition of the inequality variable. While inequality is defined as the within-region component of the Theil index T in De Dominicis (2014), we define inequality as the non-spatial component of intra-regional inequality, which is derived using the decomposition of the Gini index (see Section 4.2 for further details). Taking advantage of the spatial decomposition of the Gini index of inequality recently proposed in the literature, we aim at identifying the component of intraregional inequality that is not influenced by spatial dependence. This component can be viewed as the idiosyncratic component of inequality, thus its relationship with regional growth is analysed.

The model in (9) could be estimated, and interpreting the model parameters requires considering the direct, indirect, and total impacts (see, for further details, Appendix 1 of Deliverable D 3.2). The direct impact is the effect that a change in an exploratory variable in a single unit produces on the dependent variable within the same unit. The indirect impact refers to the effect that a change in an explanatory variable in a given unit produces on the dependent variable in neighbouring units. Finally, the total impact is defined as the sum of the direct and indirect effects. LeSage and Pace (2009) developed summary measures expressing the average direct, indirect, and total impact. A significative indirect impact gives evidence of the presence of spatial spillovers.

The model defined in (9) only accounts for spatial dependence. A further issue in geographical analysis concerns simultaneous treatment of spatial dependence and spatial heterogeneity.

Spatial heterogeneity refers to the instability of relationships and behaviours across space; it can be modelled by assuming heterogeneous coefficients in a regression model. A way to model spatial heterogeneity requires defining spatial regimes as clusters of spatial units that share some similar characteristics. The identification of spatial regimes can be based on some a priori criterion, such as the initial GDP (Ertur et al. 2006) or endogenous criteria (Durlauf and Johnson, 1995; Postiglione et al. 2010, 2013).

The definition of spatial clusters in EU regions could be also based on levels of development (Iammarino et al. 2019). Classifying regions according to the level of development is particularly relevant for EU regions, since the EU policies are mainly targeted to less-developed regions. Thus, identifying such clusters of regions could help to assess the effectiveness of EU policies targeted for these regions.

In this report, we decided to follow this last approach by defining two groups of NUTS 2 European regions based on the division between less-developed regions (first group), and transition and more-developed regions (second group), which were identified according to the EU Cohesion Policy for the 2014-2020 funding period. The choice of NUTS 2 regions as units of analysis is mainly motivated by the relevance that this geographical level assumes in the definition of EU policies. In fact, the eligibility to funds from the EU Cohesion Policy is primarily evaluated at the NUTS 2 level[5].

The economic model estimated for these two different groups of EU regions represents a heterogenous specification of model (9). In this heterogenous version, the parameters are different across the two clusters of regions. In contrast, the parameters in the homogenous equation (9) are assumed equal for all the spatial units under investigation.

The proposed analysis assumes great relevance in the EU context. First, the explicit treatment of spatial effects allows one to account for interactions between regions that are linked by a range of economic flows within the context of EU. Interregional trade and mobility, transfer of technology and all other types of regional spillovers are likely to be more intense among neighbouring regions, thus motivating the use of the described spatial econometric definition. Second, the proposed definition is likely to provide useful insights for policy makers within the EU. Analysing the conditional convergence process could suggest relevant policy interventions. In fact, while unconditional convergence implies a 'spontaneous' process of reduction of disparities across regions, conditional convergence implies a catch-up process after controlling for specific regional characteristics, that are represented by intra-regional inequality and working population growth in this analysis. These variables could be the subject of policy interventions to increase regional growth.

Finally, accounting for differences between groups of regions allows policies that embrace the goal of promoting development in different types of regions to be defined. This is consistent with the idea of defining place-sensitive policies (Iammarino et al. 2019) that are focused on different starting points and different mixes of instruments for distributing development in different groups of regions. Such policies overcome the limitations of Europe-wide, place-neutral, or general-purpose policies that appear as inappropriate for addressing spatially uneven development.

4.2 The inequality variable

The inequality variable ($Ineq_i$) must be chosen for use in the model described in Section 4.1. We consider a measure of within-region disparities that, for each NUTS 2 region, is computed over the NUTS 3 regions of which it is composed. The choice to consider within-region inequality is motivated by the idea of introducing the contribution from each region to overall inequality in our model. This idea is also consistent with De

[5] https://ec.europa.eu/eurostat/web/nuts/background

Dominicis (2014), who considered the within-group component using the Theil index. Unlike De Dominicis (2014), we consider the Gini index of inequality. We focus on this measure of inequality as the recent literature has revealed that the Gini index of inequality nests a measure of spatial dependence (Rey and Smith, 2013; Panzera and Postiglione, 2020). This facilitates the component of inequality that is not influenced by spatial interactions among regional units to be isolated. We apply this decomposition after identifying the within-region component of inequality.

The Gini index of inequality G is traditionally decomposed into a within-group component, a between-group component, and an overlapping term (Mookherjee and Shorrocks, 1982; Shorrocks and Wan, 2005). The within-group component, G_{within}, corresponds to a weighted average of the inequality within each subgroup; it expresses the contribution from each regional unit to the overall inequality as measured by the index G. The between-group component, $G_{between}$, corresponds to the value of the Gini index when the income of all units is replaced by the mean GDP per worker for the subgroup to which they belong. The interaction term, R, is a residual that is necessary to maintain the identity $G = G_{within} + G_{between} + R$.

In our empirical application, we consider the within-group inequality for each NUTS 2 region. This is calculated using the NUTS 3 regions that belong to a particular NUTS 2 region.

Considering N subgroups corresponding to the NUTS 2 EU regions, let the GDP per worker mean and the working population share for region i be respectively μ_i and p_i, where $i = 1,2, \dots, N$. The within-region component of the Gini index of inequality can be expressed as

$$G_{within} = \sum_{i=1}^{N} p_i^2 \frac{\mu_i}{\mu} G_i \qquad (10)$$

where μ denotes the mean of the GDP per worker for the overall distribution, and G_i denotes the Gini index of inequality for region i.

The spatial decompositions of the Gini index described in Section 2 (compare equation (2)) can be applied to the Gini indices G_i calculated for each region i. This allows the spatial and non-spatial components of within-region inequality to be distinguished. Decomposing inequality in its spatial and non-spatial components allows to appreciate the component of inequality that is influenced by positive spatial autocorrelation, and the idiosyncratic component of inequality, which is mainly due to specific location characteristics.

Specifically, we will focus on the latter, which can be interpreted as the within-region inequality component that is not influenced by the spatial dependence effect. Isolating this component is relevant because spatial dependence, giving rise to clusters of similar values, could lead to increasing inequality (Panzera and Postiglione, 2020). Therefore, the identification of the non-spatial component of inequality facilitates a better understanding of the phenomenon and can offer useful insights to address the impact of inequality on regional growth.

However, the spatial dependence effect is not discarded from the analysis. In fact, the inequality variable is included in the analysis by considering both its value in a particular region and the average value in the neighbouring regions (i.e. a spatially lagged inequality variable).

As discussed in Section 3.2, different spatial decompositions of the inequality measures have been proposed in the literature. Some of these decompositions relate to the Gini index of inequality. By applying the decomposition proposed by Rey and Smith (2013), the within-region component of the Gini index in (10) could be written as

$$G_{within} = \sum_{i=1}^{N} p_i^2 \frac{\mu_i}{\mu} NG_i + \sum_{i=1}^{N} p_i^2 \frac{\mu_i}{\mu} NNG_i \qquad (11)$$

where NG_i and NNG_i represent the neighbour and non-neighbour components of the inequality index calculated for each region i, respectively. The first term on the right side of equation (11) identifies the component of inequality that is not influenced by positive spatial association, while the second term on the right side expresses the spatial component of inequality that varies along the same direction as the spatial autocorrelation (see Section 2).

The spatial decomposition proposed by Panzera and Postiglione (2020) is based on expressing the Gini index of inequality in terms of covariance (see equation (6)). This decomposition could be combined with the decomposition by population subgroups for the Gini index which also focuses on covariance-based formulas, leading to the identification of an intra-group, inter-group, and overlapping component indices (Yitzhaki, 1994; Giorgi, 2011).

We take the non-spatial component of the within-region inequality, which is defined in equation (11), as the inequality variable in our growth model. This implies that the inequality variable included in model (9) can be written as

$$\ln Ineq_i = \ln p_i^2 \frac{\mu_i}{\mu} NG_i \qquad (12)$$

where NG_i is the non-spatial component of the Gini indices calculated across spatial units (i.e. NUTS 3 regions) that are included in each NUTS 2 region under investigation. We consider the non-spatial component of the within-region inequality since it can be interpreted as the component that is not influenced by spatial interactions between regions. In this sense, this component can be considered as the idiosyncratic component of inequality, thus it allows one to appreciate the real extent of the regional economic contributions to overall inequality.

5. BEYOND INCOME INEQUALITY

5.1 Synthetizing a multidimensional phenomenon

In recent years, more authors have supported the idea that well-being and development cannot be measured through the use of a single indicator, such as GDP per capita (among others, see Stiglitz et al. 2009). Among different applications, composite indicators have been defined to measure human well-being for this reason (Sen, 1985; Nussbaum, 2000).

For example, the concept of poverty refers to multiple forms of inequity, sources of exclusion, and differences in living conditions that are essential to human dignity. To this end, multidimensional poverty indicators (MPI) are considered for measuring poverty (Alkire and Santos, 2010). Moreover, composite indicators of deprivation are adopted to point out areas that present major disparities (Pampalon and Raymond, 2000). Indicators of socio-economic deprivation are appealing for policy makers in order to develop local policies, tackle disparities, and reduce the gap between advantaged and disadvantaged areas (Havard et al. 2008).

Composite indicators estimate multi-dimensional concepts that cannot be captured by a single indicator and are developed to be key tools for policy makers (Bell et al. 2007; Padilla et al. 2014; Pampalon et al. 2012).

In dealing with a multivariate reality, building a composite indicator leads to a set of choices. In fact, the results may depend on the choice of variables used in the analysis. Hence, this decision is related to key features, such as what are the stakeholders to be engaged, the policy makers to be advocated, and the scenario in which the policies are to be implemented (Kuc-Czarnecka et al. 2020).

Synthesis of a single measure may be obtained by adopting different techniques. An important option regards the potential for compensation between different dimensions that are included in the composite indicator. While synthetizing variables, compensatory approaches allow one dimension to compensate other dimensions. Adding multiple inputs to build a composite indicator is a well-known example of this approach. However, non-compensatory approaches are used because compensation is more difficult. This feature has been considered in earlier studies, particularly in the field of well-being (De Muro et al. 2011). The geometric mean and Mazziotta-Pareto index are two of these non-compensatory techniques (Mazziotta and Pareto, 2019).

Another common issue regards synthesis of simple variables. In fact, indicators are obtained as a weighted average. Consequently, a relevant problem is to derive these weights. Three important classes of approaches to set the weights exists: data-driven (or statistical), normative, and hybrid.

Statistical weights are a direct function of the data used to represent reality and are not based on any prior judgement regarding trade-offs between different dimensions. This range of approaches are data-driven so that each variable is weighted based on the evidence gathered.

Conversely, normative approaches only depend on the value judgements about the trade-offs and are not based on the actual distribution of achievements in the society being analysed. Hence, those weights are defined by the knowledge that, for example, policy makers may have regarding the relevance of each variable on the panorama of the multivariate phenomenon.

Finally, hybrid approaches represent a third class of methodologies that exploit data-driven methods and depend on some form of (often prior) evaluation of potential trade-offs between variables (see Decancq and Lugo, 2013, for an extensive overview).

If the use of statistical techniques may bring puzzling results under certain circumstances, normative (and hybrid) approaches, based on elicitation of a number of experts (as in Delphi method), sometimes present problems that can cause for long time to bring up consensus (Di Cesare et al. 2020). For this reason, we rely on principal component analysis (PCA) in the current report, which is considered a starting point for deriving a composite indicator. In the second stage, some limitations of this approach will be considered and used to introduce a subsequent methodology. The latter addresses some of the problems revealed by the use of PCA for analysing geographically distributed data.

5.2 Principal component analysis for derivation of composite indicators

Given $\mathbf{X}$, a $N \times p$ matrix containing the centred[6] variables (i.e. the single indicators considered), PCA allows independent components (in the sense that each new dimension is orthogonal to the others) to be derived and used as composite indicators (Jolliffe, 2002).

The covariance matrix $\boldsymbol{\Sigma}$ can be expressed in terms of the original data as

$$\boldsymbol{\Sigma} = \frac{1}{n}\mathbf{X}^t\mathbf{X}. \tag{13}$$

where t indicates the transpose matrix. Observed variables are than transformed to obtain the composite indicators matrix $\mathbf{Z}$ (i.e. a new coordinate system), known as scores:

$$\mathbf{Z} = \mathbf{XA} \tag{14}$$

where $\mathbf{A}$ is a projection matrix given by the decomposition of the following variance-covariance matrix:

$$\boldsymbol{\Sigma} = \mathbf{A}\boldsymbol{\Lambda}\mathbf{A}^t. \tag{15}$$

Here, $\mathbf{A}$ is the matrix of the eigenvectors (i.e. loadings) of $\boldsymbol{\Sigma}$ and $\boldsymbol{\Lambda}$ is a diagonal matrix containing the eigenvalues of $\boldsymbol{\Sigma}$. The matrix of loadings $\mathbf{A}$ represents the statistical weights (i.e. the importance of each starting indicator) of the obtained composite indicators $\mathbf{Z}$.

In PCA, not all principal components are usually considered. The essence of the methodology consists in selecting a certain number of components that account for the largest part of the information included into the original data set. Therefore, in PCA, the composite indicators are usually obtained by considering only the first q components.

Variables are weighted with the values of the loadings contained in each of the considered eigenvectors in (14). Hence, the first composite indicator, for example, is the sum of all variables included in the original dataset weighted by the first column $\mathbf{a}_1$ of the loading matrix $\mathbf{A}$.

PCA relies on the assumption that the covariance matrix has the same structure in all localities included into the sample. This means ignoring the possible presence of

[6] A centred variable is obtained by subtracting its mean from the original variable.

structural differences caused by spatial heterogeneity that assumes the form of instability across the space of behaviours and relations.

Analysing the consequences of spatial heterogeneity is crucial for capturing dissimilarities in the phenomenon across a territory and allows one to model the differences that define context-related weights. Thus, a more appropriate tool for deriving a composite indicator may be represented by geographically weighted principal component analysis (GWPCA; Harris et al. 2011). In this technique, the variance-covariance matrix is estimated at each locality in order to pursue data decomposition while accounting for local diversities.

5.3 Geographically weighted principal component analysis for derivation of composite indicators

Let (u_i, v_i) be the coordinates for each spatial unit $i = 1,2, \dots, N$. Our purpose is to relax the hypothesis of homogeneity in the variance-covariance matrix $\mathbf{\Sigma}$, thus assuming the set of weights (i.e. loadings) could change based on location. In order to model spatial heterogeneity, we must assume that the covariance matrix is expressed as a function of a pair of coordinates for each region i as $\mathbf{\Sigma}(u_i, v_i)$. Consequently, according to Harris et al. (2015), we define a geographically weighted covariance matrix for each unit i as follows:

$$\mathbf{\Sigma}(u_i, v_i) = \mathbf{X}^{\boldsymbol{t}}\mathbf{C}(u_i, v_i)\mathbf{X} \tag{16}$$

where $\mathbf{C}(u_i, v_i)$ is a diagonal matrix including geographical weights generated with a kernel function. Weights are generated according to a distance-decay logic so that closer places will have larger weights used to calculate the local covariance matrix as those regions have stronger influence over each other. In particular, a bi-square kernel function is used in this application:

$$c_{ij} = \exp\left(-\frac{1}{2}\left(\frac{d_{ij}}{\phi}\right)^2\right) \tag{17}$$

where $c_{ij}s$ are the entries of $\mathbf{C}(u_i, v_i)$, d_{ij} is the geographical distance[7] between the centroids of localities i and j, and ϕ is the bandwidth.

These GWPCA weights depend on the choice of kernel function and on the bandwidth. The optimal bandwidth also represents a very important feature, and, for this reason, it

[7] In our analysis, d_{ij} is equal to the Euclidean distance.

must be determined carefully. Therefore, a cross-validation method for bandwidth selection is conducted (Harris et al. 2015). Further, the variance-covariance matrix estimated at the local level for each unit can be decomposed using the same procedure in PCA:

$$\mathbf{\Sigma} = \mathbf{A}(u_i, v_i)\mathbf{\Lambda}(u_i, v_i)\mathbf{A}(u_i, v_i)^{\boldsymbol{t}} \tag{18}$$

where $\mathbf{A}(u_i, v_i)$ is the matrix of local eigenvectors and $\mathbf{\Lambda}(u_i, v_i)$ is a diagonal matrix containing the local eigenvalues.

Finally, we can define a composite indicator z_{is}, where $s = 1, \dots, p$ at each (u_i, v_i), as follows (Cartone and Postiglione, 2020):

$$z_{is}(u_i, v_i) = \mathbf{x}_i{}^t \mathbf{a}_1(u_i, v_i) \tag{19}$$

where z_{is} is the value of the s^{th} composite indicator for the selected locality, $\mathbf{x}_i$ is a vector of p variables measured at location i, and $\mathbf{a}_1(u_i, v_i)$ is the first column of the local loading matrix $\mathbf{A}(u_i, v_i)$ that weighs the relevance of individual indicators.

In the derivation of composite indicators, a great advantage of GWPCA is the chance to consider local context in the structure of the covariance matrix. This allows the researcher to obtain context related to a composite indicator whose set of weights is based on local characteristics. In addition, this feature generally causes the composite indicator to better stress social and economic differences that characterize the EU in its heterogeneity and that are extremely important to identify critical local issues.

Thanks to GWPCA, composite indicators can encompass more aspects of socioeconomic dynamics, and policy makers could gain more information about the structure of inequality in each locality. This could also help policy makers more precisely track inequality across different locations with the aim of developing localised policies.

6. EMPIRICAL ANALYSIS ON DIFFERENT GEOGRAPHICAL SCALES

6.1 Spatial and non-spatial components of inequality in EU NUTS 3 regions

This section presents the evolution of economic inequality in EU regions. According to the recent literature discussed in Section 3.2, we consider the Gini index of inequality and its decomposition in spatial and non-spatial components. The measures proposed by Rey and Smith (2013) and Panzera and Postiglione (2020) are applied to GDP per worker data for the 1343 NUTS 3 regions belonging to 28 EU Member States (Austria, Belgium, Bulgaria, Cyprus, Croatia, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Italy, Ireland, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, and United Kingdom). Note that French overseas departments have been excluded from the analysis. The period under consideration ranges from 1995 to 2019. The ARDECO database was used as a data source. This new database contains a set of long time-series indicators for EU regions, as well as for regions in some EFTA[8] and candidate countries, at various statistical scales (NUTS1, NUTS2, and NUTS3).

Spatial inequality could be analysed on different geographical scales. However, investigating inequality at the regional level could offer useful insights from a policy perspective because the reduction of regional differences is an explicit concern in the EU Cohesion Policy. The period under consideration includes economic crises (i.e 2008) that led to widening of regional disparities in a number of economic indicators, such as employment and GDP. Territorial diversity can be better understood by focusing on data at a more detailed geographical level. This justifies the use of data at the NUTS 3 level. The choice of investigating inequality at a detailed geographical scale is also consistent with the idea of promoting place-based regional policies, where interventions can be tailored to suit different places and to respond to the structural opportunities and constraints in each region (Barca et al. 2012).

The problem of inequality enters the socio-political agenda as a concern of spatial justice. Patterns of dependence between regions with different economic and cultural strengths and issues of justice among regional units are the product of spatial dynamics. This supports the inclusion of a spatial dimension in examining regional inequality. Including the assessment of spatial dependence in our empirical analysis offers the opportunity to verify the role that proximity and spatial interactions among regional units play on regional inequality.

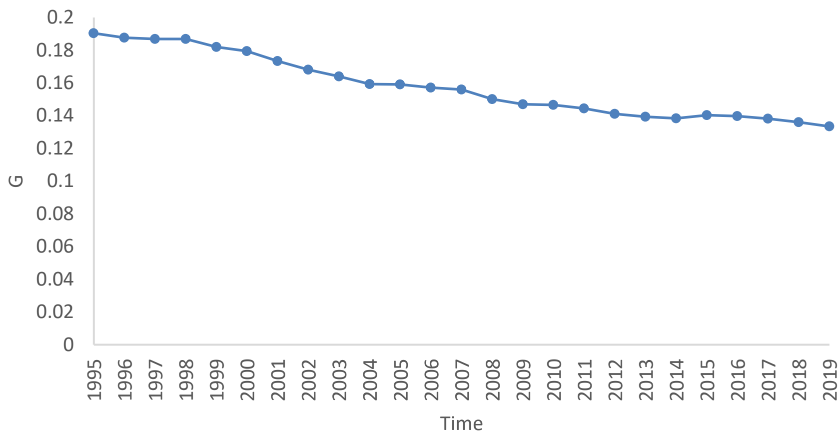
The Gini index of inequality calculated for GDP per worker from 1995 to 2019 for the aforementioned NUTS 3 regional units is shown in Figure Figure *2*.

As Figure 2 shows, the overall inequality among regions, as measured by the Gini index G, shows a general decrease throughout the period under investigation. The decline in economic inequality in EU is mainly due to the convergence process, by which GDP per worker in poorer regions increased relative to GDP per worker in richer regions (see Deliverable D 3.2). According to the considered inequality measure, disparities in the EU at the NUTS 3 level declined from around 0.19 in 1995 to around 0.13 in 2019. This

[8] The European Free Trade Association (EFTA) is a regional free trade area consisting of four countries: Iceland, Liechtenstein, Norway, and Switzerland.

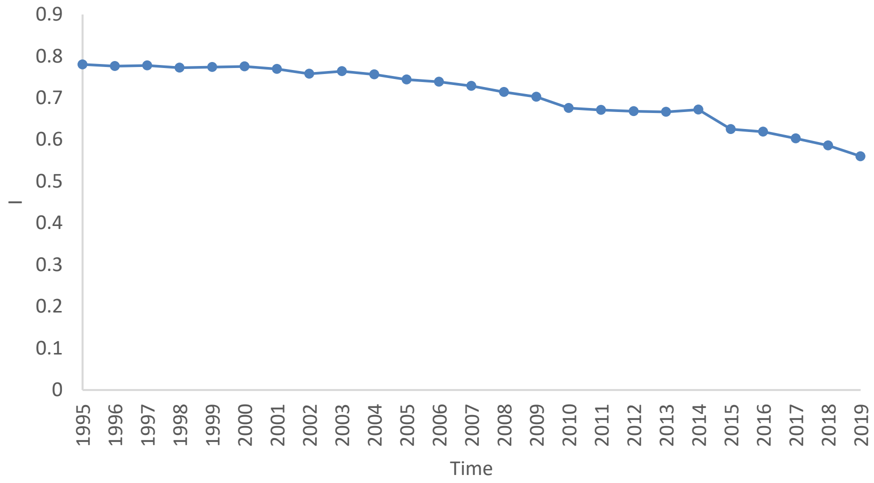
declining trend of inequality in EU 28 has been confirmed in previous studies (Butkus et al. 2018; European Commission, 2018). The decreasing trend characterized inequality also during the years of economic crises. This result could be explained by the asymmetric growth trajectories that the European regional economies experienced during crises. In fact, during this period, while the least developed regions have kept on converging toward richer regions, developed regions experienced a general income decline. Furthermore, in most of the EU, rural regions proved to be more resilient while metropolitan regions experienced a deterioration of economic performance (European Commission, 2014). This mixed territorial impact of economic crises led to a decrease in economic disparities.

Figure 2 – Gini index of inequality calculated for GDP per worker, NUTS 3 regions 1995-2019. Source: Own Elaboration on ARDECO database.



Thus, the declining trend highlighted in Figure 2 is due to the presence of an increasing similarity in GDP across regional units. A different emphasis on similarity is offered by Moran's I. This index focuses on clustering similar values across space, providing a measure of the spatial dependence of values observed in neighbouring regional units (for further details, see Section 3.1). Figure 3 shows Moran's I calculated for GDP per worker in the considered regional units from 1995 to 2019.

Figure 3 – Moran's I calculated for GDP per worker, NUTS 3 regions 1995-2019. Source: Own Elaboration on ARDECO database.



As Figure 3 shows, Moran's I exhibits a decreasing trend during the period under investigation. This indicates that the spatial dependence between the GDP values decreases from 1995 to 2019. This result is consistent with the decrease in inequality shown in Figure 2. In fact, a strong positive correlation between inequality and spatial dependence has been detected in some previous studies (Rey, 2004). However, as previously discussed, while a simple reshuffling of GDP values in the map does not impact the inequality measure, it does cause a change in Moran's I (see, in this regard, Section 3.1). This indicates the presence of differences among these two statistics and highlights the importance of considering these measures jointly to obtain findings that could not be obtained when they are used in isolation.

In order to appreciate how spatial dependence impacts the evolution of inequality, the Gini index of overall inequality is decomposed into its spatial and non-spatial components by following the approaches proposed by Rey and Smith (2013) and Panzera and Postiglione (2020). The spatial decompositions of the Gini index are calculated using the R-software.

Table 1 shows the neighbour (NG) and non-neighbour (NNG) components of inequality (Rey and Smith, 2013), while Table 2 shows the spatial Gini (G_s) and the non-spatial Gini (G_{ns}) indices calculated using the procedure presented by Panzera and Postiglione (2020). In both spatial measures, we consider a proximity matrix based on the k nearest neighbours with $k = 7$.

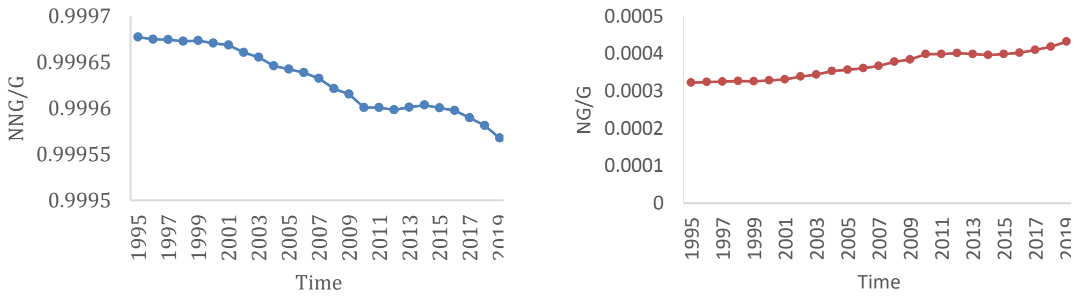
Table 1 – Gini index of inequality (G) and its neighbour (NG) and non-neighbour (NNG) components calculated for GDP per worker, NUTS 3 regions, 1995-2019. Source: Own Elaboration on ARDECO database.

Year	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**
G	0.1904	0.1877	0.1868	0.1869	0.1820	0.1795	0.1734	0.1682	0.1640
NG	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
NNG	0.1903	0.1876	0.1867	0.1868	0.1819	0.1794	0.1733	0.1681	0.1639
Year	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
G	0.1593	0.1590	0.1572	0.1559	0.1500	0.1470	0.1465	0.1443	0.1411
NG	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
NNG	0.1592	0.1589	0.1571	0.1558	0.1499	0.1469	0.1464	0.1442	0.1410
Year	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**		
G	0.1393	0.1383	0.1404	0.1398	0.1381	0.1360	0.1334		
NG	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001		
NNG	0.1392	0.1382	0.1403	0.1397	0.1380	0.1359	0.1333		

As previously mentioned, the NNG component of the Gini index can be interpreted as the spatial component of the index, which varies along the same direction as the positive spatial correlation. The values reported in Table 1 reveal the dominance of the NNG component of inequality. Furthermore, this component shows a decreasing trend, which is consistent with the decrease in the spatial autocorrelation among GDP values highlighted in Figure 3. The neighbour component of inequality, which expresses the non-spatial part, is near 0 and remains nearly stable throughout the period under investigation. The relative contribution from these spatial and non-spatial components to overall inequality is shown in Figure 4.

The relative contribution from the spatial component decreases up to 2010 and shows slight increases in 2011, 2013, and 2014. Conversely, the relative contribution from the non-spatial component shows an increasing trend, with slight decreases in 2011, 2013, and 2014.

Figure 4 – Contribution from non-neighbour (NNG) and neighbour (NG) components to overall inequality (G). GDP per worker data, NUTS 3 regions, 1995-2019. Source: Own Elaboration on ARDECO database.



However, both components only exhibit small variations. The variations in the spatial and non-spatial components of inequality are more evident when considering the decomposition proposed by Panzera and Postiglione (2020). The spatial Gini (G_s) and non-spatial Gini (G_{ns}) indices are reported in Table 2.

As reported in Table 2, the entire period under analysis is characterized by a dominance of the spatial component of inequality (G_s) relative to the non-spatial component (G_{ns}). This evidence is consistent with the results reported in Table 1.

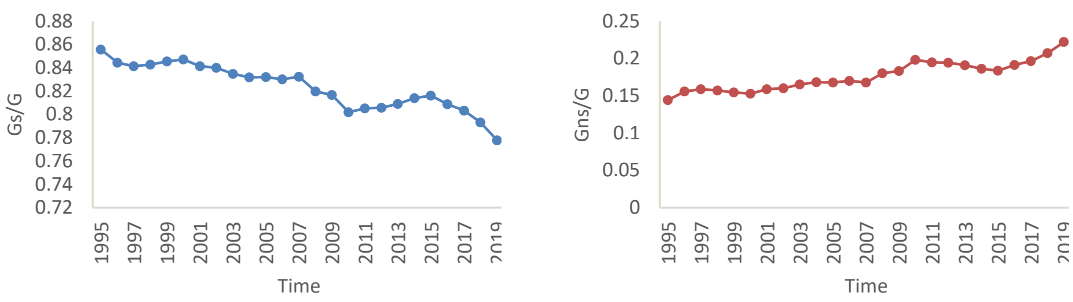
Applying the decomposition proposed by Panzera and Postiglione (2020) confirms the decreasing trend in the spatial component of inequality. This result is consistent with the trend seen in Moran's I, as shown in Figure 3.

Table 2 – Gini index of inequality (G) and its Spatial (G_s) and Non-Spatial (G_{ns}) components calculated for GDP per worker, NUTS 3 regions, 1995-2019. Source: Own Elaboration on ARDECO database.

Year	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**
G	0.1904	0.1877	0.1868	0.1869	0.1820	0.1795	0.1734	0.1682	0.1640
G_{ns}	0.0275	0.0292	0.0296	0.0294	0.0282	0.0276	0.0276	0.0270	0.0271
G_s	0.1629	0.1585	0.1572	0.1575	0.1538	0.1519	0.1458	0.1412	0.1369
Year	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
G	0.1593	0.1590	0.1572	0.1559	0.1500	0.1470	0.1465	0.1443	0.1411
G_{ns}	0.0269	0.0268	0.0268	0.0263	0.0272	0.0272	0.0290	0.0280	0.0274
G_s	0.1324	0.1322	0.1304	0.1296	0.1228	0.1198	0.1175	0.1163	0.1137
Year	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**		
G	0.1393	0.1383	0.1404	0.1398	0.1381	0.1360	0.1334		
G_{ns}	0.0266	0.0258	0.0259	0.0268	0.0272	0.0282	0.0296		
G_s	0.1127	0.1125	0.1145	0.1130	0.1109	0.1078	0.1038		

The relative contribution to overall inequality from the spatial and non-spatial components is shown in Figure 5. Figure 5 confirms the decreasing and increasing trends from spatial and non-spatial components of inequality, respectively. These results indicate the non-spatial component, which expresses the component that is not influenced by the effects of neighbours, increases during the period under consideration.

Figure 5 – Contribution from spatial (Gs) and non-spatial (Gns) components to overall inequality (G). GDP per worker data, NUTS 3 regions, 1995-2019. Source: Own Elaboration on ARDECO database.



Applying the spatial decomposition of inequality allows the spatial structure within a particular region to be accounted for. In fact, regional interactions could play a role in shaping the distribution of GDP values, and thus in determining similarities or disparities across regional units. The dominance of the spatial component of inequality that we empirically observed indicates that the geographical proximity among regional units has a great influence in determining the observed pattern of inequality in GDP values within EU regions. Specifically, the relevance assumed by the spatial component of inequality shows that differences in GDP values with respect to surrounding regions strongly influence the GDP within a particular region. These dynamics could be discovered only by isolating the spatial and non-spatial components of inequality. The role played by the geographical position of data on inequality suggests the relevance of defining place-based policies. The empirical results of our analysis suggest that policies aimed at reducing regional disparities should be defined for a specific region while accounting for potential influences from neighbouring regions. Furthermore, understanding the value of the influence due to proximity allows one to account for actual inequality that is caused by the fact that none regions is *per se*.

Our results also reveal that the decrease in regional inequality observed for the 1995-2019 period is primarily due to the decreasing of the component of inequality that is determined by the spatial interactions among regional units. Furthermore, geographical proximity has less influence on overall inequality during the period under investigation. Differences or similarities in GDP across EU regions are thus increasingly determined by specific factors that are internal to each region. Thus, our empirical results indicate that spatial spillovers are gradually less evident during the period under investigation. The simultaneous increase in the non-spatial inequality reveals an effective increase in inequality which is masked when considering only overall inequality. This suggests the importance of decomposing inequality into its different components in order to appreciate the contribution of regional interconnections to inequality.

6.2 The impact of inequality on economic growth in EU NUTS 2 regions

In this section, the impact of inequality on growth in EU regions is assessed by using the empirical model defined in equation (9) to describe data related to NUTS 2 regions in 26 EU Member States. Cyprus and Luxembourg are excluded from the analysis because of the coincidence between the NUTS 2 and the NUTS 3 level of classification, which does not allow the within-region component of inequality to be calculated. Furthermore, just as was the case at the NUTS 3 level, French overseas departments are excluded from the subsequent analysis. Model estimation is performed using STATA software.

Our model is a spatially augmented Solow growth model, as defined by Ertur and Koch (2007), where investment in physical capital is expressed as a function of an inequality measure. Expressing investment as a function of inequality allows one to assess the effect of inequality on regional growth. Adopting a spatial perspective could provide a deeper understanding of this relationship because regional interdependences driven by geographical proximity are accounted for in the analysis.

The proposed model differs from the model proposed by De Dominicis (2014) in some important aspects. While De Dominicis (2014) relies on the within-region component of the Theil index as a measure of inequality, we measure the contribution from each region to inequality by considering the within-region component of the Gini index. With respect to the Theil index, the Gini index is easier to interpret, and its values can be compared directly across time and space. In fact, the Theil index has zero as its lower bound, while its upper bound depends on population size, making direct comparisons between different groups difficult (Buitelaar et al. 2017).

Following Rey and Smith (2013), we decompose the within-region Gini index into its neighbour and non-neighbour components and focus on the former, which can be interpreted as the component of inequality that is not influenced by spatial dependence. This inequality decomposition allows the different components of regional inequality to be distinguished and, thus, to appreciate the real extent of the phenomenon. The spatial interactions among regional economies are included in the model by considering the spatial lag of the inequality variable.

Considering inequality within NUTS 2 regions, using NUTS 3 units might provide useful information on the extent of inequality. In fact, the observed inequality decrease, as measured with the global Gini index G, does not exclude an increase in the within-region component of inequality. Focusing on disparities within each region provides policy makers with valuable information. In fact, if inequality in the EU follows from differences within regions, rather than from disparities between regions, it would make sense to develop interventions targeted to the context-specific needs and assets of each region.

Some evidence on within-region inequality and its spatial and non-spatial components from 1995 to 2019 is shown in the following figures. Within-region inequality is calculated for the NUTS 2 regions. Figure 6 shows the contribution from within-region inequality to total inequality, as measured by the Gini index G. As Figure 6 shows, regional disparities throughout the period under investigation primarily arise due to inequality within the NUTS 2 regions. Decomposition of the Gini index by population subgroups (Mookherjee and Shorrocks, 1982; Shorrocks and Wan, 2005) shows that the residual contribution in the definition of inequality can be attributed to interregional disparities and the overlapping term. The within-region component accounts for about 85% of the total disparities at the beginning of the period (i.e. in 1995), reaching its minimum of approximately 83% at the end of the period. During the period under consideration, the proportion of inequality originated by the within-region component goes through periods of expansion and decline, reaching some peaks in 2001, 2004, 2008, and 2013. The increasing contribution from within-region inequality has as its counterpart a reduction in interregional inequality that is mostly driven by success of regions in Cohesion countries at converging with the rest of the EU (European Commission, 2001). Since 2013, the proportion of within-region inequality shows a markedly decreasing trend, thus the contribution from between-region components in determining disparities in EU increases.

Figure 6 – Contribution from the within-region component to overall inequality as measured by the Gini index. GDP per worker data at NUTS 3 level grouped according to NUTS 2 regions. Source: Own Elaboration on ARDECO database.

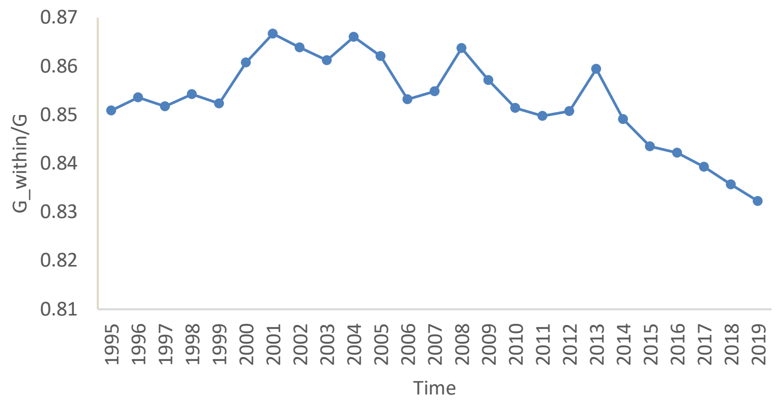


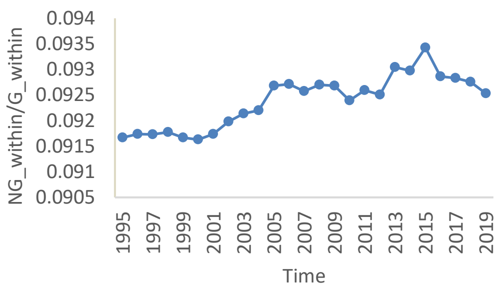
Table 3 – Within-region Gini index of inequality (G_{within}) and its neighbour (NG_{within}) and non-neighbour (NNG_{within}) components calculated for GDP per worker for NUTS 2 regions at the NUTS 3 level, 1995-2019. Source: Own Elaboration on ARDECO database.

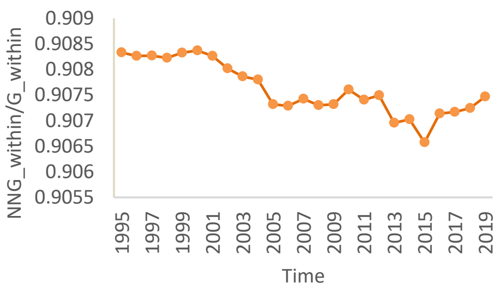
Year	**1995**	**1996**	**1997**	**1998**	**1999**	**2000**	**2001**	**2002**	**2003**
G_{within}	0.1619	0.1601	0.1590	0.1595	0.1550	0.1543	0.1502	0.1452	0.1411
NG_{within}	0.0148	0.0147	0.0146	0.0146	0.0142	0.0141	0.0138	0.0133	0.0130
NNG_{within}	0.1471	0.1454	0.1444	0.1449	0.1408	0.1402	0.1364	0.1319	0.1281
Year	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**	**2011**	**2012**
G_{within}	0.1378	0.1369	0.1339	0.1330	0.1294	0.1258	0.1246	0.1225	0.1198
NG_{within}	0.0127	0.0127	0.0124	0.0123	0.0120	0.0117	0.0115	0.0113	0.0111
NNG_{within}	0.1251	0.1242	0.1215	0.1207	0.1174	0.1141	0.1131	0.1112	0.1087
Year	**2013**	**2014**	**2015**	**2016**	**2017**	**2018**	**2019**		
G_{within}	0.1195	0.1172	0.1182	0.1175	0.1157	0.1134	0.1108		
NG_{within}	0.0111	0.0109	0.0110	0.0109	0.0107	0.0105	0.0112		
NNG_{within}	0.1084	0.1063	0.1072	0.1066	0.1050	0.1029	0.0996		

Table 3 shows the within-region inequality (G_{within}) values from 1995 to 2019 decomposed into its neighbour and non-neighbour components (NNG_{within} and NG_{within}, respectively). As previously mentioned, the non-neighbour component of the within-region component of the Gini index G expresses the component of the within-region inequality determined by positive spatial correlation. As revealed by the results in Table 3, this component is higher than the neighbour component throughout the period under investigation. This indicates that, from 1995 to 2019, the within-region component of inequality arises due to the positive spatial dependence, implying that similar dynamics drive internal inequality for neighbouring regional units.

The relative contributions from the neighbour and non-neighbour components of within-region inequality are depicted in Figure 7. Figure 7 shows that the contribution from the neighbour component to within-region inequality exhibits an increasing trend since 2000 and starts to decrease in 2016. This accounts for the inequality component that is not influenced by spatial dependence.

Figure 7 – Contribution from neighbour and non-neighbour components to within-region inequality. GDP per worker data at the NUTS 3 level grouped according to NUTS 2 regions. Source: Own Elaboration on ARDECO database.





Our empirical results reveal that the spatial structure where a region is located greatly influences its level of internal disparities. Such spatial contamination should operate via geographical proximity as well as through technological and economic linkages (Benos et al. 2015). These interregional interactions allow disparities within each region to be analysed in terms of the inequality in neighbouring regions, offering valuable information to policy makers. In fact, these influences should be accounted for when defining regional policies. Quantifying the spatial component of within-region inequality also allows a non-spatial component to be identified, which shows that inequality is not influenced by interregional interactions and is primarily determined by region specific factors. This latter component is introduced as an explanatory variable in the empirical model in equation (9).

Specifically, in model (9), the dependent variable, g_i, is the natural logarithm of the average growth rate in GDP per worker in PPS[9] during the 1995-2019 period. The natural logarithm of the initial (1995) GDP per worker, y_{it-T}, is included in the model to assess the process of conditional β-convergence. The inequality variable, $\ln(Ineq_i)$, is the natural logarithm of the neighbour component of the within-region Gini index, as defined in equation (12).

The proposed analysis focuses on 274 NUTS 2 regions. The within component of the Gini index is calculated using GDP per worker related to the 1341 NUTS 3 regions belonging to the aforementioned NUTS 2 units[10] and is decomposed into its neighbour and non-neighbour components, according to equation (11). The inequality measure included in the growth model is referred to the initial period (1995). The variable n_i is the natural logarithm of the average growth rate of the working population from 1995 to 2019. Note that $(l + k)$ = 0.05, which follows a common assumption in the literature (Mankiw et al. 1992).

The proximity relationships among regional units are summarized by their spatial weight matrix as defined according to the k-nearest neighbours, with $k = 7$. The same criterion is considered when defining the weight matrix that summarizes proximities among NUTS 3 regions and that is used to define the spatial component of G_{within}.

The choice of $k = 7$ is consistent with the circumstance that EU regions have on average 5-6 contiguous neighbours (Le Gallo and Ertur, 2003). However, in this report, several values of k have been tried in the definition of the proximity matrix. Different specifications of k did not give rise to significant differences in the results.

[9] The purchasing power standard, abbreviated as PPS, is an artificial currency unit. Theoretically, one PPS can buy the same amount of goods and services in each country.

[10] Among the EU NUTS 2 regions, there are some regions composed by a relatively large number of NUTS 3 regions, and other regions that are composed by a single NUTS 3 region. For the latter regions, the Gini index is zero. In order to include these regions in our analysis, we follow the same approach proposed by De Dominicis (2014). In particular, when a NUTS 2 region is composed of a single NUTS 3 region, the Gini index for this region is equal to the average value of the Gini index computed for the NUTS 2 regions in the same country.

Summary statistics of the dependent and independent variables are shown in Table 4. Table 4 shows that the coefficient of variation is higher for the GDP per worker growth rate, while the initial GDP per worker has lower dispersion around the mean value.

The estimates obtained for the model in equation (9) are shown in Table 5, where the spatial and non-spatial models are compared. The non-spatial model is estimated using ordinary least squares (OLS) regression. The SDM model is estimated using the maximum likelihood method and, because the Breusch-Pagan test reveals the presence of heteroscedasticity (i.e. non-constant error variance), we also apply a generalised spatial two-stage least squares (GS2SLS) calculation to treat error as heteroscedastic (Kelejian and Prucha, 2010). The estimates obtained with this approach are reported in Appendix 1. They have similar magnitude as the maximum likelihood estimates and give rise to similar interpretation.

The use of the spatial specification corresponding to SDM is theoretically motivated. In fact, as previously mentioned, we focus on the spatial augmented Solow growth model that has the SDM as its empirical counterpart (Ertur and Koch, 2007). However, the appropriateness of the SDM specification has been also empirically verified, following the procedure for model comparison suggested by Elhorst (2010). Details are given in Appendix 2.

Table 4 – Summary statistics of dependent and independent variables in the regional growth model (1995-2019), 274 NUTS 2 regions. Source: Own Elaboration on ARDECO database.

Variables	Min	Q_1	Median	Mean	Q_3	Max	CV
g	0.0069	0.0189	0.0226	0.0261	0.0289	0.0792	0.4513
$\ln y_{t-T}$	8.5510	10.2330	10.5360	10.3830	10.6520	11.1270	0.0430
$\ln Ineq_i$	-14.1950	-10.9730	-10.2890	-10.3510	-9.6130	-7.128	0.1048
$\ln(n + 0.05)$	-3.7220	-2.9440	-2.8690	-2.8780	-2.8000	-2.5440	0.0497

In Table 5, the parameter estimates are shown with some diagnostic and performance measures. Model representativeness is assessed using conventional statistical measures, such as the Akaike information criterion, AIC, and the coefficient of determination, R^2. Moran's I statistic on the regression residuals is considered as diagnostics for spatial dependence.

Table 5 – Estimation results for the spatial and non-spatial growth models (1995-2019) using cross-sectional data, 274 NUTS 2 EU regions. Source: Own Elaboration on ARDECO database.

	Non- Spatial model	SDM
Variable	Coefficient (standard error)	Coefficient (standard error)
Constant	0.3214*** (0.0216)	0.1445***(0.0404)
$\ln y_{t-T}$	-0.0248 ***(0.0012)	-0.0195***(0.0017)
$\ln Ineq$	0.0011***(0.0004)	0.0007**(0.0003)
$\ln(n + 0.05)$	0.0093***(0.0035)	0.0040 (0.0035)
$W \ln y_{t-T}$		0.0089***(0.0029)
$W \ln Ineq$		-0.0002 (0.0007)
$W \ln(n + 0.05)$		0.0024 (0.0059)
ρ		0.6463***(0.0692)
λ (Convergence Rate)	3.87%	2.67%
Moran's I	0.3174***	
AIC	-1,990.10	-2,049.78
R^2	0.7068	

Significance levels ***1%, ** 5%, * 10%

As shown in Table 5, in the non-spatial model, all coefficients are highly significant. The β coefficient associated with the initial GDP per worker is negative, revealing β-

convergence among the considered spatial units. Regarding the other variables included in the non-spatial model, we find positive and significant impacts on inequality for the GDP per worker growth rate, and the coefficient associated with the working population growth rate is positive and significant. The convergence rate for the non-spatial model is $\lambda = 3.87\%$. Moran's I statistic for spatial autocorrelation applied to the OLS residuals is positive and highly significant. This reveals the presence of spatial dependence that must be considered in the model definition.

Moving from the non-spatial to the spatial specification, we note a decreased AIC value for the spatial model. In the SDM specification, the coefficient associated with the initial GDP per worker is positive and highly significant. The coefficient associated with the inequality variable is positive and significant. This result supports the predicted sign of the inequality variable. The coefficient associated with the working population growth rate is positive but not significant. The coefficients associated with the spatially lagged inequality variable and the spatially lagged working population growth rate are both not significant. The coefficients associated with the spatial lag in the initial GDP per worker and spatial autocorrelation parameter ρ are both positive and highly significant. The speed of convergence for the spatial model, $\lambda = 2.67\%$, is lower than that obtained for the non-spatial model. This result is consistent with findings in previous studies that showed how the introduction of spatial effects tends to reduce the estimated speed of global convergence (Monfort, 2008).

The model estimation appears to support the idea that the steady state position of a particular region is related to inequality in neighbouring regions, since $\frac{\rho_2}{\beta_1} = \frac{-0.0002}{-0.0195} = 0.010$ (see Section 4.1).

As pointed out by LeSage and Pace (2009), in order to interpret the SDM coefficients correctly, the direct, indirect, and total impacts need to be calculated (LeSage and Fischer, 2008). In fact, because of the spatial dependence incorporated in the model, a change in an explanatory variable for a single spatial unit has a direct impact on the dependent variable at the same location and could indirectly affect the dependent variable at different locations. The average direct impact provides a summary measure of the impact arising from changes in the i^{th} observation of a covariate on the dependent variable from region i; it also includes feedback effects. The average indirect impact reflects spatial spillovers. The sum of the average direct and indirect impacts measures the average total impact. Table 6 reports the average direct, indirect, and total impacts that would arise from changing each explanatory variable in the SDM. A set of 2,000 MCMC draws are used to produce impact estimates and inferences of their statistical significance.

As Table 6 shows, the average total impact associated with $\ln y_{t-T}$ is negative and significant. This estimate reveals that a 1% increase in the initial GDP per worker is associated with a 0.0297% decrease in the average growth rate. This result is consistent with the conditional β-convergence hypothesis. The total impact is obtained by summing the direct and indirect effects. The average direct impact associated with $\ln y_{t-\tau}$ is negative and highly significant. This indicates that an increase in the initial GDP per worker in a specific region determines a decreasing in the region's subsequent growth rate. The indirect impact estimate reveals that the initial GDP per worker in neighbouring regions also determines a decrease in the growth rate of GDP per worker. The total impact associated with the inequality variable is positive but not significant. The total impact is derived from the sum of positive direct and indirect effects, but only the direct effect is significant. A positive but not significant total impact is reported for the working population growth rate.

Table 6 – Average direct, indirect, and total impacts of the explanatory variables for SDM, 274 NUTS 2 European regions, 1995-2019. Source: Own Elaboration on ARDECO database.

Variable	Average Direct Impact	Average Indirect Impact	Average Total Impact
$\ln y_{t-T}$	-0.0200***	-0.0102**	-0.0302***
$\ln Ineq$	0.0008**	0.0008	0.0016
$\ln(n + 0.05)$	0.0047	0.0136	0.0182
Significance levels ***1%, **5%, *10%			

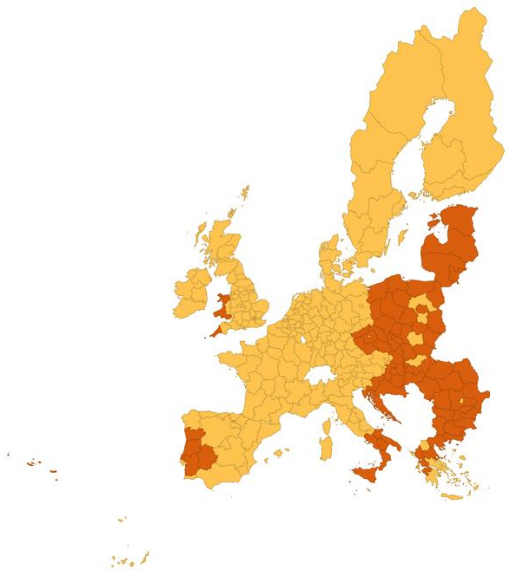
These results reveal that, for the EU NUTS 2 regions, inequality within regional units increases the growth rate in GDP per worker. This growth-promoting effect of inequality can be explained by the incentive to increase work productivity and thus to increase investment in education to reach higher income level, as well as by the higher saving propensity of high-income persons that leads to higher investments. However, this positive relationship is not significant when considering the inequality within neighbouring regions. In other terms, our empirical findings indicate that higher internal disparities in surrounding regions do not significantly impact GDP growth in a particular region.

As a further aspect, there may not be a single, universal mechanism governing the relationship between inequality and regional growth. In this regard, some recent empirical studies highlighted that, although inequality increases overall growth, such a positive effect appears to apply primarily to richer economies (Besley and Burgess, 2003;

Ravallion, 2007; United Nations, 2020). To verify this relationship, we estimate the SDM for two different clusters of regions corresponding to less-developed regions, transition regions, and more-developed regions. The two clusters of regions are depicted in Figure 8. The darker colour shows less-developed regions, while the light colour shows transition regions and more-developed regions.

The less-developed regions correspond to the NUTS 2 regions with GDP per capita less than 75% of the EU average. The transition regions have GDP per capita between 75% and 90% of the EU average, and more-developed regions have GDP per capita higher than 90% of the EU average. All these regions are eligible for funding under different funds and programmes under the EU regional policy. However, the less-developed regions receive most of the support, since the European funds are primarily targeted at reducing regional disparities in income, wealth, and opportunities.

Figure 8 – Clusters of less-developed regions, transition regions, and more developed regions (EU Cohesion Policy, 2014-2020).



The maximum likelihood estimation results in SDM for both the described clusters are reported in Table 7[11]. The results reported in Table 7 reveal that the effect of intraregional inequality on increased regional growth is significant only for the transition and more-developed regions. This implies that higher internal disparities could increase

[11] For the sake of simplicity, we do not report the table of impacts for this heterogenous version of the model.

growth in wealthy regions. On the other side, this relationship is not significant for less-developed regions.

The convergence process throughout the considered period is confirmed for both groups of regions, but the convergence rate reported for the less-developed regions is higher than that reported for the other regions. This result indicates that economic growth proceeds in less-developed regions faster than in other regions. This more rapid growth could be partly explained by their drivers primarily related to EU funding. The appearance of these different convergence clusters seems to support the relevance of identifying differentiated political responses.

As our empirical results suggest, regions characterized by different levels of development experience differences in growth, and intraregional inequality plays a different role in this process. The positive correlation between inequality and growth is observed for more-developed regions. Hence, this empirical evidence reveals that internal disparities could have a growth-promoting effect at higher levels of regional development. This relationship is not confirmed when considering poorer areas that do not experience a significant positive impact of internal disparities on regional growth. This result sheds light on the differences between the groups of regions under consideration and supports the relevance of policies targeted at less-developed areas to promote regional growth and to favour spatial justice, which are intended as a more equitable allocation of resources and opportunities to use them (Soja, 2009).

Table 7 – Parameter estimates for cluster of EU regions, transition and more developed regions and less developed regions (1995-2019). Source: Own Elaboration on ARDECO database.

	SDM Transition and More Developed Regions	SDM Less-Developed Regions
Variable	Coefficient (standard error)	Coefficient (standard error)
Constant	0.1467*** (0.0474)	0.2651***(0.0550)
$\ln y_{t-T}$	-0.0203*** (0.0024)	-0.0249***(0.0028)
$\ln Ineq$	0.0008** (0.0003)	0.0005 (0.0007)
$\ln(n + 0.05)$	0.0136*** (0.0048)	-0.0025 (0.0047)
$W \ln y_{t-T}$	0.0084** (0.0035)	0.0071* (0.0038)
$W \ln Ineq$	-0.0004 (0.0008)	0.0008 (0.0018)
$W \ln(n + 0.05)$	-0.0110 (0.0075)	0.0232***(0.0078)
ρ	0.5888*** (0.0670)	
λ (Convergence Rate)	2.83%	3.89%
AIC	-2,066.7	

Significance levels ***1%, **5%, *10%

To further investigate the hypothesis related to the presence of significant growth promoting effects of inequality in more-developed regions, we present results obtained with the model described in Section 4.1 for a group of regions that does not include regions belonging to the new accession Member States, such as countries that joined the EU in 2004 (Estonia, Latvia, Lithuania, Poland, Czech Republic, Slovakia, Slovenia, and Hungary), in 2007 (Bulgaria and Romania), and 2013 (Croatia). Our analysis focuses on the 206 NUTS 2 regions belonging to 14 EU Countries (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Ireland, Netherlands, Portugal, Spain, Sweden, and United Kingdom). Note that Luxembourg is excluded from the analysis because of the coincidence between the NUTS 2 and the NUTS 3 levels, which does not allow the inequality variable to be calculated. In this analysis, we consider a wider time period from 1980 to 2019 because data on the variable of interest are available for the entire period for all units under investigation. The ARDECO database was used as the data source. The growth rate in GDP per worker and working population growth rate were calculated from 1980 to 2019. The initial GDP per worker and inequality variable are calculated for the year 1980. The inequality variable is calculated as described in Section 4.2 by focusing on GDP per worker data related to the 1023 NUTS 3 regions belonging to the aforementioned NUTS 2 geographical units.

Results from estimation of the non-spatial and spatial growth models are reported in Table 8, along with some performance and diagnostic measures. As Table 8 shows, all the coefficients in the non-spatial model are highly significant and have the expected sign, with the only exception being the coefficient for the working population growth rate. Moran's I for the regression residuals is highly significant, revealing the presence of a positive spatial correlation. Spatial dependence is included in the model using SDM, which performs better than the non-spatial model, as revealed by the lower AIC value. In this model, most of the coefficients and the coefficients associated with initial GDP per worker, its spatial lag, and inequality variable are significant. Both models confirm a convergence process is taking place, suggesting initial intraregional inequality increases regional growth. The spatial model predicts a convergence rate of 4.60%. The convergence rate found for the SDM estimated on the 206 NUTS 2 regions from 1980 to 2019 is higher than that reported for the 274 NUTS 2 regions considered from 1995 to 2019. This result indicates that the gap in GDP per worker between regions that are more similar (i.e. the regions considered from 1980 to 2019) should be halved with respect to the convergence process among regional units that are more heterogeneous.

The proper interpretation of the SDM requires computing the direct, indirect, and total effects, which are reported in Table 9.

Table 8 – Estimation results for spatial and non-spatial growth model (1980-2019) using cross-sectional data, 206 NUTS 2 EU regions. Source: Own Elaboration on ARDECO database.

	Non-Spatial model	SDM
Variable	Coefficient (standard error)	Coefficient (standard error)
Constant	0.2634***(0.0184)	0.0845***(0.0295)
$\ln y_{t-T}$	-0.0198***(0.0012)	-0.0210***(0.0010)
$\ln Ineq$	0.0008***(0.0003)	0.0006**(0.0002)
$\ln(n+0.05)$	0.0096***(0.0034)	0.0076**(0.0030)
$W \ln y_{t-T}$		0.0160***(0.0019)
$W \ln Ineq$		0.0001 (0.0006)
$W \ln(n+0.05)$		-0.0008 (0.0056)
ρ		0.7036***(0.0613)
λ (Convergence Rate)	3.94%	4.60%
Moran's I	0.4134***	
AIC	-1636	-1713
R^2	0.5805	

Significance levels ***1%, **5%, *10%

Table 9 – Average direct, indirect, and total impacts of the explanatory variables for SDM, 206 NUTS 2 European regions, 1980-2019. Source: Own Elaboration on ARDECO database.

Variable	Average Direct Impact	Average Indirect Impact	Average Total Impact
$\ln y_{t-T}$	-0.0208***	0.0039	-0.0169***
$\ln Ineq$	0.0007**	0.0014	0.0020
$\ln(n+0.05)$	0.0083***	0.0146	0.0230
Significance levels ***1%, **5%, *10%			

The results in Table 9 show the presence of significant direct effects for all variables under consideration, and a significant total effect for initial GDP per worker. These findings confirm the presence of a convergence process and a positive correlation between intraregional inequality and growth for the group of regions under consideration.

The presence of no significant indirect impacts indicates the selected explanatory variables observed for a region only impact growth in that region. However, some feedback influences from other regions are accounted for in the estimated average direct impacts. A positive significant spatial autocorrelation coefficient was obtained in all analyses and supports the conclusion that regional growth rates has a significant positive spatial dependence.

The obtained empirical results appear to support some previous studies, where a positive correlation between inequality and growth was observed in wealthy regions (Artelaris and Petrakos, 2016). According to our empirical findings, more developed and transition regions could benefit from higher internal disparities, while they do not take advantage of high disparities in neighbour regions. This positive relationship is insignificant in less-developed regions. In more-developed regions, growth-promoting aspects of inequality could dominate because of less-serious credit market constraints, which facilitates investments and hence growth (Barro, 2000; Voitchovsky, 2005). Our empirical evidence is in part supported by some previous studies. As an example, focusing on a sample of countries at different level of development, including the United States and other major OECD countries, as well as countries from Sub-Saharan Africa and Latin America, Barro (2000) found that inequality encourages economic growth in richer countries while retards growth in poorer places. Similarly, Voitchovsky (2005), considering a large number of countries including, among others, countries from Europe, the United States, and Mexico, verified that inequality at the top and bottom ends of income distribution has different implications on economic growth. Specifically, inequality at the top-end of income distribution appears to encourage growth, while disparities at the bottom-end of the distribution retards growth. Conversely, empirical studies focused on EU regions, obtained different evidences.

Our empirical findings differ from the evidence obtained by De Dominicis (2014), who found inequality has a positive impact on growth in less-developed regions. The difference in the estimated impact of inequality on regional growth could be explained by differences in the considered variables. Our analysis provides new evidence by isolating the idiosyncratic component of inequality from the part of inequality driven by spatial proximity relationships. In other words, isolating the intraregional disparity components, which are determined by regional interdependencies driven by geographical proximity, allows a specific intraregional disparity component to be identified and linked to region-specific factors. Assessing the impact of this component on regional growth leads to discovery a growth-promoting effect of inequality. This positive effect is significant only for more developed and transition regions. Hence, differences in internal factors produce different levels of regional development, yet differences in the level of development across regions could determine differences in the factors driving regional growth. These differences support the differentiation of policies targeted to regions in different categories. In light of our empirical evidence, policies targeted to poorer areas should focus on increasing their level of development and ensure spatial justice. A more equitable distribution of resources within regions could be relevant for these areas, where intraregional disparities do not produce significant growth-promoting effects.

6.3 Going beyond GDP at low geographical scale: A measure of local economic development at the municipality level

Single indicators are often used to simplify and interpret economic and social phenomena. Unfortunately, the reality is more complicated and is composed of interconnected, multidimensional aspects (Stiglitz et al. 2009). To solve this problem, many researchers have tried to describe complex phenomena by combining sets of different variables into composite indicators. Composite indicators are an alternative way to represent economic, social, and environmental problems (OECD, 2008). A large number of studies have focused on developing composite indicators to increase stakeholder awareness of serious problems (Kuc-Czarnecka et al. 2020). Moreover, composite indicators are not limited to a comparison of different dimensions, but they enable researchers and analysts to communicate to the general public and, consequently, increase accountability (Nardo and Saisana, 2008).

There are three aspects that may lead us to explore economic performance beyond GDP. As noted by Carrozza (2014), composite indicators could be considered by experts in different fields. In addition, Boulanger (2014) noted that composite indicators can be instrumental in the development of a new public. A larger effort towards defining composite indicators allow us to specify measures that target and support specific policies (Kuc-Czarnecka et al. 2020). Composite indicators are not just numbers pointing to an issue; they also entail a specific policy aim or social change desired by proponents of the measure. Hence, the ability of a composite indicator to provide greater insight into complex phenomena, such as economic performance, could support local bottom-up strategies.

Composite measures may be built on different definitions that must be strictly connected to the aim of the policy under consideration. In WP1, it was noted that territorial cohesion has not resulted in commonly agreed indicators, while there have been various attempts to operationalize this concept in a multivariate sense. In this direction, composite indicators offer the opportunity to come closer to a wider concept of territorial cohesion. Indeed, composite measure may contain a certain degree of subjectivity. However, their approach may be of great help in the search, identification, and strengthening of territorial instruments at the local, regional, national, and European level in the post-2020 period. This feature has been considered as very relevant in the European development strategy[12].

In this section, we synthesise several indicators that are available at the municipality level for these three countries, which are among the most populated in Europe. We take advantage of the dataset provided by WP2 to complement the results previously presented in the report. In the specific case, we perform the analysis merely at a very localized level (i.e. municipality level) to provide a spatially deeper picture and a multivariate focus on local disparities.

The analysis reported in this section must be considered as explorative within the sphere of local economic development as it is based on data from the IMAJINE dataset. However, as already indicated in WP1 (see D 1.1), multidimensional monitoring of economic well-being may bring the analysis closer to the concept of economic cohesion. On a more practical level, a multidimensional composite indicator that goes beyond the

12 Radim Srsen, Regional Councillor of Olomouc (Czechia), stated on 8 October 2019 during the 136th European Committee of Regions that ‘Community-led local development has proved to be a very successful tool of local development delivering European values by local means through a strong engagement of citizens and a bottom-up approach’.

use of single indicators at low geographical scale offers a different perspective for policy makers compared to the insight gained from examining per capita GDP, and it may be useful for evaluating local policies and integrated territorial investments[13].

Five different dimensions are considered based on data in the IMAJINE dataset. First, economic wealth is considered as evidence of performance at a regional level. Hence, we include disposable income of households at the municipality level (*wealth*). The second dimension considered is poverty, for which we consider the at-risk-of-poverty rate (AROPE) before social transfers, calculated as the share of people having an equivalent disposable income before social transfers that is below the at-risk-of-poverty threshold (*aro*). This indicator is often taken as a relevant feature at regional level by the EU[14]. Unfortunately, data for AROPE at the municipality level in Italy are unavailable, therefore this measure will not be included in the definition of a composite indicator for Italy.

A dimension that also appears relevant in terms of local economic performance is education (European Commission, 2001). In order to consider lack of educational attainment at a local level, we consider the rate of people (illiterate or not) who did not receive formal education (*edu*). Local labour (*lab*) is also considered in the unemployment rate and is included in the composite indicator as a key input (Potter and Marchese, 2010). Certain studies have also emphasized share of agriculture as an important indicator of the sectorial mix at the local level (*agr*). For example, Paci and Pigliaru (1999) explored the remarkable differences in the sectorial mix for Italian regions and found that agricultural prevalence characterizes lagging regions, which struggle to have a deep structural change in their economies and face stagnating growth (Schuh et al. 2019). Hence, the share of the agricultural sector, calculated as the ratio of agricultural employees to total employees within a municipality, is considered in the development of a composite measure.

PCA is used to define weights and synthesise a composite indicator (Greyling and Tregenna, 2017). Moreover, particular attention is paid to the spatial characteristics of the given data, and a spatial extension of PCA, namely GWPCA, is used. GWPCA allows weights at the unit level to be defined, which can be used to tackle substantial

13 An overview of different attempts to measure territorial and economic cohesion is provided in D 1.1 of the IMAJINE Project.

14 The urban and regional dimension of the crisis: Eighth progress report on economic, social and territorial cohesion, Report from the Commission, June 2013.

heterogeneity. This feature may be appealing for understanding the importance of each dimension in space and help design regional policies that consider heterogeneity within regions.

As most of the variables have theoretically negative relationships with respect to local economic performance, household income is reversed in sign. In this fashion, a lower value of the composite measure indicates a higher potential for local economic performance. Conversely, units characterized by a larger composite indicator must be interpreted as being located in more-disadvantaged areas. As single inputs are expressed in different unit measures, all variables are standardized to have zero mean and unit variance (Becker et al. 2017). The year for the analyses is 2011 due to the data available in the WP 2 IMAJINE dataset.

6.3.1 Local performance at the municipality level in France

The first country to be analysed is France. As a first step, the weights for each variable are reported in Table 10. Weights are obtained as loadings using standard PCA at the municipality level. For the sake of simplicity, only loadings from the first component indicator are included.

Table 10 – Loadings from the first component of PCA for French municipalities. Source: Own Elaboration on IMAJINE WP 2 database.

Variables	Weights (PCA)
Wealth (-)[15]	0.496
AROPE	0.628
Unemployment	0.528
Agricultural sector	0.219
No education	0.178
Proportion of variance explained	0.351

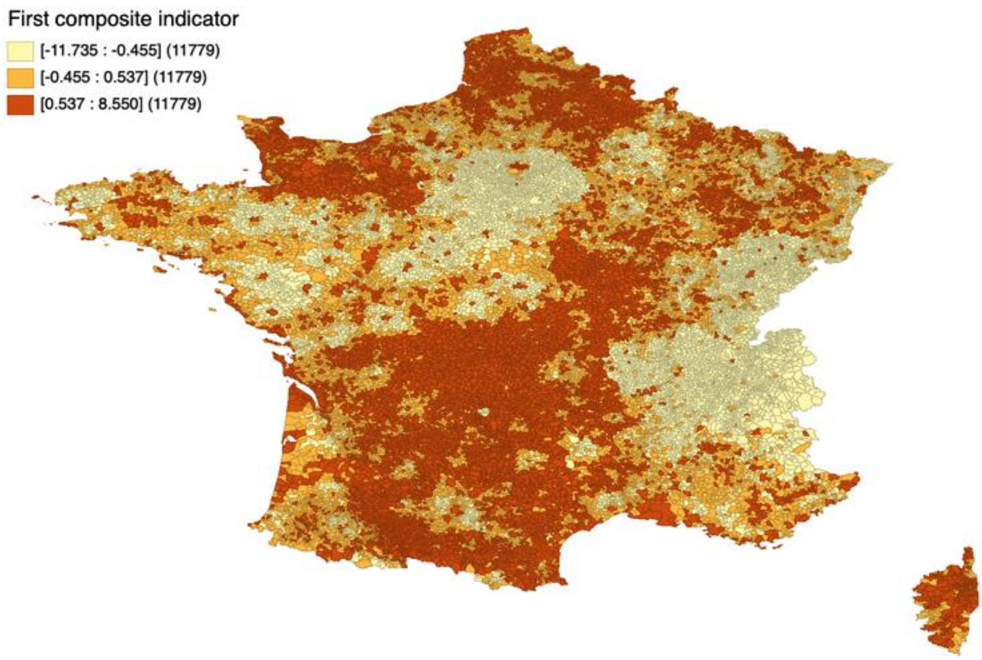
As expected, all loadings have the same sign and all variables are positively related to the final measure. Accordingly, higher levels of the composite indicator suggest local disadvantages related to lower income, employment, and educational attainment, as well as higher rate of poverty and relative size of the agricultural sector. As reported in Table 10, the most important variable in the definition of the composite indicator is

[15] The variable ‘wealth’ is included with the sign minus in brackets (-) to point out that the input is reversed in its sign.

AROPE, as it presents the highest loading. Unemployment is also very relevant. Surprisingly enough, wealth (-) is only the third variable in terms of importance. Share of agricultural sector and lack of educational attainment tend to be less relevant in the structure of the data for France. The first composite indicator encompasses approximately 35% of the information included in the data.

In PCA, the principal component scores (see equation (14)), obtained as the product of inputs for their respective weights, represent the values of composite indicators. Therefore, a way to explore the results obtained could be to map the scores for the first component. Figure 9 shows a quantile map of our composite measure obtained using the weights in Table 10.

Figure 9 – Quantile map of the first composite indicator from PCA for municipalities in France. Source: Own Elaboration on IMAJINE WP 2 database.



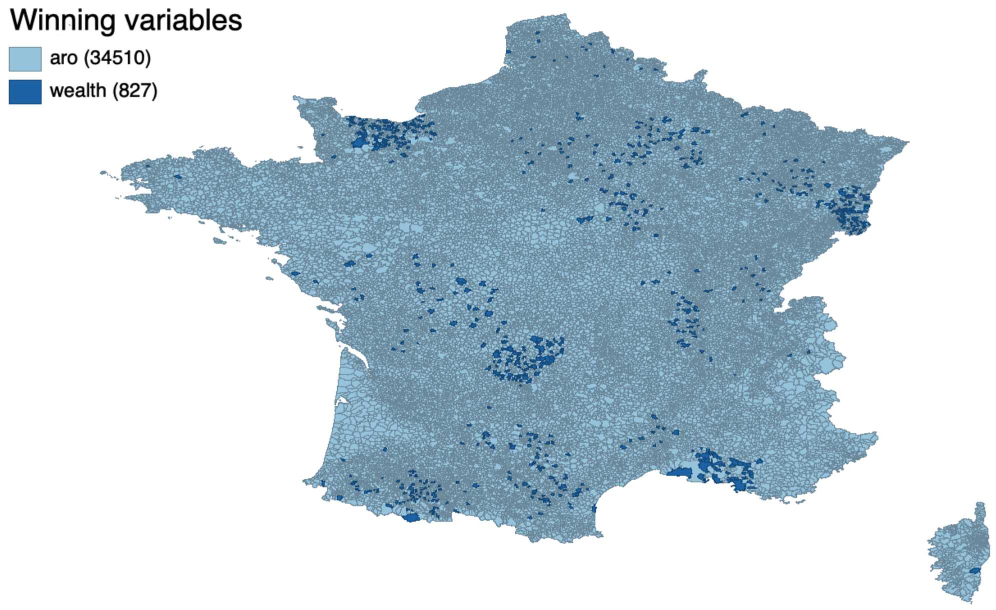
As shown in Figure 9, lower values of the composite indicators are situated closer to the most urbanized areas, corresponding to the two major cities in France (Lyon and Paris - Ile de France). Moderately low levels of the composite indicators are also located near the city of Toulouse, Loire valley, and in the southeast. Conversely, more rural areas in the *Massif Central* and in the North Departments exhibit higher values. This representation provides high spatial resolution, which can aid identification of very local disparities, with particular attention paid to urban-rural dualism. Despite that, the values reported in Figure 9 may suffer from misspecification and lead to misinterpretation as the weights for all municipalities are held constant. It may be a too general assumption to consider each input as equally relevant throughout France. For this reason, it might be desirable to allow the weights to vary.

To this end, GWPCA was used to analyse the data from France by adopting a bi-square kernel and an adaptive[16] bandwidth, which is obtained using a cross-validation procedure (Harris et al. 2015). Results from GWPCA are extensive and may be difficult to summarize. One well-known option is to use winning variables (Harris et al. 2011), defined as variables with the highest local loading in absolute value in each location. In terms of composite indicators, the winning variables suggest which variable has the greatest influence over the final indicator in each unit. Figure 10 shows a map of the winning variables for the first composite indicator defined with GWPCA at the municipality level.

Figure 10 reveals some differences at very local levels. Two inputs are identified as winning variables at the municipality level, namely AROPE and wealth. In particular, if AROPE is the most important variable for identifying critical situations in local economic performance throughout nearly the entire country, wealth tends to be more important in many locations, including municipalities in the *Marseille* region, near the German border, and in several units in the Channel area. Figure 11 shows a map of the local indicator calculated with GWPCA for first component.

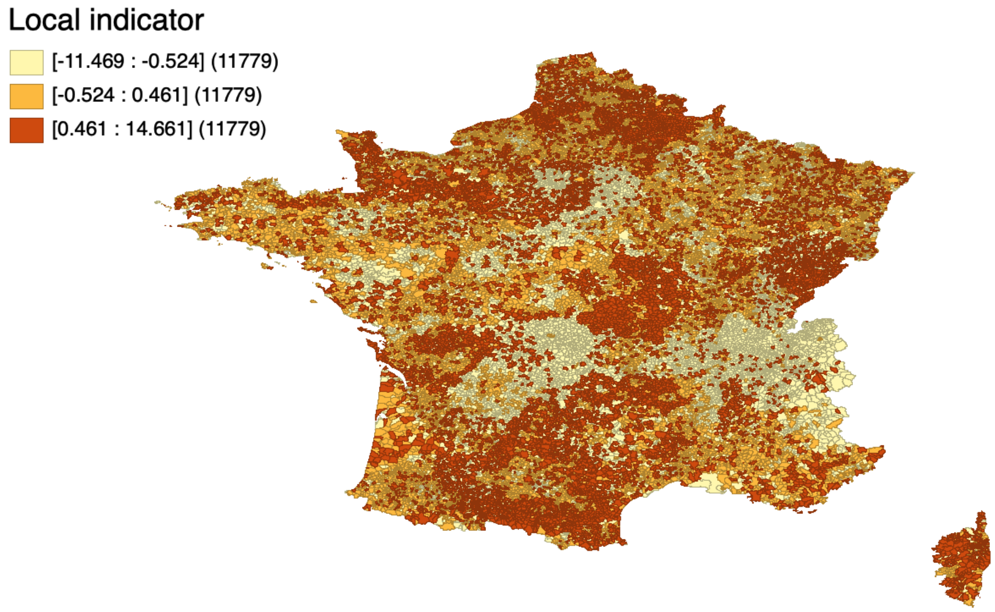
[16] Adaptive kernels consider a certain number of neighbours for each municipality. Contrast this with a fixed bandwidth, where neighbours are only accounted for in a fixed range. Adaptive bandwidth is preferable for regions irregular borders and patchwork regions (Fotheringham et al. 2002).

Figure 10 – Map of the winning variables for France at the municipality level. Source: Own Elaboration on IMAJINE WP 2 database.



In Figure 11, the local indicator is defined using loadings of the first local principal component at the municipality level. Many differences can be observed in this map compared to Figure 9. The geography of the indicator differs from that obtained using standard PCA. In fact, a more detailed picture of the local differences within regions emerges. For example, less economically developed areas (in dark brown) are more clustered around mountainous and rural areas (particularly in the *Massif Central*). In addition, many of the refined urban-rural differences can be seen from this indicator, as for *Languedoc-Roussillon* administrative region (southwest), which includes urban areas specialized in education and services (e.g. Montpellier) and rural areas.

Figure 11 – Quantile map of the local indicators for the first component determined from GWPCA for French municipalities. Source: Own Elaboration on IMAJINE WP 2 database.



6.3.2 Local performance at the municipality level in Spain

The extent of local multivariate economic performance is also analysed for Spain. Table 11 reports the weights obtained using standard PCA in Spain. Loadings for the first composite indicator are reported.

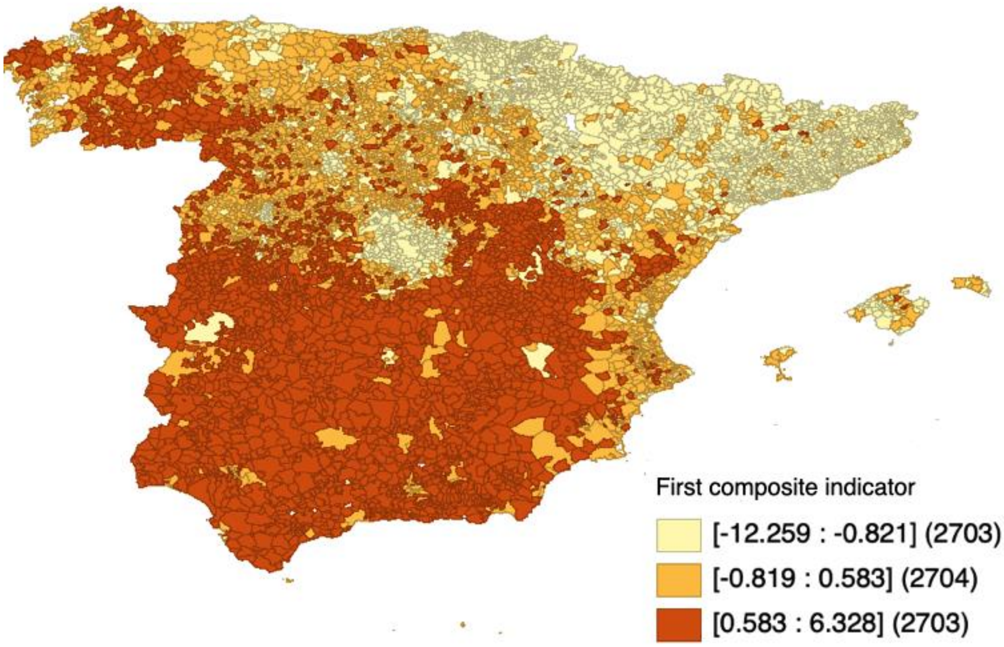
Values for weights obtained with PCA in Spain differ from those obtained in France. In this case, unemployment and lack of educational attainment are the most influential variables, followed by AROPE and household income. Compared to the case of France, the first component indicator is justified, in a statistical sense, by greater variance (58%), meaning a great percentage of total information is accounted for. Generally, all loadings have the expected sign. A map of the scores for the first composite indicator are shown in Figure 12.

Table 11 – Loadings from the first component of PCA for Spanish municipalities. Source: Own Elaboration on ARDECO database.

Variables	Weights (PCA)
Wealth (-)	0.460
AROPE	0.461
Unemployment	0.517
Agricultural sector	0.213
No education	0.512
Proportion of variance explained	0.580

Figure 12 shows how lower values of the composite indicator are situated closer to the most urbanized areas (Madrid and Barcelona). Moreover, a relevant north-south polarization emerges, as northern municipalities tend to be more economically developed than those in the south. A mixed pattern of moderate and low levels of composite indicators are also observed in the northern regions, especially in Asturias, while southern areas are more disadvantaged. In particular, municipalities in Andalusia and Extremadura exhibit higher indicators with a few exceptions for major cities and regional capitals. Moderate economic performance, as calculated with our explorative indicator, is also observed along the Mediterranean coast and localities nearby Valencia.

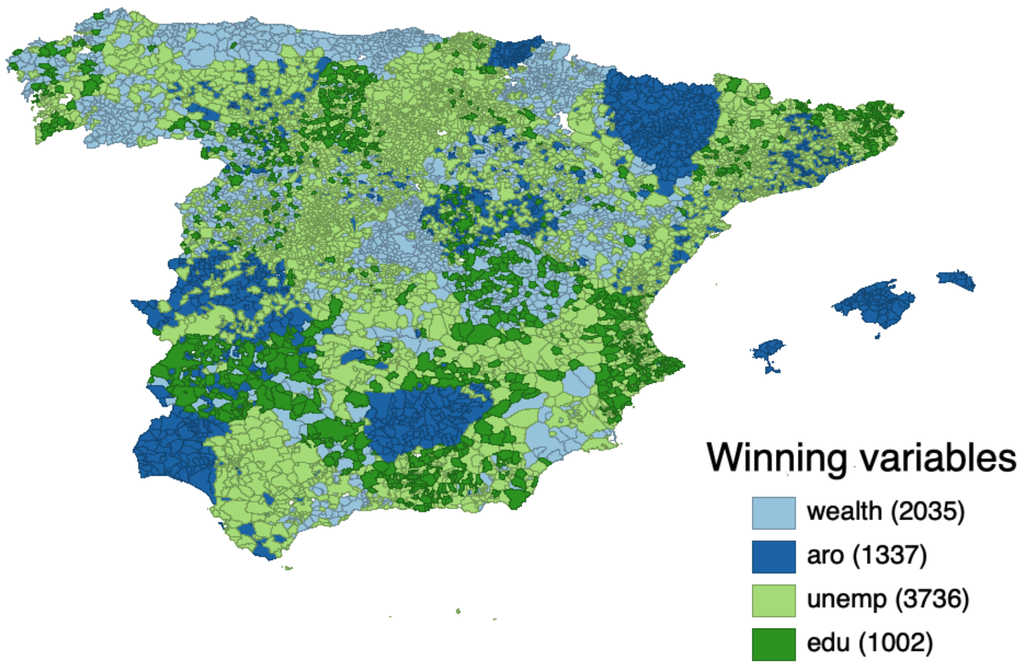
Figure 12 – Quantile map of the first composite indicator from PCA for Spanish municipalities. Source: Own Elaboration on IMAJINE WP 2 database.



As for France, we proceed by relaxing the assumption of homogeneity on the weights for the composite indicator by performing GWPCA. The winning variables for the first component determined with GWPCA are reported in Figure 13. In this case, a bi-square kernel with an adaptive bandwidth was used.

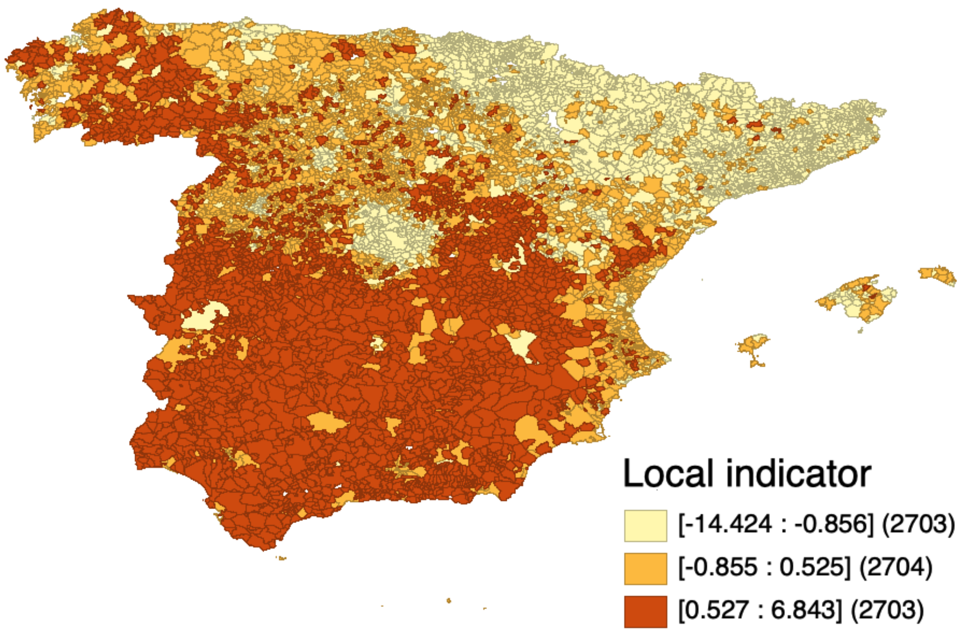
Figure 13 stresses the presence of local differences in Spain. The winning variable is different in various municipalities, even within the same region. Unemployment characterizes most Spanish municipalities. Household income (wealth (-)) is found to be the leading variable, especially in the Atlantic area. AROPE characterizes municipalities in very clustered areas of Aragon and Andalusia. Lack of educational attainment appears very scattered throughout the country. However, dark green clusters associated with this variable are primarily localised in southern Spain. Figure 14 shows a map of the local indicator calculated with GWPCA for the first component.

Figure 13 – Map of the winning variables for Spain at the municipality level. Source: Own Elaboration on IMAJINE WP 2 database.



In Figure 14, the local indicator shows some differences compared to the results from standard PCA. Even if those differences are less evident compared to France, the local indicator changes rank for approximately 100 municipalities. By comparing this map to Figure 12, more municipalities in the southern area shift to the median interval, which increases their rank. In addition, a series of changes appears in the *Castilla-y-Leon* region, where the number of dark brown units increases. Globally, the analysis performed with GWPCA confirms north-south polarization with more exceptions compared to the results found using standard PCA. Hence, a more detailed picture of local differences within regions emerges.

Figure 14 – Quantile map of the local indicators for the first component determined from GWPCA for Spanish municipalities. Source: Own Elaboration on IMAJINE WP 2 database.



6.3.3 Local performance at the municipality level in Italy

An analysis of local economic performance was conducted with refined spatial resolution also in Italy. Table 12 shows weights determined using standard PCA for the first composite indicator in Italy. As mentioned earlier, only four inputs were used, as AROPE data for Italy is unavailable in the IMAJINE dataset. Other variables have the same definitions as those in the previous analyses for France and Spain.

Table 12 – Loadings from the first component of PCA for Italian municipalities. Source: Own Elaboration on ARDECO database.

Variables	**Weights (PCA)**
Wealth (-)	0.528
Unemployment	0.509
Agricultural sector	0.354
No education	0.529
Proportion of variance explained	0.587

Here, the educational attainment appears to be the most important variable. However, both wealth and unemployment significantly affect economic performance. Share of the agricultural sector is less relevant when compared to other variables, but it has larger weight compared to France and Spain. The proportion of variance accounted by the first composite indicator is similar to that found in Spain. Figure 15 shows the first component scores calculated in Italy.

Figure 15 – Quantile map of the first composite indicator from PCA for Italian municipalities. Source: Own Elaboration on IMAJINE WP 2 database.

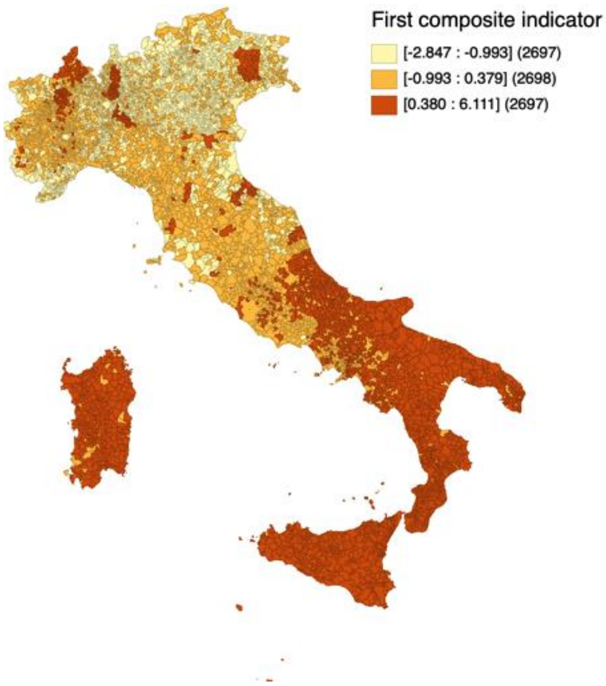


Figure 15 returns a highly polarized pattern. While municipalities in the northern part of Italy report very low values of the indicator, southern municipalities appear more disadvantaged. In fact, the country seems divided into three major clusters. Due to the large presence of heterogeneity that affects Italy (Panzera and Postiglione, 2014), considering local changes in weights may also be an appropriate choice for analysing Italian data. Hence, GWPCA was used with a bi-square kernel and adaptive bandwidth.

Figure 16 – Map of the winning variables for Italy at the municipality level. Source: Own Elaboration on IMAJINE WP 2 database.

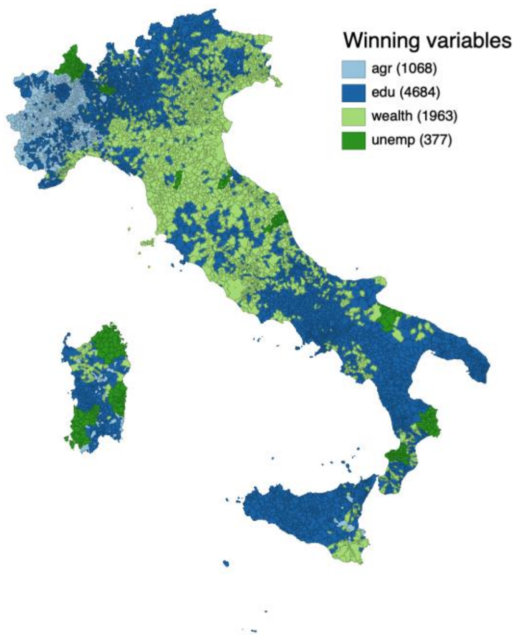
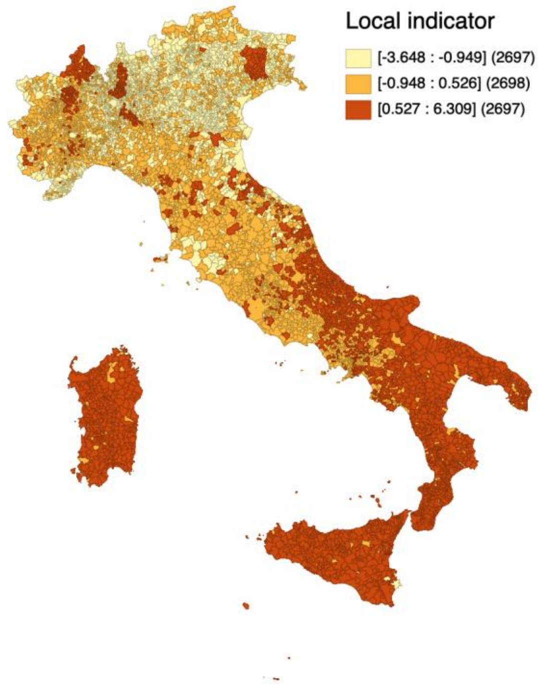


Figure 16 shows the winning variables for the first component in Italy determined using GWPCA.

Interestingly, lack of educational attainment is the leading variable in a large number of southern municipalities. Household income can often be considered the winning variable, especially in the central regions (wealth (-)). Unemployment is found to be the winning variable in a few clustered areas in *Mezzogiorno*, Calabria, Sardinia, and Sicily. In the central part of the country, major differences between inland areas and municipalities on the Tyrrhenian and Adriatic seas are also seen. Levels for the first local indicator determined with GWPCA are reported in Figure 17.

Figure 17 – Quantile map of the local indicators for the first component determined from GWPCA for Italian municipalities. Source: Own Elaboration on IMAJINE WP 2 database.



A major difference that emerges is a higher indicator value in the inland areas, especially in central areas (i.e. Abruzzo, Lazio, and Marche) and in the north (especially Liguria) compared to the results found with standard PCA (see Figure 15). Also, regarding the local indicator, greater economic performance persists primarily around most urbanized areas in the north (Lombardy, Emilia-Romagna, and Veneto).

6.3.4 Local differences and multivariate reality

The evidence shown in this section is a preliminary attempt to capture the structure of local economic performance at a very refined scale. We acknowledge that this analysis cannot be considered a definition for a comprehensive composite indicator due to data constraints. In fact, more indicators should be considered, and our discussion of economic performance should be narrowed. This is the case for inputs that account for technology and innovation. However, the results confirm how a plurality of factors contributes to a description of local economic performance.

In addition, the analysis stresses the need to pursue multidimensionality at a local level to prevent policy makers from setting *ad hoc* policies and interventions. For example, the EU has been supporting integrated territorial investments to provide member States with a flexible tool for enabling efficient implementation of integrated actions through simplified financing. Importantly, those policies underline a geographical dimension that must be considered. Local composite indicators allow us to better consider differences at low geographical scale, target disadvantaged areas within certain regions, and develop very accurate policies that may help ensure policy effectiveness. This advantage is justified, from a technical perspective, by a large increase in representativeness of the first component indicator, where approximately 80% of variance is explained within each country thanks to the use of GWPCA. Finally, as the preliminary results show, the use of spatial techniques cannot be considered as merely related to academic interest. Rather, they may be used to develop local measures that can be considered by policy makers.

7. FINAL REMARKS

Regional units are the primary focus of the EU Cohesion Policy. This makes an analysis of the relationship between inequality and growth at regional level relevant. The study of regional growth and regional inequality requires considering the spatial dimension of these occurrences. In fact, spatial interactions between regional units may influence the respective dynamics of inequality and growth. Moreover, the consideration of the relative position of EU regions in relation to one another aids identification of the peculiarities of each regional economy and in developing place-based policies.

This report has focused on introducing the spatial dependence effect for use in analysing regional inequality and for studying the impact of inequality on regional growth. Specifically, the results presented in this report stress the inadequacy of traditional inequality measures that do not consider the geographical position of data, thus discarding the spatial interactions among neighbouring regional units. In this report, we extended the analysis of disparities among regional units by relying on recent methodological approaches that account for spatial and non-spatial components of inequality, where the former expresses the component of inequality driven by the spatial dependence relationships. The empirical analysis carried out at the NUTS 3 level highlights the dominance of the spatial component of inequality with respect to the non-spatial component throughout the period under consideration (1995-2019). However, the contribution from the non-spatial component to overall inequality appears to increase during economic crises, suggesting a reduction of spatial interactions among regional units during such crises.

The non-spatial component of inequality has been included in a spatially augmented Solow growth model to investigate the relationship between inequality and growth. The proposed model relates growth in GDP per worker with the non-spatial component of within-region inequality, which expresses the contribution from each region to overall inequality. This report aims to address an open research question that concerns the link between inequality and growth. The link between these phenomena has been widely investigating in the theoretical and empirical literature, often leading to conflicting results. Using the SDM, we found a positive correlation between inequality and growth. Specifically, our analysis was carried out at the NUTS 2 level, and the results show intraregional inequality drives greater regional growth. However, this positive relationship is particularly significant for more developed and transition regions. Hence, the promoting-growth effect of inequality is more obvious in regions with higher levels of development. The positive impact of intraregional disparities on growth is insignificant in less-developed regions that, hence, should benefit from policies tailored to suit their specific needs and that may be different from those implemented in transition and more-developed regions. Growth and convergence are assessed for these groups of regions, with less-developed regions showing a faster growth rate than other regions. These growth rates are likely facilitated by EU policies that are specifically targeted to this group of regions.

The complexity of the economic phenomena addressed in this report requires simplification. However, simplicity must come without neglecting important mechanisms that drive economies and societies. The idea of local economic performance and development as a multidimensional phenomenon has been discussed in this report, and a proposal for synthesising these different aspects as a composite indicator has been introduced. The Joint Research Centre of the European Commission has been working for several years on composite indicators that can provide policy makers with the ‘big picture’ on relevant matters. Such areas include lifelong learning, innovation, environmental pressure, education, and competitiveness. Hence, in this report we attempted to take advantage of the Horizon 2020 IMAJINE dataset to provide greater insight into economic performance at a local level in a multivariate sense.

Extensive evidence that supports a broader consideration of spatial effects and spatial scale in the debate of inequality and cohesion policies has been presented in this report. Finally, it is important to note that a low geographical scale helps ensure the models produce accurate results while enriching empirical findings. In addition, a low geographical scale may shed light on how space could be considered a relevant dimension to ensure *all places* are considered in the discussion of inequality.

APPENDIX 1

The Breusch Pagan test calculated for the SDM provides evidence of heteroscedasticity (i.e. non-constant variance) in the error terms. This motivates the estimation of a heteroscedastic spatial model using GL2SLS. The estimation results are reported in Table 13, while Table 14 displays the average direct, indirect, and total impact estimates.

Table 13 – Estimation results for spatial heteroscedastic model (1995-2019) using cross-sectional data, 274 NUTS 2 EU regions. Source: Own Elaboration on ARDECO database.

	Spatial Heteroscedastic Model GS2SLS
Variable	Coefficient (standard error)
Constant	0.1100 (0.1780)
$\ln y_{t-T}$	-0.0195***(0.0025)
$\ln Ineq$	0.0007**(0.0004)
$\ln(n + 0.05)$	0.0043 (0.0050)
$W \ln y_{t-T}$	0.0112 (0.0124)
$W \ln Ineq$	-0.0003 (0.0009)
$W \ln(n + 0.05)$	0.0001 (0.0138)
ρ	0.7323*(0.4295)
λ (Convergence Rate)	2,67%

Significance levels ***1%, **5%, *10%

Table 14 – Average direct, indirect, and total impacts of the explanatory variables for spatial heteroscedastic model, 274 NUTS 2 European regions, 1995-2019. Source: Own Elaboration on ARDECO database.

Variable	Average Direct Impact	Average Indirect Impact	Average Total Impact
$\ln y_{t-T}$	-0.0200***	-0.0109	-0.0309***
$\ln Ineq$	0.0007*	0.0008	0.0015
$\ln(n+0.05)$	0.0048	0.0116	0.0164
Significance levels ***1%, **5%, *10%			

The estimates obtained for the heteroscedastic spatial model have the same sign and are similar in magnitude to the estimates reported for SDM (see Table 5 and Table 6). Significant direct impacts are reported for the initial GDP per worker (giving evidence of conditional convergence) and for the inequality variable (revealing positive correlation with regional growth). Indirect impacts are significant only for the initial GDP per worker.

APPENDIX 2

As a robustness check for the SDM specification, we used the approach proposed by Elhorst (2010). According to this approach, we first calculated the Lagrange Multiplier (LM) tests. These tests, in their classical and robust versions, allow to assess if the spatial lag model or the spatial error specification are preferable to the non-spatial model (OLS). While the spatial lag model introduces a spatial lag on the dependent variable, the spatial error model introduces a spatial lag on the error term. The SDM generalizes both these specifications. According to Elhorst (2010) if, based on the LM tests, the spatial lag or the spatial error models or both are preferable to the OLS specification, the SDM should be estimated. Subsequently, the Likelihood Ratio (LR) test should be performed to assess if the SDM could be simplified by the spatial lag or the spatial error model (i.e. Common Factor hypothesis test). Results for LM and LR test are reported in Table 15.

Table 15 - LM Tests on spatial dependence and spatial error autocorrelation, LR tests on parameter restrictions (p-value in parenthesis). Source: Own Elaboration on ARDECO database.

LMerror	102.1700
	(0.0000)
LMlag	59.8890
	(0.0000)
Robust LMerror	46.3700
	(0.0000)
Robust LMlag	4.0906
	(0.0431)
LR err	7.0100
	(0.0715)
LR lag	17.3900
	(0.0006)

Results for the classic and robust LM tests indicate that the OLS model should be rejected in favour of both the spatial error and the spatial lag specifications (both LMerror and LMlag are significant, as well as Robust LMerror and Robust LMlag). These results indicate that the SDM specification should be estimated. The LR err and the LR lag tests are both significant. These results confirm the appropriateness of the SDM specification, that could not be simplified by the spatial error and the spatial lag models.

REFERENCES

Aghion, P., Howitt, P. (1997). *Endogenous growth theory*. Cambridge, MA: MIT Press.

Aghion, P., Caroli, E., Garcia-Penalosa, C. (1999). Inequality and economic growth: the perspective of the new growth theories. *Journal of Economic Literature*, 37: 1615-1660.

Alesina, A., Rodrik, A. (1994). Distributive politics and economic growth. *Quarterly Journal of Economics*, 109: 465-490.

Alkire, S., Santos, M.E. (2010). Acute multidimensional poverty: A new index for developing countries. *United Nations Development Programme Human Development Report Office Background Paper (2010/11)*.

Amos, O. (1988). Unbalanced regional growth and regional income inequality in the latter stages of development. *Regional Science and Urban Economics*, 18: 549–566.

Anselin, L. (1988). *Spatial econometrics: Methods and models*. Kluwer Academic Publishers, Dordrecht.

Arbia, G. (2001). The role of spatial effects in the empirical analysis of regional concentration. *Journal of Geographical Systems*, 3: 271-281.

Arbia, G., Benedetti, R., Espa, G. (1996). Effects of the MAUP on Image Classification. *Geographical Systems*, 3: 123-141.

Arbia, G., Piras, G. (2009). A new class of spatial concentration measures. *Computational Statistics & Data Analysis*, 53: 4471-4481.

Artelaris, P. (2017). Geographies of crisis in Greece: A social well-being approach. *Geoforum*, 84: 59-69.

Artelaris, P., Petrakos, G. (2016). Intraregional spatial inequalities and regional income level in the EU: Beyond the inverted-U hypothesis. *International Regional Science Review*, 39: 291-317.

Artelaris, P., Tsirbas, Y. (2018). Anti-austerity voting in an era of economic crisis: Regional evidence from the 2015 referendum in Greece. *Environment and Planning C*, 36: 589-608

Atems, B., Jones, J. (2015). Income inequality and economic growth: A panel VAR approach. *Empirical Economics*, 48: 1541–1561.

Azariadis, C. (1996). The economics of poverty traps part one: Complete markets. *Journal of Economic Growth*, 1: 449-486.

Azariadis, C., Drazen, A. (1990). Threshold externalities in economic development. *Quarterly Journal of Economics*, 105: 510-526.

Azzoni, R. (2001). Economic growth and regional income inequality in Brazil. *The Annals of Regional Science*, 35: 133-152.

Bandyopadhyay, D., Tang X. (2011). Understanding the economic dynamics behind growth–inequality relationships. *Journal of Macroeconomics*, 33: 14-32.

Barca, F., McCann, P., Rodríguez-Pose, A. (2012). The case for regional development intervention: Place-based versus place-neutral approaches. *Journal of Regional Science*, 52: 134-152.

Barro R. J. (1991). Economic growth in a cross section of countries. *Quarterly Journal of Economics*, 106: 407-443.

Barro, R. J. (2000). Inequality and growth in a panel of countries. *Journal of Economic Growth*, 5: 5-32.

Barro, R.J., Sala-i-Martin, X. (1992). Convergence. *Journal of Political Economy*, 100: 223-251.

Barro, R.J., Sala-i-Martin, X. (1995). *Economic Growth*. Boston, McGraw Hill.

Bartiaux, F., Maretti, M., Cartone, A., Biermann, P., Krasteva, V. (2019). Sustainable energy transitions and social inequalities in energy access: A relational comparison of capabilities in three European countries. *Global Transitions*, 1: 226-240.

Baumol, W. (1986). Productivity growth, convergence, and welfare: What the long-run data show. *American Economic Review*, 76: 1072-1085.

Becker, W., Saisana, M., Paruolo, P., Vandecasteele, I. (2017). Weights and importance in composite indicators: Closing the gap. *Ecological Indicators*, 80: 12-22.

Beckfield, J. (2019). *Unequal Europe: Regional integration and the rise of European inequality*. Oxford University Press.

Bell, N., Schuurman, N., Oliver, L., Hayes, M.V. (2007). Towards the construction of place-specific measures of deprivation: a case study from the Vancouver metropolitan area. *The Canadian Geographer/Le Géographe Canadien*, 51: 444-461.

Benabou, R. (1996). Inequality and growth. In B.S. Bernanke, J. J. Rotemberg (Eds). *NBER Macroeconomics Annual*, 11, MIT Press, pp. 11-92.

Benos, N., Karagiannis, S., Karkalakos, S. (2015). Proximity and growth spillovers in European regions: The role of geographical, economic and technological linkages. *Journal of Macroeconomics*, 43: 124-139.

Berry, B.J., (1988). Migration reversals in perspective: the long-wave evidence. *International Regional Science Review*, 11: 245–251.

Besley, T., Burgess, R. (2003). Halving global poverty. *Journal of Economic Perspectives*, 17: 3-22.

Bickenbach, F., Bode, E. (2008). Disproportionality measures of concentration, specialization and localization. *International Regional Science Review*, 31: 359-388.

Borts, G., Stein, J. (1964). *Economic growth in a free market*. New York, Columbia University Press.

Boulanger, P.M. (2014). *Elements for a comprehensive assessment of public indicators*. JRC Scientific and Policy Reports, Luxembourg: Publications Office of the European Union.

Brakman, S., Garretsen, H., van Marrewijk, C. (2015). Regional resilience across Europe: On urbanization and the initial impact of the Great Recession. *Cambridge Journal of Regions, Economy and Society*, 8: 141-148.

Buitelaar, E., Weterings, A., Ponds, R. (2017). *Cities, economic inequality and justice. Reflections and alternative perspectives.* London: Routledge.

Butkus, M., Cibulskiene, D., Maciulyte-Sniukiene, A., Matuzeviciute, K. (2018). What is the evolution of convergence in the EU? Decomposing EU disparities up to the NUTS 3 level. *Sustainability*, 10, 1552; doi:10.3390/su10051552.

Capello, R., Caragliu, A., Fratesi, U. (2015). Spatial heterogeneity in the costs of the economic crisis in Europe: Are cities sources of regional resilience? *Journal of Economic Geography*, 15: 951-972.

Carrozza, C. (2014). Democratizing expertise and environmental governance: Different approaches to the politics of science and their relevance for policy analysis. *Journal of Environmental Policy & Planning*, 17: 108–126.

Cartone, A., Postiglione, P. (2020) Principal component analysis for geographical data: The role of spatial effects in the definition of composite indicators. *Spatial Economic Analysis*, DOI: 10.1080/17421772.2020.1775876

Chatterji, M., Dewhurst, J.H. (1996). Convergence clubs and relative economic performance in Great Britain: 1977-1991. *Regional Studies*, 30: 31-40.

Cheshire, P., Malecki, E. (2004). Growth, development, and innovation: A look backward and forward. *Papers in Regional Science*, 83: 249-267.

Christopherson, S., Martin, R., Pollard, J. (2013), Financialisation: Roots and repercussions. *Cambridge Journal of Regions Economy and Society*, 6: 351–357.

Cingano, F. (2014) Trends in income inequality and its impact on economic growth. *OECD Social, Employment and Migration Working Paper* No 163. OECD Publishing.

Conley, T. J., Topa, G. (2002). Socio-economic distance and spatial patterns in unemployment. *Journal of Applied Econometrics*, 17: 303-327.

Corrado, L., Fingleton, B. (2012). Where is the economics in spatial econometrics? *Journal of Regional Science*, 52: 210-239.

Cowell, F. A. (2011). *Measuring inequality. Third edition.* Oxford: Oxford University Press.

Davies, S. (2011). Regional resilience in the 2008–2010 downturn: Comparative evidence from European Countries. *Cambridge Journal of Regions, Economy and Society*, 4: 369-382.

Davies, S., Hallet, M. (2002). Interactions between national and regional development, *HWWA Discussion Paper No. 207*.

Dawkins, C.J. (2004). Measuring the spatial pattern of residential segregation. *Urban Studies*, 41: 833-851.

De Dominicis, L. (2014). Inequality and growth in European regions: Toward a place-based approach. *Spatial Economic Analysis*, 9: 120-141.

De Dominicis, L., Florax, R., de Groot, H. L. F. (2008). A meta-analysis on the relationship between income inequality and economic growth. *Scottish Journal of Political Economy*, 55: 654-682

De Muro, P., Mazziotta, M., Pareto, A. (2011). Composite indices of development and poverty: an application to MDGs. *Social Indicators Research*, 104: 1-18.

Decancq, K., Lugo, M.A. (2013). Weights in multidimensional indices of wellbeing: An overview. *Econometric Reviews*, 32: 7-34.

Deininger, K., Squire, N. (1996). A new data set measuring income inequality. *World*

Bank Economic Review, 10: 565-591.

Di Cesare, S., Cartone, A., Petti, L. (2020). A New Scheme for the Evaluation of Socio-Economic Performance of Organizations: A Well-Being Indicator Approach. In M. Traverso, Petti L., Zamagni A. (Eds). *Perspectives on Social LCA. Contributions from the 6th International Conference*. Springer, pp. 25-34.

Dijkstra, L., Poelman, H., Rodríguez-Pose, A. (2020). The geography of EU discontent. *Regional Studies*, 54: 737-753.

Đokić, I., Fröhlich, Z., Rašić Bakarić, I. (2016). The Impact of the economic crisis on regional disparities in Croatia. *Cambridge Journal of Regions, Economy and Society*, 9: 179-195.

Donald, B., Glasmeier, A., Gray, M., Lobao, L. (2014). Austerity in the city: Economic crisis and urban service decline? *Cambridge Journal of Regions, Economy and Society*, 7: 3-15.

Dunford, M. (1993). Regional disparities in the European Community: Evidence from the REGIO databank. *Regional Studies*, 27: 727-743.

Durlauf, S. N., Johnson, P. A. (1995). Multiple regimes and cross-country growth behaviour. *Journal of Applied Econometrics*, 10: 365-384.

Easterlin, R. (1958). Long term regional income changes: Some suggested factors. *Papers and Proceedings of the Regional Science Association*, 4: 313-325.

Elhorst, J.P. (2001) Dynamic models in space and time. *Geographical Analysis,* 33: 119-140.

Elhorst, J.P. (2010) Applied spatial econometrics: raising the bar. *Spatial Economic Analysis*, 5(1): 9-28.

Elhorst, J.P., Piras, G., Arbia, G. (2010). Growth and convergence in a multiregional model with space-time dynamics. *Geographical Analysis*, 42(3):338-355.

Ertur, C., Le Gallo, J., Baumont, C. (2006). The European regional convergence process, 1980-1995: Do spatial dependence and spatial heterogeneity matter? *International Regional Science Review*, 29: 3-34.

Ertur, C., Koch, W. (2007). Technological interdependence and spatial externalities: Theory and evidence. *Journal of Applied Econometrics*, 22: 1033-1062.

European Commission (2001). *Unity, solidarity, diversity for Europe, its people and its territory. Second report on economic and social cohesion*. Luxembourg, Office for Official Publications of the European Communities.

European Commission (2014). *Investment for jobs and growth: Promoting development and good governance in EU regions and cities. Sixth report on economic, social and territorial cohesion*. Luxembourg, Publications Office of the European Union.

European Commission (2018). *The EU-wide income distribution: Inequality level and decomposition*. Luxembourg, Publications Office of the European Union.

Fan, C., Casetti, E. (1994). The spatial and temporal dynamics of US regional income inequality, 1950-1989. *The Annals of Regional Science*, 28: 177-196.

Forbes, K. (2000). A reassessment on the relationship between inequality and growth. *American Economic Review*, 90: 869-887.

Frank, M. (2008). Inequality and growth in the United States: Evidence from a new state-level panel of income inequality measures. *Economic Inquiry*, 47: 55-68.

Fujita, M., Krugman, P., Venables, A. (1999). *The spatial economy*. Cambridge, MA: MIT Press.

Galor, O. (1996). Convergence? Inferences from theoretical models. *Economic Journal*, 106: 1056-1069.

Galor, O. (2000). Income distribution and the process of development. *European Economic Review*, 44: 706-712.

Gini, C. (1912). *Variabilità e mutabilità*. Contributo allo studio delle distribuzioni e delle relazioni statistiche. Bologna, C. Cuppini.

Giorgi, G.M. (2011). The Gini inequality index decomposition. An evolutionary study. In J. Deutsch, J. Silber (Eds). *The measurement of individual well-being and group inequalities: Essays in memory of Z.M. Berrebi*. London: Routledge, pp. 185-218.

Greyling, T., Tregenna, F. (2017). Construction and analysis of a composite quality of life index for a region of South Africa. *Social Indicators Research*, 131: 887-930.

Haining, R. P. (2003). *Spatial data analysis: theory and practice*. Cambridge university press, Cambridge.

Harris, P., Brunsdon, C., Charlton, M. (2011). Geographically weighted principal components analysis. *International Journal of Geographical Information Science*, 25: 1717-1736.

Harris, P., Clarke, A., Juggins, S., Brunsdon, C., Charlton, M. (2015). Enhancements to a geographically weighted principal component analysis in the context of an application to an environmental data set. *Geographical Analysis*, 47: 146-172.

Havard, S., Deguen, S., Bodin, J., Louis, K., Laurent, O., Bard, D. (2008). A small-area index of socioeconomic deprivation to capture health inequalities in France. *Social Science & Medicine*, 67: 2007-2016.

Heshmati, A. (2006). The world distribution of income and income inequality: A review of the economics literature. *Journal of World-Systems Research*, 12: 61-107.

Hirschman, A. (1958). *The strategy of economic development*. New Haven: Yale University Press.

Iammarino, S., Rodríguez-Pose, A., Storper, M. (2019). Regional inequality in Europe: evidence, theory and policy implications. *Journal of Economic Geography*, 19: 273-298.

Janikas, M., Rey, S. (2005). Spatial clustering, inequality and income convergence in the U.S.: 1969-2001. *Région et Développement*, 21: 45-64.

Jolliffe, I.T. (2002). *Principal component analysis. Second edition*. New York: Springer-Verlag.

Kaldor, N. (1956). Alternative theories of distribution. *Review of Economic Studies*, 23: 83-100.

Kaldor, N. (1970). The case for regional policies. *Scottish Journal of Political Economy*, 17: 337-348.

Kelejian, H.H., Prucha, I.R. (2010). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. *Journal of Econometrics*, 157: 53–67.

Knowles, S. (2005). Inequality and economic growth: The empirical relationship reconsidered in the light of comparable data. *The Journal of Development*, 41: 135-159.

Koo, J., Song, Y. (2016). The relationship between income inequality and aggregate saving: an empirical analysis using cross-country panel data. *Applied Economics*, 48: 892-901.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99: 183-199.

Kuc-Czarnecka, M., Piano, S.L., Saltelli, A. (2020). A quantitative storytelling in the making of a composite indicator. *Social Indicators Research*, 249: 775-802.

Kuznets, S. (1955). Economic growth and income inequality. *American Economic Review*, 45: 1-28.

Le Gallo, J., Ertur, C. (2003). Exploratory spatial data analysis of the distribution of regional per capita GDP in Europe: 1980-1995. *Papers in Regional Science*, 82: 175-201.

Lerman, R., Yitzhaki, S. (1984). A note on the calculation and interpretation of the Gini index. *Economics Letters*, 15: 363-368.

LeSage, J., Fischer, M. M. (2008). Spatial growth regressions: Model, specification, estimation and interpretation. *Spatial Economic Analysis*, 3: 275-304.

LeSage, J. P., Pace, R. K. (2009). *Introduction to spatial econometrics*. Boca Raton, Taylor & Francis.

Lessmann, C. (2014). Spatial inequality and development. Is there an inverted–U relationship? *Journal of Development Economics*, 106: 35-51.

Li, H., Zou, H. (1998). Income inequality is not harmful for growth: Theory and evidence. *Review of Development Economics*, 2: 318-334.

Lucas, R. (1988). On the mechanics of economic development. *Journal of Monetary Economics*, 22: 3-42.

Magrini, S. (2004). Regional (DI)Convergence. In J.V. Henderson, J.F. Thisse (Eds.), *Handbook of Regional and Urban Economics*, Elsevier, vol.4, pp. 2741-2796.

Mankiw, N.G., Romer, D., Weil, D.N. (1992). A contribution to the empirics of economic growth. *The Quarterly Journal of Economics*, 107: 407-437.

Márquez, M. A., Lasarte-Navamuel, E., Lufin, M. (2019). The role of neighbourhood in the analysis of spatial economic inequality. *Social Indicators Research*, 141: 245-273.

Mazziotta, M., Pareto, A. (2019). Use and misuse of PCA for measuring well-being. *Social Indicators Research*, 142: 451-476.

Monastiriotis, V. (2011). Making geographical sense of the Greek austerity measures: compositional effects and long-run implications. *Cambridge Journal of Regions, Economy, and Society*, 4: 323-337.

Monfort, P. (2008). *Convergence of EU regions: Measures and evolution*. European Commission, Regional Policy, 01/2008, 1-20.

Mookherjee, D., Shorrocks, A. (1982). A decomposition analysis of the trend in UK income inequality. *Economic Journal*, 92: 886-902.

Myrdal, G. (1957). *Economic theory and underdeveloped regions*. London: Duckworth.

Munda, G., Nardo, M. (2009). Noncompensatory/nonlinear composite indicators for ranking countries: a defensible setting. *Applied Economics*, 41: 1513-1523.

Nardo, M., Saisana, M. (2008). OECD/JRC handbook on constructing composite indicators. Putting theory into practice. In *Proceedings of the NTTS (New Techniques and Technologies for Statistics) Seminar* (p. 16).

Nijkamp, P. (1990). Spatial developments in the united states of Europe: glorious victories or ignominious defeats? *Papers in Regional Science*, 69: 1-10.

Nussbaum, M. (2000). Women's capabilities and social justice. *Journal of Human Development*, 1: 219-247.

OECD (2008). Handbook on constructing composite indicators: Methodology and user guide. Paris: OECD Publishing.

OECD (2014). *Regional outlook: Regions and cities: Where policies and people meet*. Paris: OECD Publishing.

Paci, R., Pigliaru, F. (1999). Is dualism still a source of convergence in Europe? *Applied Economics*, 31: 1423-1436.

Padilla, C.M., Kihal-Talantikite, W., Vieira, V. M., Rossello, P., Le Nir, G., Zmirou-Navier, D., Deguen, S. (2014). Air quality and social deprivation in four French metropolitan areas—A localized spatiotemporal environmental inequality analysis. *Environmental Research*, 134: 315-324.

Paelinck, J. (1978). Spatial econometrics. *Economics Letters*, 1: 59-63.

Palaskas, T., Psycharis, Y., Rovolis, A., Stoforos, C. (2015). The asymmetrical impact of the economic crisis on unemployment and welfare in Greek urban economies. *Journal of Economic Geography*, 15: 973-1007.

Pampalon, R., Raymond, G. (2000). A deprivation index for health and welfare planning in Quebec. *Chronic Disease Canada*, 21: 104-113.

Pampalon, R., Hamel, D., Gamache, P., Philibert, M. D., Raymond, G., Simpson, A. (2012). An area-based material and social deprivation index for public health in Québec and Canada. *Canadian Journal of Public Health/Revue Canadienne de Santé Publique*, 103: S17-S22.

Panizza, U. (2002). Income inequality and economic growth: Evidence from American data. *Journal of Economic Growth*, 7: 25–41

Panzera, D., Postiglione, P. (2014). Economic growth in Italian NUTS 3 provinces. *The Annals of Regional Science*, 53: 273-293.

Panzera, D., Postiglione, P. (2020). Measuring the spatial dimension of regional inequality: an approach based on the Gini correlation measure. *Social Indicators Research*, 148: 379-394.

Partridge, M.D. (1997). Is inequality harmful for growth? Comment. *American Economic Review*, 87: 1019-1032.

Partridge, M.D. (2005). Does income distribution affect US state economic growth? *Journal of Regional Science*, 45: 363-394.

Pekkala, S. (2000). Aggregate economic fluctuations and regional convergence: The Finnish case 1988–1995. *Applied Economics*, 32: 211-219.

Perotti, R. (1996). Growth, income distribution, and democracy: What the data say. *Journal of Economic Growth*, 1: 149-187.

Persson, T., Tabellini, G. (1994). Is inequality harmful for growth? Theory and evidence. *American Economic Review*, 84: 600-621.

Perugini, C., Martino, G. (2008). Income inequality within European regions: determinants and effects on growth. *Review of Income and Wealth*, 54: 73–406.

Petrakos, G., Artelaris, P. (2009). European regional convergence revisited: a weighted least squares approach. *Growth and Change*, 40: 319-331.

Petrakos, G., Saratsis, Y. (2000). Regional inequalities in Greece. *Papers in Regional Science*, 79: 57-74.

Petrakos, G., Psycharis, Y. (2016). The spatial aspects of economic crisis in Greece. *Cambridge Journal of Regions, Economy and Society*, 9: 137-152.

Petrakos, G., Kallioras, D., Anagnostou, A. (2011). Regional convergence and growth in Europe: understanding patterns and determinants. *European Urban and Regional Studies*, 18: 375-391.

Petrakos, G., Rodriguez-Pose, A., Rovolis, A. (2005). Growth, integration and regional disparities in the European Union. *Environment and Planning A*, 37: 1837-1855.

Plummer, P., Taylor, M. (2001). Theories of local economic growth: A model specification and empirical validation. *Environment and Planning A*, 33: 385-398.

Postiglione, P., Benedetti, R., Lafratta, G. (2010). A regression tree algorithm for the identification of convergence clubs. *Computational Statistics & Data Analysis*, 54: 2776-2785.

Postiglione, P., Andreano, M.S., Benedetti, R. (2013). Using constrained optimization for the identification of convergence clubs. *Computational Economics*, 42: 151-174.

Potter, J., Marchese, M. (2010). A review of local economic and employment development policy approaches in OECD countries: case studies of regional economic development approaches, OECD.

Ravallion, M. (2007). Inequality is bad for the poor. In S. Jenkins, J. Micklewright (Eds), *Inequality and poverty re-examined*. Oxford: Oxford University Press.

Rey, S.J. (2004). Spatial analysis of regional income inequality. In M.F. Goodchild, D.G. Janelle (Eds). *Spatially integrated social science*. Oxford: Oxford University Press, pp. 280-299.

Rey, S.J, Montouri, B. (1999). US regional income convergence: A spatial econometric perspective. *Regional Studies*, 33: 143-156.

Rey, S.J., Janikas, M.V. (2005). Regional convergence, inequality, and space. *Journal of Economic Geography*, 5: 155-176.

Rey, S.J., Smith, R.J. (2013). A spatial decomposition of the Gini coefficient. *Letters in Spatial and Resource Sciences*, 6: 55-70.

Rodríguez-Pose, A., Fratesi, U. (2007). Regional business cycles and the emergence of sheltered economies in the southern periphery of Europe. *Growth and Change*, 38: 621-648.

Romer, P. (1986), Increasing returns and long run growth. *Journal of Political Economy*, 94: 1002-1037.

Royuela, V., García, G. (2015). Economic and social convergence in Colombia. *Regional Studies*, 49: 219-239.

Royuela, V., Veneri, P., Ramos, R. (2019). The short-run relationship between inequality and growth: evidence from OECD regions during the Great Recession. *Regional Studies*, 53: 574-586.

Saltelli, A. (2007). Composite indicators between analysis and advocacy. *Social Indicators Research,* 81: 65-77.

Schechtman, E., Yitzhaki, S. (1987). A measure of association based on Gini's mean difference. *Communications in Statistics - Theory and Methods*, 16: 207-231.

Schuh, B., et al. (2019). *Research for AGRI Committee – The EU farming employment: current challenges and future prospects*. European Parliament, Policy Department for Structural and Cohesion Policies, Brussels.

Sen, A. (1980). Equality of what? In McMurrin S., (ed.) *Tanner Lectures on Human Values, volume I.* Cambridge, Cambridge University Press - Salt Lake City, University of Utah Press.

Sen, A. (1985). Well-being, agency and freedom: The Dewey lectures 1984. *The Journal of Philosophy*, 82: 169-221.

Shorrocks, A.F. (1980). The class of additively decomposable measures. *Econometrica*, 48: 613-625.

Shorrocks, A., Wan, G. (2005). The spatial decomposition of inequality. *Journal of Economic Geography*, 5: 59-81.

Silber, J. (2012). *Handbook of income inequality measurement*. Springer Science & Business Media.

Soja, E. (2009). The city and spatial justice. *Spatial Justice*, 1: 31-38.

Solow, R.M. (1956). A contribution to the theory of economic growth. *Quarterly Journal of Economics*, 70: 65-94.

Stiglitz, J., Sen, A., Fitoussi, J. (2009). Report by the commission on the measurement of economic performance and social progress.

Terrasi, M. (1999). Convergence and divergence across Italian regions. *The Annals of Regional Science*, 33: 491-510.

Thorbecke, E., Charumilind, C. (2002). Economic inequality and its socioeconomic impact. *World Development*, 30: 1477-1495.

Tobler, W. R. (1970). A computer movie simulating urban growth in the Detroit Region. *Economic Geography Supplement*, 46: 234-240.

Tobler, W. R. (2004). On the first law of geography: A reply. *Annals of the Association of American Geographers*, 94: 304-310.

Townsend, P. (1987). Deprivation. *Journal of Social Policy*, 16: 125-146.

United Nations (2020). *World social report 2020. Inequality in a rapidly changing world*. United Nations Publication.

Voitchovsky, S. (2005). Does the profile of income inequality matter for economic growth? *Journal of Economic Growth*, 10: 273-296.

Wang, Z., Ge, Z. (2004). Convergence and transition auspice of Chinese regional growth. *The Annals of Regional Science*, 38: 727-739.

Weide, R., Milanovic, B. (2018). Inequality is bad for growth of the poor (but not for that of the rich). *The World Bank Economic Review*, 32: 507-530.

Williamson, J.G. (1965). Regional inequality and the process of national development: a description of the patterns. *Economic Development and Cultural Change*, 13: 3-45.

Yitzhaki, S. (1994). Economic distance and overlapping of distributions, *Journal of Econometrics*, 61: 147-159.