



Universidad Autónoma
de Madrid

Biblos-e Archivo
Repositorio Institucional UAM

Repositorio Institucional de la Universidad Autónoma de Madrid
<https://repositorio.uam.es>

Esta es la **versión de autor** del artículo publicado en:
This is an **author produced version** of a paper published in:

Regional Studies 53.8 (2019): 1207 - 1220

DOI: <https://doi.org/10.1080/00343404.2019.1566698>

Copyright: © 2019 Taylor and Francis

El acceso a la versión del editor puede requerir la suscripción del recurso
Access to the published version may require subscription

The Drivers of Local Income Inequality: A Spatial Bayesian Model Averaging Approach

Miriam Hortas-Rico ^{*}and Vicente Rios [†]

December 11, 2018

Abstract

This study analyses the drivers of local income inequality in Spain. We derive a novel data set of inequality metrics for a sample of municipalities over the period 2000-2006. Spatial Bayesian Model Selection and Model Averaging techniques are used in order to examine the empirical relevance of i) spatial functional forms, ii) spatial weight matrices and iii) a large set of factors that could affect inequality. Our findings suggest that local inequality is mainly explained by human capital, economic factors and local politics. In addition, the use of Bayesian Geographically Weighted regressions provides evidence in favour of spatially heterogeneous effects.

Keywords: Income Inequality, Spatial Econometrics, Model Averaging, Spanish Municipalities.

JEL codes: C11, C15, C21, D31.

FUNDING: This work was supported by the Fundacion Ramon Areces under Grant number CISP16A4784.

^{*}Department of Economics and Public Finance, Universidad Autónoma de Madrid, Ciudad Universitaria de Cantoblanco, 28049 Madrid, Spain. E-mail: miriam.hortas@uam.es

[†]Department of Economics, Universidad Pública de Navarra, Campus Arrosadia, 31006 Pamplona, Spain. E-mail: vicente.rios@unavarra.es

INTRODUCTION

The study of income inequality is a central issue in economic research and politics, as it is considered to be “one of the biggest social, economic and political challenges of our time” (The Economist, 2012). Although it is widely recognized that some degree of inequality in the distribution of income is necessary in order to provide incentives to the most skilled and productive individuals, inequality has a number of negative consequences on the functioning of the society. In their study, Dabla-Norris et al. (2015) assert that inequality: (i) erodes the functioning and quality of the political systems, as it can concentrate political power in the hands of a few elites, resulting in misallocation of resources, corruption and nepotism, (ii) raises the risk of fiscal and financial crises, causing economic instability and reducing growth, (iii) hampers poverty reduction, and (iv) may damage social cohesion and fuel civil and social conflicts, as it lowers intergenerational income mobility and opportunities of the poor.

The widespread rise of income inequality has exacerbated the gap between the rich and the poor. Available data show that the magnitude of inequality and its sources vary markedly not only across countries but also between municipalities (see Berube & Holmes, 2015; Musterd et al., 2016). Yet, relatively little is known about income inequality at the local level.

In this paper, we take advantage of this intermunicipal variation to present a comprehensive portrait of the patterns and causes of local inequality. We draw on a novel data set of local inequality metrics and implement Spatial Bayesian Model Averaging techniques (SBMA, hereinafter) to analyse the drivers of inequality in a representative sample of Spanish municipalities over the period 2000-2006. The paper makes several novel contributions to the literature.

First, following the methodology described in Hortas-Rico et al. (2014), we derive a novel data set on local income distributions employing Personal Income Tax (PIT) micro-data. With this information in hand, a set of income-related summary measures (including average income and inequality measures) is calculated for those municipalities with more than 5,000 inhabitants.

Second, given the uncertainty surrounding spatial interactions among municipali-

ties we perform a Spatial Bayesian Model Selection following LeSage (2014), Da Silva et al. (2017) and Rios (2017). In doing so, we extend previous research on local disparities, so as to account for different patterns of spatial interaction and different spatial weight matrices to describe the levels of income inequality among Spanish municipalities. It represents, as such, a novel application at the local level, given that previous studies of local income inequality (see, e.g. Levernier et al., 1998, Glaeser et al., 2009 or Florida & Mellander, 2016) omitted the role of spatial interactions in shaping local inequality patterns.

Third, income inequality is analysed by employing a set of thirty possible explanatory variables that are expected to affect local inequality patterns. Compared with the limited set of regressors considered in the existing empirical literature, this study rigorously assesses model uncertainty over a larger set of inequality determinants and minimizes omitted variable bias. These covariates can be grouped into six categories: (i) economic factors, (ii) demographic characteristics (iii) human capital factors, (iv) fiscal policy, (v) local politics and (vi) local amenities, deemed important for location decisions and population sorting.

Fourth, we draw on the work by LeSage & Parent (2007) and Crespo-Cuaresma et al. (2014) and employ SBMA techniques to perform inference. This methodology is particularly useful to address model uncertainty when there is spatial dependence in the data. Spatial dependence can lead to biased and inconsistent least-squares estimates that at best will be inefficient, invalidating the use of conventional Bayesian model averaging (BMA) techniques. Hence, contrary to previous studies on inequality where inference is based in single econometric model analysis, the SBMA approach has the advantage of minimizing the likelihood of producing (i) biased estimates and (ii) artificially low confidence intervals (Moral-Benito, 2015).

Finally, we consider the heterogeneity in the estimated effects of the different regressors across the Spanish geography by means of a Bayesian Geographically Weighted Regression (BGWR) analysis, following the methodology developed by LeSage (2004). This analysis is intended to provide a better understanding of the drivers of local income inequality across the spatial dimension, which is essential to formulate efficient public policies aimed at curbing local income disparities and to evaluate their effectiveness.

The reminder of the paper is organized as follows. The next section reviews the

literature on the determinants of local income inequality. The third section describes the data used in the analysis and presents an exploratory spatial analysis on the importance of spatial dependence. The fourth section describes the econometric methodology, while the main results are presented in the fifth section. Finally, the last section concludes.

THE CAUSES OF INCOME INEQUALITY: THEORETICAL CONSIDERATIONS

The literature analysing the drivers of inequality at the country-level has identified various driving forces behind the increase in global inequality levels during the last three decades: (i) technological change and skill premium, (ii) globalization and financialization, (iii) the decline of some labour market institutions such as trade-union membership, traditionally responsible of the compression in the distribution of wages, and (iv) the progressivity decline of the tax systems in many advanced economies (Dabla-Norris et al., 2015). However, local-level income inequality is likely to be different from national-level income inequality, as it neither responds to the same factors, nor creates the same policy implications (Glaeser et al., 2009). Therefore, the explanatory power of cross-country stylized facts could be compromised at the local level (Levernier et al., 1998).

The dynamics of urban structure are evolving over time. Municipalities have experienced a reorganization of their production and consumption activities, driven by supply factors (e.g., knowledge spillovers and agglomeration economies) as well as demand factors (e.g., preferences for local amenities or redistributive policies). As a result, individuals with different endowments of education and abilities have been differentially exposed to the opportunities offered by this new economic landscape (Florida & Mellander, 2015).

To date, however, only a handful of studies have paid attention to the local dimension of income inequality (e.g. Glaeser et al., 2009). The study of income disparities at the local level is particularly relevant given that not only the 54% of the world population lives in cities but also because the widening gap between the bottom- and the upper-tail of the income distribution is leading to important income segregation

in many cities worldwide which, in turn, can compromise the social stability and the competitive power of cities as engines of growth (Musterd et al., 2017).

According to the literature, income inequality is driven by a myriad of factors. Next we summarize the main findings of previous theoretical and empirical studies, distinguishing between inequality-enhancing and inequality-hindering factors.

Human Capital Factors

Human capital is widely viewed as a major factor determining the distribution of income. Access to education increases the job opportunities and the earning potential of the poor, facilitating their upward mobility. On the other hand, it allows for a more informed participation in the market economy, thus reducing the lobbying ability of the rich (Rodriguez-Pose & Tselios, 2009). Educational attainment has an equalizing effect on income distributions and, therefore, educational inequality is expected to be positively correlated to income inequality. Florida (2002) popularized the creative class, an occupational skill concept of a person's capability that reflects accumulated experience, creativity, intelligence, innovativeness and entrepreneurial abilities (Florida et al., 2008). A greater share of these high-skill workers is expected to be a source of income inequality because of their significantly higher-than-average earnings level, given the existence of positive returns to investment in human capital (Mincer, 1958; Cloutier, 1997). In addition, as noted in Florida (2002), income inequality is an unavoidable externality of the rise of the creative class given that a higher proportion of higher-income professionals increases the demand for low-income workers, hence widening the income gap. On the other hand, an increasing share of low-skilled workers could decrease inequality given the lower dispersion in bottom tail of the income distribution and their compression effect on the overall distribution of income (Izquierdo & Lacuesta, 2007). Nonetheless, its effect ultimately depends on the level of minimum wages and the skill distribution across jurisdictions.

Economic Factors

Other studies identify the connection between economic factors and income inequality. Even though the a priori effect of average income on inequality is uncertain

-depending on whether the increase in income is pro-poor, pro-rich or neutral-, most studies find a positive link between these two variables, reflecting that economic development seems to increase the occupational choices and the earning opportunities of the rich (Rodriguez-Pose & Tselios, 2009; Florida & Mellander, 2016). An additional factor that may affect local inequality is the municipality's share of value added over the total value added, whose impact on inequality is certainly ambiguous. The effect of unemployment is also undetermined. A higher unemployment rate decreases the access to wages, which is the main source of income, but its net impact on income inequality would depend on whether unemployment inflows have a larger impact at lower or higher wage segments. In addition, job protection and unemployment benefits are key factors in shaping income distributions, as they can lower inequality through smaller income dispersion (OECD, 2012a). A large literature has documented the industry mix effects on inequality. On the one hand, manufacturing, services and construction sectors have historically allowed low qualified workers to earn relatively high-wages, suggesting that increases in these sectoral shares should lead to more equal income distributions (Cloutier, 1997). On the other hand, the agriculture sector is expected to increase income inequality because the strong variation in farm size leads to a greater dispersion of earnings between farmers (Levernier et al., 1995). As concerns female participation, some previous empirical evidence finds a negative effect (Levernier et al., 1998; Rodriguez-Pose & Tselios, 2009). However, this result is somewhat unexpected, since its net effect ultimately depends on whether increasing female labour-force participation contributes or not to narrowing the existing gender-wage gap (Gonzales et al., 2015). If female have lower-than-average earnings due to either shorter working hours or wage discrimination in the labour market, then an increase in female participation should have an income inequality-enhancing effect (OECD, 2012a). Other studies have focused their attention on housing values, since increasing housing prices favour landlords and rentiers with respect to renters, thus contributing to an increasing net capital income gap between the two groups (Rognlie, 2015). Recent evidence suggests that differences in income have been capitalized into housing prices, reducing returns to living in productive places net of housing costs for unskilled workers, hence forcing them to move out and locate in less productive areas (Ganong Shoag, 2017). In addition, increasing housing prices may exacerbate inequality and block access to opportunities for upward mobility given that as rent consumes an increasing share of a family's budget, educational expenditures decrease, undermining youth prospects and impeding mobility. In the same vein, Bonnet et al. (2014) cast some doubt on the impact of homeownership rates on income inequality, since the

inter-generational inequality could be balanced by intra-generational equality gains.

Demographic Factors

Demographic characteristics are also deemed important in shaping income distributions. For instance, the share of immigrants is expected to be a factor in widening inequality, since they tend to earn less than natives (even for similar levels of education) and are likely competing for low-skill jobs, increasing the supply of less-skilled labour and depressing the wages of low-income workers, hence exacerbating existing wage differentials (Topel, 1994). Similarly, older populations tend to be characterized by higher income inequality as a result of the accumulation of sequences of transitory shocks that hit households at the lower end of the wage and salary distribution (Güvenen et al., 2015). On the contrary, increasing household size is expected to produce consumption complementarities and consumption scale economies, reducing the need for high per capita income and thickening the bottom-tail of income distribution (OECD, 2012b). Variations in population size reflect differences in agglomeration economies and cost-of-living differentials. Yet, the relationship between population size and income inequality is ambiguous. On the one hand, highly populated areas are expected to exhibit higher inequality levels as existing urban agglomeration effects disproportionately benefit high-skilled workers given that knowledge tends to diffuse at a faster rate than in less populated places, which ultimately reinforces the skill premium (Levernier et al., 1995, 1998). On the other hand, a greater population with increased opportunities for suitable job matches might favour less-skilled workers if they are less geographically mobile (Levernier et al., 1998). In the same vein, Cloutier (1997) argues that increases in population size tend to reduce inequality through improved capital markets that increase investment in human capital and reduce the average rate of return.

Local Government Factors

Governments have also been active players driving income inequality. As noted in previous studies, income inequality also depends on local policy decisions (Glaeser et al., 1995). As explained in Afonso et al., (2010), the level of taxation and its pro-

gressivity is perhaps the most direct factor reducing income inequality, whereas public spending can affect income distribution directly (e.g., via income transfers and cash payments to poorer individuals) or indirectly (e.g., via spending decisions that improve productivity and job access to the less well off). Similarly, intergovernmental transfers are expected to reduce income inequality, since they provide local governments with an additional source of revenues for public service provision.

Income inequality may also depend on factors related to the political preferences and stability of the local council (Tiebout, 1956). Thus, a higher share of votes to left-wing parties is expected to reduce inequality, as parties on the left of the political spectrum tend to promote redistributive policies. The role of government strength and partisan ideological alignment is uncertain, as it depends on the amount of grants received by local governments as well as on the composition of public expenditures. Finally, corruption is expected to be positively associated to income inequality given that it can increase rent-seeking and bribe-taking, thus distorting local governments' efficiency and resource allocation, increasing the operating costs of government and reducing the amount of revenues available for other (welfare/social) services. Similarly, corrupt governments are likely to take biased decisions favouring the well-connected individuals, which are commonly the wealthy elites (Gupta et al., 2002).

Local Amenities

Attempting to explain changes in the income distribution, economists have also considered the impact of local amenities (such as crime rates, urban blight, accessibility or leisure activities, among others). The effect of the amenity covariates is somewhat uncertain and depends on the population shifts they induce and how they affect the population composition in both the origin-destination municipalities. According to the Tiebout's model, individuals are mobile and sort themselves into jurisdictions according to their preferences, so that they create homogeneous communities of like income or race. This mobility hypothesis is particularly relevant among rich people for whom mobility costs are lower. In this setting, certain positive amenities (such as road accessibility) may draw high-income residents to a particular location, whereas certain city problems (such as crime or blight) might encourage their flight from blight.

DATA

In the analysis that follows, municipalities are taken as our geographical unit and the Gini Index as our measure of income inequality. The sample data covers almost all Spanish municipalities with 5,000 or more inhabitants. Data availability forced us to exclude 280 observations, leaving us with a sample of 977 municipalities.¹ This is sufficiently representative given that they account for about 85% of the total population. In addition, municipalities above 5,000 inhabitants have the advantage of corresponding to natural political units and they are large enough to provide inequality measures with statistical precision, which makes them suitable for analysing and discussing public policy.

The analysis covers the 2000-2006 time period. In particular, the income inequality variable is measured for the year 2006 and most of the explanatory variables are taken in 2001 in order to avoid reverse causality problems. Note that the period of study is particularly relevant to the aim of this paper, since it covers the years of economic growth preceding the financial crisis.

The set of explanatory variables considered in this study has been selected according to the literature review presented in the previous Section. Their definitions, abbreviations, descriptive statistics, data sources and expected effects are presented in Appendix Table (A1).

Income Inequality Data

The initial aim is to develop an accurate measure of income inequality so that the driving forces of such phenomenon can be empirically tested. Despite its importance, local income data and, thus, inequality data, remain a key missing element within the official statistics of many developed countries, with Spain being no exception. Household income and expenditure surveys have a territorial representation limited to the regional level, as they do not have a sufficient sampling size to offer a reasonable precision at a smaller level.

To redress this lack of information, a wide range of statistical techniques have been developed over the last two decades aimed at providing reliable estimates of lo-

cal income. The majority often use micro-data information from surveys, combined with aggregate information about relevant variables for the considered population subgroups. There is, however, a growing body of empirical literature focusing on tax-based research (see, for instance, Atkinson et al., 2011). PIT samples have emerged as an interesting alternative for overcoming the aforementioned territorial representativeness limitations shown by household surveys when analysing personal income distributions.

Available micro data on PIT returns from the Spanish Tax Administration Office enable us to derive income distributions at the local level for a representative sample of municipalities. In this study we make use of the 2006 PIT sample, which includes 964,489 records extracted from a population of 17,840,783 personal income tax returns. These micro-level PIT samples are only representative at the provincial level and, therefore, a re-weighting procedure needs to be implemented to derive a representative income sample at the municipal level. As in Hortas-Rico et al. (2014), the methodology employed here relies on a distance function optimization-based approach for survey re-weighting that consists of adjusting the original micro-data sample weights. Then, local income distributions and selected summary measures can be derived. These data are uniquely qualified to the purposes of the paper; they are the only data containing detailed information on income inequality measures for Spanish municipalities. In particular, we calculate the Gini coefficient at the municipal level as our (pretax) income inequality measure. We use this index as the baseline measure of inequality, mainly because it is the ubiquitous standard in the inequality literature. This index is defined as:

$$G(y) = 1 - 2 \int_0^1 L(p; y) dp \quad (1)$$

where the Lorenz curve of income, $L(p; y)$, at such p -values of ranked relative cumulated-population (so that, $p \in (0, 1)$) can be defined mathematically by the expression:

$$p = F(q) \Rightarrow L(p; y) = \int_0^q y f(y) \frac{dy}{d\mu_y} \quad (2)$$

where p is a percentile function, $F(q)$ is the distribution function measuring the proportion of individuals of the population having a standard of living below or equal to

q and μ_y denotes the average income. Note that $G(y)$ takes values between 0 (perfect equality) and 1 (complete inequality).²

Preliminary Evidence on the Role of Space

With the aim of providing a first insight into the spatial pattern of inequality at the local level, Figure (1a) plots the contour map of the Gini index across Spanish municipalities. As shown, income inequality exhibits considerable spatial variation, ranging from a low of 0.25 to a high more than double that, 0.6. The spatial distribution of income inequality is quite complex, as marked spatial clusters of very low- and very-high income inequality characterize it.

[insert Figure (1) about here]

In addition, Figure (1b) (left panel) displays the Moran's scatter plot. The slope of the regression line is the Moran's I statistic, a measure of spatial correlation, and takes a value of 0.21 (p-value=0.00). This suggests that municipalities with high inequality are surrounded by municipalities with high inequality.³

As a further check on the role played by the spatial location of jurisdictions in explaining income inequality outcomes, we estimate a stochastic kernel following the methodology outlined by Magrini (2009).⁴ Stochastic kernel estimation allows capturing the transitions between the original distribution and the neighbour-relative income inequality distribution. To read this diagram note that a value of 1 on the horizontal axes indicates the Spanish average inequality rate, a value of 2 indicates twice the Spanish average, and so on. On the other hand, contour lines gives the probability that any particular municipality will experience that relative rate of income inequality. Estimated stochastic kernel results shown in Figure (1b) (right panel) reveal that the probability mass tends to be located parallel to the axis corresponding to the original distribution. Accordingly, spatial effects appear as a relevant factor explaining the observed variability in local inequality. These findings regarding the role of space suggest that it is necessary to accommodate such interdependence in the modelling process and that an explicit accounting for spatial effects is required by means of spatial econometric models.

ECONOMETRIC STRATEGY

Econometric studies of Levernier et al., (1998), Glaeser et al. (2009) or Florida & Mellander (2016) analyse inequality at the local level, but treat the units of analysis as isolated entities, ignoring the spatial characteristics of the data and the potential role of space in modulating the economic evolution of local income inequality.⁵ Nevertheless, insofar every municipality evolves interacting with other municipalities, as suggested by the preliminary evidence in Figure (1), major problems may arise if the spatial characteristics of the data are ignored. The consequences of omitting these interactions from the model specification are potentially important from an econometric perspective, and may cause estimates to become biased, inconsistent and/or inefficient (Elhorst, 2014).

Spatial Bayesian Model Selection

We begin by considering the *Spatial Error Model* (SEM) and the *Spatial Durbin Error Model* (SDEM) specifications which are given by Equations (3) and (4), respectively:

$$\begin{aligned} y &= \alpha \iota_n + X\beta\epsilon \\ \epsilon &= \lambda W\epsilon + v \end{aligned} \tag{3}$$

and

$$\begin{aligned} y &= \alpha \iota_n + X\beta + WX\theta + \epsilon \\ \epsilon &= \lambda W\epsilon + v \end{aligned} \tag{4}$$

where y denotes a $N \times 1$ dimensional vector consisting of observations for the Gini index in the year 2006 for municipality $i = 1, \dots, N$ and X is an $N \times K$ matrix of exogenous aggregate political, socioeconomic and economic covariates with associated response parameters β contained in a $K \times 1$ vector. α reflects the constant term, ι_n is a $N \times 1$ vector of ones while λ is the spatial diffusion coefficient which captures spatially correlated shocks working through the error term $W\epsilon$. W is a $N \times N$ row-standardized matrix of known constants describing the spatial arrangement of the

municipalities in the sample. Finally, $v = (v_1, \dots, v_N)'$ is a vector of i.i.d disturbances whose elements have zero mean and finite variance σ^2 . In addition, the SDEM includes the spatial lag of the rest of control variables (exogenous effects), WX , whose impact is reflected by the $K \times 1$ vector of coefficients θ . Additionally, we also consider the *Spatial Lag Model* (SLM) and the *Spatial Durbin Model* (SDM) which are given by Equations (5) and (6)

$$y = \alpha\iota_n + \rho Wy + X\beta + v \quad (5)$$

$$y = \alpha\iota_n + \rho Wy + X\beta + WX\theta + v \quad (6)$$

The SDEM/SEM do not require a theoretical model for spatial or social interaction processes, as it is common in the case in spatial models including endogenous interactions such as the SDM/SLM. Indeed, as explained by Gibbons & Overman (2012) and Halleck-Vega & Elhorst (2015), spatial models containing endogenous interactions such as the SDM/SLM are generally difficult to justify from a theoretical basis. In the context of income inequality, endogenous interactions would lead to a scenario where changes in one jurisdiction set in motion a sequence of adjustments in (potentially) all units in the sample such that a new long-run steady state equilibrium of income inequality arises.⁶ A feature of the SDEM/SEM type of models is that they highlight the presence of omitted variables with explicit spatial patterns.

Another, relevant source of model uncertainty in spatial econometrics is the spatial weights matrix. Given that this is a relevant issue in spatial econometric modelling, a broad range of alternative specifications of W are considered. The first spatial weights matrix is based on the concept of first order contiguity, according to which $w_{ij} = 1$ if jurisdictions i and j are physically adjacent and 0 otherwise. Secondly, several matrices based on the k -nearest neighbours computed from the great circle distance between the centroids of the various jurisdictions. Furthermore, as is common practice in applied research, all the matrices are row-standardized, so that it is relative, and not absolute, distance that matters.

In order to choose between different potential specifications of the spatial weight matrix W , as well as to choose between SDM, SLM, SDEM and SEM specifications a Bayesian model comparison approach is applied following Da Silva et al. (2017) and

Rios (2017). This approach determines the posterior model probabilities (PMP) of the alternative specifications given a particular W , as well as the PMP of different spatial weight matrices given a particular model specification. However, a difference with respect to the Bayesian approach of model selection implemented in Da Silva et al. (2017) and Rios (2017) is that we do not use the full set of regressors for each spatial model (SM) and each W matrix in a single regression. To derive the probability of each spatial model SM conditional on W and X ($P(SM|W, X)$) and the probability of the W conditional on the spatial model SM and X ($P(W|SM, X)$), we averaged model probabilities over a large number of subsets of X .⁷ Proceeding in this way, we find the SEM appears to be the preferred spatial model specification, as its probability ranges from 74% to 99% depending on the W . Conditional to the SEM specification, the spatial weight matrix displaying the highest probability is the 7-nearest neighbour matrix (25.2%). Thus, the SBMA analysis presented in the next section relies on the SEM specification with a 7-nearest neighbour's spatial weight matrix. The results and technical details of the implementation of this procedure are reported in Table (A2) in the Appendix.

Spatial Bayesian Model Averaging

Spatial Bayesian Model Averaging (SBMA) allows us to consider all possible combinations of potential determinants of income inequality and takes a weighted average of the coefficients. Sub-structures of the model in Equation (3) are given by subsets of coefficients $\eta^k = (\delta^k, \lambda)$ and regressors X_k . Assuming that the total number of possible explanatory variables is K , the total number of possible models is 2^K and $k \in [0, 2^K]$. Inference on the parameters of the variables X explicitly takes into account model uncertainty and it is based on probabilistic weighted averages of parameter estimates of individual models:

$$p(\eta|y, X) = \sum_{k=1}^{2^K} p(\eta_k|M_k, y, X) p(M_k|y, X) \quad (7)$$

The weights, the PMP's are given by:

$$p(M_k|y, X) = \frac{p(y, X|M_k) p(M_k)}{\sum_{k=1}^{2^K} p(y, X|M_k) p(M_k)} \quad (8)$$

Model weights can be obtained using the marginal likelihood of each individual model after eliciting a prior over the model space. The marginal likelihood of model M_k is given by:⁸

$$p(y, X|M_k) = \int_0^\infty \int_{-\infty}^\infty \int_{-\infty}^\infty p(y, X|\delta, \lambda, \sigma, M_k) d\delta d\lambda d\sigma \quad (9)$$

The Posterior Mean (PM) of the distribution of η is:

$$E(\eta|y, X) = \sum_{k=1}^{2^K} E(\eta_k|M_k, y, X) p(M_k|y, X) \quad (10)$$

while the Posterior Standard Deviation (PSD) reads as:

$$PSD = \sqrt{Var(\eta|y, X)} \quad (11)$$

where the $Var(\eta|y, X)$ is given by:

$$Var(\eta|y, X) = \sum_{k=1}^{2^K} Var(\eta_k|M_k, y, X) p(M_k|y, X) + \sum_{k=1}^{2^K} (E(\eta_k|M_k, y, X) - E(\eta|y, X))^2 p(M_k|y, X) \quad (12)$$

where the first term reflects the variability of estimates across different regression models and the second term captures the weighted variance across different models. We compute the posterior inclusion probability (PIP) for a variable h as the sum of the PMP's including the variable h :

$$PIP = p(\eta_h \neq 0|y, X) = \sum_{k=1}^{2^K} p(\eta_k|M_k, y, X) p(M_k|\eta_h \neq 0, y, X) \quad (13)$$

Finally, we compute the conditional posterior positivity of a parameter h as:

$$p(\eta_h \geq 0|y, X) = \sum_{k=1}^{2^K} p(\eta_{k,h}|M_k, y, X) p(M_k|y, X) \quad (14)$$

where values of conditional positivity close to 1 indicate that the parameter is positive in the vast majority of considered models. Conversely, values near 0 indicate a predominantly negative sign.

We use the Monte Carlo Markov Chain Model Composition (MC^3) methodology for spatial models developed by LeSage & Parent (2007). The key feature of this econometric procedure is that it eliminates the need to consider all possible models by constructing a sampler that explores relevant parts of the large model space. The algorithm operates in the model space as follows. If we let M denote the current state of the chain, models are proposed using a neighbourhood, $nb(M)$ which consists on the model itself and models containing either one variable more (*birth step*) or one variable less (*death step*) than M . A transition matrix q , is defined by setting $q(M \rightarrow M') = 0$ for all $M' \notin nb(M)$ and $q(M \rightarrow M')$ constant for all $M' \in nb(M)$. The proposed model M' , is compared with the current model state M using the acceptance probability:

$$P = \min \left[1, \frac{p(M'|y)}{p(M|y)} \right] \quad (15)$$

The vector of log-marginal values for the current model M and the proposed alternative models M' are scaled and integrated to produce Equation (15). In addition to the birth and death steps, the sampler employed here includes the *move step*, which replaces randomly variables in X with variables not included currently in the model, which leaves the model proposal M' with the same dimension as M .

RESULTS

Main Results

Table (1) reports the results obtained under the SEM specification when implementing the MC^3 algorithm for the 10,000 top models out of the 17,365 generated by the sampler and a W matrix based on the 7-nearest neighbour's.⁹ However, before continuing with the discussion of the results in Table (1), it is worth mentioning the problems that the methodology applied here is able to solve and those problems that

may persist, affecting the quality of the estimates. The strong point of the SBMA methodology employed here is that it accounts for the uncertainty of the parameter estimates across different models when there is spatial dependence in the data, while controlling for omitted variable bias (LeSage & Parent, 2007; Moral-Benito, 2015). However, it does not correct for the potential negative effect of endogeneity caused by reverse causal relationships or measurement errors. In fact, how to tackle the issue of endogeneity in a model averaging framework is an important line of open research.¹⁰ To minimize the potential problems caused by reverse causality the explanatory variables are lagged five years. On the other hand, to deal with the effect of measurement errors and outliers, we allow for heteroskedasticity and perform a variety of robustness checks.

As usual in BMA exercises, the concentration of the posterior density in this context is very high. In particular, the top 1% models concentrate the 59.48% of mass, while the top 5% concentrate the 84.46%. We scale the PIPs of the different variables in quartiles to classify evidence of robustness of inequality regressors into four categories so that regressors with $PIP \in [0 - 25\%]$ are considered as weak determinants, variables with $PIP \in [26 - 50\%]$ as moderate determinants, with $PIP \in [50 - 75\%]$ as substantive and with $PIP \in [75 - 100\%]$ as highly important.

As observed, the group of very important determinants consists on the creative class (99.9%) and a wide range of economic factors, including the share of employment in services (99.9%) or manufacturing (99.9%), housing values (98.5%), average income (94.8%) or the unemployment rate (84.8%). In the group of substantive determinants we find the total spending at the regional level (74.8%), the educational inequality (56.1%), the level of corruption (55.2%) and the share of elderly population (51.5%). The local council's ideology (37.9%) and the total resident population (33.8%) are found within the set of moderate determinants. Finally, weak inequality drivers include other economic and demographic factors (e.g. female participation, the share of employment in agriculture or construction, the share of immigrants or low-skilled workers, and agglomeration economies) and local amenities (e.g. road accessibility, crime rates or urban blight). Overall, our findings suggest that, on the one hand, human capital and economic factors are the key factors shaping local income distributions, even though politics and fiscal policy factors also play a non-negligible role.

[insert Table (1) about here]

Columns (2) to (5) show the mean and the standard deviation of the posterior parameters distributions, along with the lower and upper 95% bounds, conditional on the variable being included in the model.¹¹ To complement these statistics, Column (6) reports the fraction of models where the t-stat of the corresponding variables is higher than 1.96 (which implies statistical significance at the 5% level), while Column (7) presents the results of the posterior sign certainty, which measures the posterior probability of a positive coefficient expected value, conditional on inclusion.

With the exception of the effect of female participation in the labour market, which is (i) positively related to inequality with certainty in all regressions in which the variable is included, and (ii) appears to be significant at 5 % level in the 68 % of the models, the results obtained for the other weak determinants do not allow to draw clear conclusions on the effect exerted on local income inequality. The reasons are twofold. First, in many cases the posterior sign certainty of these regressors is around 0.3-0.7, suggesting that both positive and negative effects can be observed. Consequently, the causal relationships for this group of variables are not robust. Second, the fraction of regressions where these variables exhibit t-stats above the 5% significance level is always below the 43% and even virtually 0% in the majority of cases. On the contrary, the groups of moderate, substantial and highly important determinants of inequality display robust sign effects (either positive or negative) and are significant at the 5% level in 88 to 100% of the models. Therefore, for the reminder of the paper we will only discuss the results for the regressors with a PIP above 25%.

First, we consider the impact of human capital. As expected, the creative class and the educational inequality are the primary indicators of income inequality. Both variables increase the proportion of higher-than-average income workers, all else equal, widening the income gap. In general, our findings confirm those of previous studies that find evidence of a positive relationship between higher proportions of high-skilled workers and higher inequality in educational attainments, both at the regional (Florida et al., 2008; Florida & Mellander, 2016; Rodriguez-Pose & Tselios, 2009) and at the local level (Glaeser et al., 2009).

Similarly, most of the economic controls are statistically significant and have the expected sign. The greater the average income, the higher the income inequality. This

result is in line with Rodriguez-Pose & Tselios' (2009) findings, but clashes with those reported in Florida & Mellander (2016) and Glaeser et al. (2009), where higher levels of income are associated with more equal income distributions. Higher housing values also exert a positive effect on income inequality, as suggested by Ganong & Shoag (2017). The share of employment in services and manufacturing (sectors associated with relatively high earnings for relatively low-skilled workers) correlates negatively with local income inequality, as in Rodriguez-Pose & Tselios (2009) or Cloutier (1997), among others. Additionally, the share of female workforce also has a positive effect, which likely reflects the lower-than-average earnings of this labour group due to either shorter working hours or wage discrimination in the labour market. Evidence supporting this hypothesis has been found in Wheeler (2008). Somewhat surprisingly, the local unemployment rate helps reducing the income gap. Job protection and unemployment benefits could explain this result, as they help equalizing income distributions (OECD, 2012a). In this regard, local initiatives aimed at promoting a city's urban and economic structure renewal have also been crucial. ¹²

As regards the demographic factors we find that, on the one hand, the income gap increases with the percentage of individuals older than age 65 in the local population whereas, on the other hand, household size is associated with a decrease in income inequality, as consumption scale economies are maximized. Total resident population also exhibits a positive impact, suggesting that greater populations facilitate job matches among less-skilled workers and improve their relative position in the income distribution (Levernier et al., 1998). This result confirms Rodriguez-Pose & Tselios' (2009) hypothesis that highly urbanized regions seem to be more prosperous and less unequal.

The fiscal policy factors also affect income inequality. As expected, public spending exerts a redistributive role on the local economies and helps narrowing income distributions (Martinez-Vazquez et al., 2012).

Considering the political factors, the empirical results show that income inequalities are lower in those municipalities with a greater vote share of the left-wing parties in the local council. This result is in line with our expectations, since the theoretical role of parties to the left of the political spectrum is to promote an active intervention in the economy and encourage the redistribution of income through public policies (see, e.g., Swank, 2002). Finally, we find evidence of the distributional consequences

of corruption. In particular, corruption is positively associated with local income inequality, as it creates permanent distortions that affect the government role in resource allocation - reducing the level of social services available to the poor - and it is also likely to accrue to the better-connected individuals in society, who belong mostly to the high-income groups (Gupta et al., 2002). The effects of both variables are robust, since they are present in all the models and always exhibit the same sign.

We now turn our attention to the spatial autocorrelation parameter (λ), which indicates the average spatial diffusion intensity of a shock in one jurisdiction to the other jurisdictions. It exhibits a positive and significant coefficient at the 5% confidence level in the 100% of the regressions with an average estimated effect of 0.27. This implies that a 1% shock to the estimated error term of inequality in one location propagates to all the other locations of the sample with an average exponential decay of the 0.27% as distance increases. These findings confirm that the SEM used in this analysis is suitable for the study of local inequality.

In addition, we have explored the sensitivity of our results in a number of different ways. We check whether the results hold when (i) we allow for heteroscedastic disturbances using a robust version of the Spatial Bayesian estimator (see Table A.3), (ii) we exclude the population outliers from the spatial data sample, that is, we drop the 50 biggest municipalities and the top 1% municipalities with extreme Gini values, respectively (see Tables A4 and A5 in the Appendix). Overall, these findings allow us to validate our main estimates, giving us confidence in the robustness of the results.

Bayesian Geographically Weighted Regression Estimates

With the aim of obtaining further policy implications suitable at the local level, a final issue considered here is the heterogeneity of the estimated parameters across space. Figure (A1) in the Appendix provides graphical evidence on the existence of hotspots in income inequality by means of the Local Moran's I and its p-values that could be explained by the existence of spatially varying relationships. As a further check, tests on individual parameter stationarity as described in Fotheringham et al. (2002) are performed. The results provide evidence of spatial non-stationarity for half of the top determinants of inequality (see Column (5) in Table (2)). Thus, relaxing the assumption of homogeneous effects on income inequality seems appropriate in this

context.

To that end, we carry out a BGWR analysis following the methodology developed and extensively discussed in LeSage (2004). The main contribution of the BGWR approach is the use of distance-weighted sub-samples of the data to produce locally linear regression estimates for every point in space. However, to carry out inferences in this context, we do not report a single BGWR model estimate, but an unweighted average of the results obtained from 1,000 model combinations of X , drawn from a Binomial distribution with an expected model size of 10.¹³ In particular, we estimate the following BGWR:

$$\tilde{y}_i = \tilde{X}_i \beta_i + \epsilon_i \quad (16)$$

$$\beta_i = (w_{i1} \otimes I_k \dots w_{in} \otimes I_k) \begin{pmatrix} \beta_i \\ \vdots \\ \beta_n \end{pmatrix} + u_i \quad (17)$$

where $\tilde{y}_i = W_i y$, $\tilde{X}_i = W_i X$ and β_i is a $k \times 1$ parameter vector associated with observation i . Here W_i is a distance-based weight vector based on an exponential decay function such that $W_i = e^{\left(\frac{-d_i}{\theta}\right)}$ where θ denotes the bandwidth. In this setting β_i is smoothed using a combination of neighbouring areas. Following LeSage (2004), to complete the model specification we add distribution for the terms: $\epsilon_i \sim N[0, \sigma^2 V_i]$, $V_i = \text{diag}(v_1, \dots, v_n)$ $u_i \sim N[0, \delta^2 \sigma^2 (X' W_i X)^{-1}]$ with δ^2 acting as the scale factor.¹⁴

The information of the BGWR estimation on the determinants of income inequality with PIPs above 25% is summarised in Table (2). Additional maps with the spatial variation of the regression estimates are provided in Figures (A2) and (A3) in the Appendix.

[insert Table (2) about here]

Columns (4) and (6) of Table (2) report the coefficient of variation and the sign certainty of the estimated parameters across the sample of municipalities. The results show that the effect of the various regressors considered here has marked heterogeneous spatial patterns across the country.

As refers to the factors that enhance inequality, we find that the positive sign is consistent across the Spanish geography for the creative class, the educational inequality, the average income and the housing values, whereas it is not for corruption and the elderly. In particular, increasing the level of corruption may have a zero or negative effect on inequality in some of the southern municipalities belonging to the regions of Andalucia and Extremadura. This finding can be explained by the strong historical roots of political clientelistic networks and its persistence in those regions. On the other hand, the elderly exert a negative effect on inequality in the north-eastern municipalities. The effect of the creative class and the average income exhibits a west-east division. Both tend to have stronger effects raising inequality in the municipalities of the east while in those located in the west the effect is smaller. Educational disparities increase income inequality more strongly in the north-east. However, the contour-map of the effect of educational inequality shows a different spatial variation pattern. In this case, increasing educational inequality appears to have relevant inequality-enhancing effects in the central area that covers Madrid and its neighbouring municipalities. As regards the effect of housing values, we find that their effect is higher in the south-eastern and south-western parts of the country.

As refers to the variables that exert a negative effect on inequality, we find that an increase in the share of left votes, the unemployment rate and the share of employment in manufacturing always decrease inequality, whereas the effect of increasing the population size, the share of service employment and public spending in some specific areas may be positive. In particular, increasing population size raises inequality in some south-eastern municipalities, whereas its effect is negative in the rest of the country. We also observe important spatial differences in the link between government spending and inequality. The effect is strongly negative in north-eastern municipalities whereas in the north-west of the country it becomes almost zero or even positive. Increasing unemployment and manufacturing employment displays a similar geographical pattern, producing stronger inequality reductions in the north-west and north-east and smaller reductions in the centre and the south. The effect of service employment is highly heterogeneous. It is positive in the north-west and in the north-east and negative in the rest of the country. Finally, the presence of left parties produces a quite homogeneous result reducing inequality, with the exception of the north-western municipalities of Galicia.

Taken together, these results clearly suggest that there is not a unique solution

or policy mix that fits equally well in all places, given that there is a high degree of spatial complexity in the determinants of inequality.

CONCLUSIONS AND POLICY IMPLICATIONS

This paper seeks to contribute to the existing empirical literature on the determinants of income inequality. To that aim, we focus on the case of Spain and draw on a novel dataset of inequality metrics for a representative sample of Spanish municipalities over the pre-crisis period 2000-2006. According to the literature, the large number of potential determinants of income inequality at the local level results in substantial uncertainty. Hence, this paper analyses the determinants of income inequality in Spain using SBMA techniques taking into account the spatial dimension of the data and the role of space in modulating the evolution of income inequality within municipalities.

Several interesting findings arise from our empirical study. First, the spatial Bayesian Model Selection implemented in the paper points to a SEM, which excludes the presence of global and local spillovers and highlights the presence of omitted variables with explicit spatial patterns. Second, the posterior inclusion probability analysis, the sign certainty estimates and the fraction of regressions with t-stats implying significance at the 5% level help to identify a group of robust determinants of inequality. In particular, we find that local inequality outcomes are mainly determined by (i) human capital, (ii) economic factors (including income per capita and sectoral composition of employment) and, to a lesser extent, (iii) the level of regional spending, and (iv) corruption.

The SBMA analysis points to the creative class as the key variable to explain differences in local income inequality, a variable that, according to Florida et al. (2008) is the engine of development. Thus, an important policy implication stemming from our study is that policy makers might face a trade-off between growth-promoting policies and inequality-reducing policies. Additional policy implications for governments aiming to tackle the income distribution issue emerge from our results. Central and regional governments can modify income distributions via progressive taxes, redistributive public policies or regulations affecting the labour market and favouring the development of specific sectors. Similarly, local policy makers have a handful of tools

(such as social services spending, economic development to housing or zoning regulations) they could use for minimizing educational inequality and improving social mobility. In the same vein, local authorities could also help improving performance and local political accountability, since policies that reduce corruption are likely to reduce income inequality as well.

The BGWR analysis provides additional information on the suitability of policies and government actions depending on their specific spatial location. In north-western locations, the inequality-enhancing effect of corruption, coupled with the effect of population aging and the impact of public spending, suggests a policy mix focused at improving institutional quality and the attraction of younger workers. In the south-west, policies aiming to attract younger populations might mitigate inequality, but they should be accompanied by actions in the housing market due to the strong observed effect of housing prices on inequality in this geographical context. On the other hand, south-eastern municipalities may benefit not only from a housing prices reduction but also from institutional quality improvements that ultimately reduce corruption. In those municipalities the strong negative link between public spending and inequality suggests that public budget expansions and intervention are likely to be effective tools for inequality reduction. Finally, in the north-eastern jurisdictions, the spectrum of inequality-reduction policies to be applied can range from educative policies aiming at reducing educational inequality to policies encouraging further development of the manufacturing sector. Despite its importance, these results are mainly exploratory, as there is not a clear theory on why some determinants should be more effective than others in different locations. Overall, these results provide interesting findings on spatial heterogeneous effects that should be addressed in future research.

Notes

¹In particular, the municipalities in the region of Navarra, the Basque Country and the Canary Islands together with the cities of Ceuta and Melilla are dropped from the analysis.

²Confidence intervals via bootstrap re-sampling methods have been calculated.

³To compute this statistic we employ a $k = 7$ nearest neighbour row-normalized spatial weight

matrix W :

$$W = \begin{cases} w_{ij}(k) = 0 & \text{if } i = j \\ w_{ij}(k) = 0 & \text{if } i \neq j, j \notin \text{nb}(i)_k \\ w_{ij}(k) = \frac{1}{k} & \text{if } i \neq j, j \in \text{nb}(i)_k \end{cases}$$

where w_{ij} terms denote the spatial weights connecting i and j , $\text{nb}(i)_k$ denotes the neighbourhood of i given k .

⁴The estimation of the stochastic kernel relies in Gaussian kernel smoothing functions and it is performed by employing the L-stage Direct Plug-In estimator with an adaptive bandwidth scaling pilot estimates of the joint distribution by $\alpha = 0.5$.

⁵The only exceptions are Ezcurra (2007) and Rodriguez-Pose & Tselios (2009), but these analysis are carried out at the regional level.

⁶A global spillover implies that a change in jurisdiction A would lead to a reaction by neighbouring jurisdictions B to change their levels of inequality, which in turn produces a game-theoretic (feedback) response of jurisdiction A, and also responses of jurisdictions C, who are neighbours to neighbouring jurisdictions B and so on, which is implausible.

⁷In particular, Table (A.2) results are obtained averaging over 1,000 subsets of X of size Ξ generated by drawing from a Binomial distribution $\Xi \sim \text{Bin}(K, \phi)$, where the parameter ϕ was used to adjust the size of the subset of regressors. Results using different values of ϕ did not change significantly the $P(SM|W, X)$ and the $P(W|SM, X)$.

⁸We use the same prior distribution configurations for the parameters δ , σ and λ employed in the model selection analysis (see footnote of Table A2). However, $p(\delta_k)$ is adjusted following the convention in BMA analysis by means of the g-prior hyper-parameter which takes the value of $g_k = 1/\max(n, K^2)$ such that:

$$p(\delta_k|\sigma^2) \sim N \left[0, \sigma^2 \left(g_k X_k' X_k \right)^{-1} \right]$$

The employment of the g-prior scales the variance of the coefficients in δ_k reflecting the strength of the prior. Lastly, we employ a binomial prior on the model space $p(M_k) = \phi^k (1 - \phi)^{K-k}$, where each covariate k is included in the model with a probability of success ϕ . We set $\phi = 1/2$ which assigns equal probability $p(M_k) = 2^{-K}$ to all the models under consideration.

⁹The number of draws to carry out the sampling exercise on the model space was 50,000.

¹⁰This is because in the context of endogenous regressors the model posterior probabilities are based on pseudo-likelihoods that are not fully comparable across models.

¹¹The key difference with respect to unconditional posterior estimates of Equations (10) and (11) is that conditional posterior estimates for a particular variable are obtained as the weighted average over the models where the variable is included. On the contrary, the unconditional posterior estimate is

the averaged coefficient over all models, including those in which the variable does not appear, hence having a zero coefficient. Thus, the unconditional posterior mean can be computed by multiplying the conditional mean in Column (3) times the PIP in Column (1)

¹²See, for instance, Viladecans & Arauzo-Carod (2012) for an example of the success of a local policy that promoted the formation of a cluster of knowledge-based activities in a poor neighbourhood of the city of Barcelona.

¹³We proceed in this way in order to match the average model size implied by the MC^3 algorithm and to minimize the bias implied by omitted variables in a single BGWR. Another reason to take this approach is that employing only the top regressors or the full set of regressors did not allowed to obtain BGWR estimates, as the correlation structure of the set of variables in X implied linear dependence problems.

¹⁴In the empirical analysis conducted we use an empirical prior for the bandwidth parameter θ based on the minimization of the score function such that $\theta = \operatorname{argmin} \sum_{i=1}^n [y_i - \hat{y}_i(\theta)]^2$, a prior for robustification against outliers $v_i \sim \chi_r^2$ with $r = 4$, and an improper prior for δ , by setting $\delta = 1$. For more details see LeSage (2004).

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the managing editor and two anonymous referees for their numerous insightful comments and constructive suggestions that have helped to significantly improve the work.

REFERENCES

- Afonso, A., Schuknecht, L., & Tanzi, V. (2010). Income Distribution Determinants and Public Spending Efficiency. *Journal of Economic Inequality*, 8, 367-389. doi:10.1007/s10888-010-9138-z
- Atkinson, A.B., Piketty, T., & Saez, E. (2011). Top Incomes in the Long Run of History. *Journal of Economic Literature*, 49(1), 3-71. doi:10.1257/jel.49.1.3
- Berube, A., & Holmes, N. (2015). *Some cities are still more unequal than others: an update*. Brookings Institution Report.

Bonnet, O., Bono, P.H., Chapelle, G., & Wasmer, E. (2014). Does Housing Capital Contribute to Inequality? A comment on Thomas Piketty's Capital in the 21st Century. *Sciences Po. Department of Economics Discussion Paper*, 2014-07. <https://core.ac.uk/download/pdf/35310497.pdf>

Cloutier, N.R. (1997). Metropolitan Income Inequality during the 1980s. The Impact of Urban Development, Industrial Mix, and Family Structure. *Journal of Regional Science*, 37(3), 459-478. doi:10.1111/0022-4146.00064

Crespo-Cuaresma, J., Doppelhofer, G., & Fedkircher (2014). The Determinants of Economic Growth in European Regions. *Regional Studies*, 48(1), 44-67. doi:10.1080/00343404.2012.6

Dabla-Norris, E., Kochar, K., Suphaphiphat, N., Ricka, F., & Tsounta, E. (2015). Causes and Consequences of Income Inequality: A Global Perspective. *IMF Staff Discussion Note*, 15/13. <https://www.imf.org/external/pubs/ft/sdn/2015/sdn1513.pdf>

Da Silva, D.F.C, Elhorst, J.P., & Da Mota Silveira Neto, R. (2017). Urban and Rural Population Growth in a Spatial Panel of Municipalities. *Regional Studies*, 51(6), 894-908. doi:10.1080/00343404.2016.1144922

Elhorst, P. (2014). *Spatial Econometrics. From Cross-sectional Data to Spatial Panels*. Springer-Verlag Berlin Heidelberg.

Ezcurra, R. (2007). Is Income Inequality Harmful for Regional Growth? Evidence from the European Union. *Urban Studies*, 44(10), 1953-1971. doi:10.1080/00420980701471877

Florida, R. (2002). *The Rise of the Creative Class*. New York, NY: Basic Books.

Florida, R., & Mellander, C. (2015). *Segregated City. The Geography of Economic Segregation in Americas Metros*. Toronto: Martin Prosperity Institute.

Florida, R., & Mellander, C. (2016). The Geography of Inequality: Difference and Determinants of Wage and Income Inequality across US Metros. *Regional Studies*, 50(1), 79-92. doi:10.1080/00343404.2014.884275

Florida, R., Mellander, C. & Stolarik, K. (2008). Inside the black box of regional development: human capital, the creative class and tolerance. *Journal of Economic Geography*, 8 (5), 615-649. doi:10.1093/jeg/lbn023

Fotheringham, S., Brunsdon, C., & Charlton, M. (2002). *Geographically Weighted Regression. The analysis of spatially varying relationships*. Chichester, UK: John Wiley & Sons.

Ganong, P., & Shoag, D. (2017). Why has regional income convergence in the U.S. declined? *Journal of Urban Economics*, 102, 76-90. doi:10.1016/j.jue.2017.07.002

Gibbons, S., & Overman, H.G. (2012). Mostly Pointless Spatial Econometrics? *Journal of Regional Science*, 52(2), 172-191. doi:10.1111/j.1467-9787.2012.00760.x

Glaeser, E., Resseger, M., & Tobbia, C. (2009). Inequality in Cities. *Journal of Regional Science*, 49(4), 617-646. doi:10.1111/j.1467-9787.2009.00627.x

Glaeser, E.L., Scheinkman, J.A., & Shleifer, A. (1995) Economic Growth in a Cross-Section of Cities. *Journal of Monetary Economics*, 36(1), 117-143. doi:10.1016/0304-3932(95)01206-2

Gonzales, C., Jain-Chandra, S., Kochhar, K., Newiak, M., & Zeinullayev, T. (2015). Catalyst for Change: Empowering Women and Tackling Income Inequality. *IMF Staff Discussion Notes*, 2015/20.

Gupta, S., Davoodi, H., & Alonso-Terme, R. (2002). Does corruption affect income inequality and poverty? *Economics of Governance*, 3, 23-45. doi:10.1007/s101010100039

Guvenen, F., Karahan, F., Ozkan, S., & Song, J. (2015). What Do Data on Millions of U.S. Workers Reveal about Life-Cycle Earnings Risk? *NBER Working Papers 20913*, National Bureau of Economic Research, Inc.

Halleck-Vega, S., & Elhorst, J.P. (2015). The SLX Model. *Journal of Regional Science*, 55(3), 339-363. doi:10.1111/jors.12188

Hortas-Rico, M., Onrubia, J. & Pacifico, D. (2014) Estimating the personal income distribution in Spanish municipalities using tax micro-data, *International Center for Public Policy Working Paper Series, paper 1419*, Andrew Young School of Policy Studies.

Izquierdo, M., & Lacuesta, A. (2007). Wage Inequality in Spain: Recent Developments. *European Central Bank Working Paper Series*, 781.

LeSage, J.P., & Parent, O. (2007). Bayesian Model Averaging for Spatial Econometric Models. *Geographical Analysis*, 39(3), 241-267. doi:10.1111/j.1538-4632.2007.00703.x

LeSage, J.P. (2004). A Family of Geographically Weighted Models. In: Anselin, L., Florax, R.J.G.M., & Rey, S.J. (eds) *Advances in Spatial Econometrics. Advances in Spatial Science*. Springer-Verlag Berlin Heidelberg.

LeSage, J.P. (2014). Spatial econometric panel data model specification: A Bayesian approach. *Spatial Statistics*, 9, 122-145. doi:10.1016/j.spasta.2014.02.002

Levernier, W., Rickman, D., & Partridge, M. (1995). Variation in U.S. State income inequality: 1960-1990. *International Regional Science Review*, 18(3), 355-378. doi:10.1177/016001769501800305

Levernier, W., Partridge, M.D., & Rickman, D.S. (1998). Differences in Metropolitan and Non-metropolitan U.S. Family Income Inequality: A Cross-County Comparison. *Journal of Urban Economics*, 44(2), 272-290. doi:10.1006/juec.1997.2070

Magrini, S. (2009). Why should we analyse convergence using the distribution Dynamics approach? *Scienze Regionali - Italian Journal of Regional Science*, 8 (1), 5-34.

Martinez-Vazquez, J., Vulovic, V., & Moreno-Dodson, B. (2012). The impact of tax and expenditure policies on income distribution: evidence from a large panel of countries. *Hacienda Pública Española / Review of Public Economics*, 200 (4), 95-130.

Mincer, J. (1958). Investment in human capital and personal income distribution. *Journal of Political Economy*, 66(4), 281-302. doi:10.1086/258055

Moral-Benito, E. (2015). Model averaging in economics: An overview. *Journal of Economic Surveys*, 29(1), 46-75. doi:10.1111/joes.12044

Musterd, S., Marcinczak, S., Van Ham, M., & Tammaru, T. (2017). Socioeconomic segregation in European capital cities. Increasing separation between poor and rich. *Urban Geography*, 38(7), 1062-1083. doi:10.1080/02723638.2016.1228371

OECD (2012a). Inequality in labour income. What are its drivers and how can it be reduced? *OECD Economics Department Policy Notes*, 8, 1-8.

OECD (2012b). *Economic Policy Reforms 2012: Going for Growth*. Paris: OECD Publishing. <http://dx.doi.org/10.1787/growth-2012-en>

OECD (2017). *Understanding the socio-economic divide in Europe*. Paris: OECD Publishing.

Rios, V. (2017). What drives unemployment disparities in European regions? A dynamic spatial panel approach. *Regional Studies*, 51(11), 1599-1611. doi:10.1080/00343404.2016.121

Rodriguez-Pose, A., & Tselios, V. (2009). Education and income inequality in the regions of the European Union. *Journal of Regional Science*, 49(3), 411-437. doi:10.1111/j.1467-9787.2008.00602.x

Rognlie, M. (2015). Deciphering the fall and rise in the net capital share: accumulation or scarcity? *Brookings Papers on Economic Activity* 2015 (1), 1-69. doi: 10.1353/eca.2016.0002

Swank, D. (2002). *Global capital, political institutions and political change in developed welfare states*. New York, NY: Cambridge University Press.

The Economist (2012, October 13). For richer, for poorer. Special report: The World Economy. *The Economist*, 2-4.

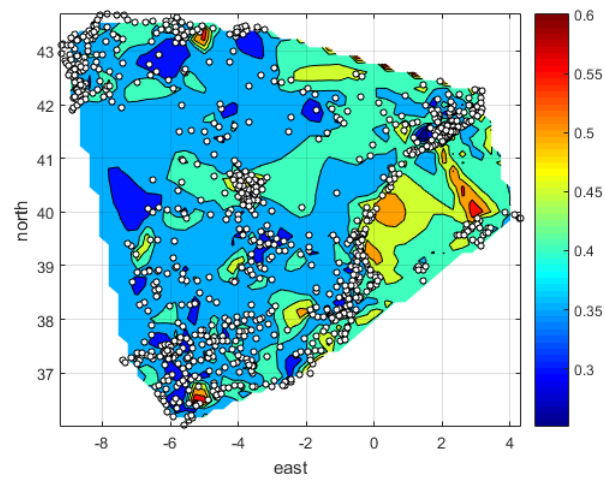
Tiebout, C.M. (1956). A Pure Theory of Local Expenditures. *The Journal of Political Economy*, 64(5), 416-424. doi:10.1086/257839

Topel, R. (1994). Regional Labor Markets and the Determinants of Wage Inequality. *American Economic Review, Papers and Proceedings*, 84(2), 17-22.

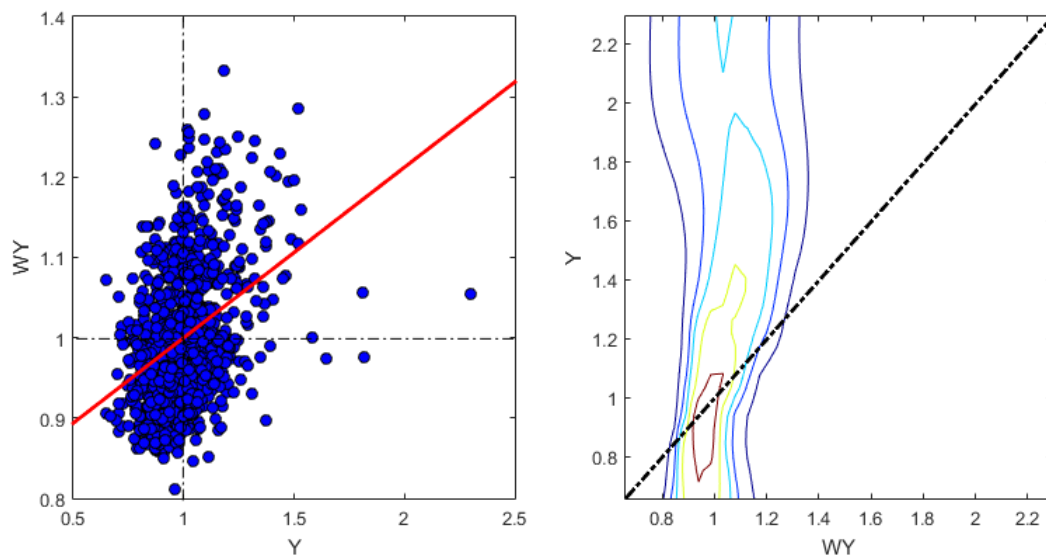
Viladecans, E., & Arauzo-Carod, J.M. (2012). Can a knowledge-based cluster be created? The case of the Barcelona 22@ district. *Papers in Regional Science*, 91(2), 377-400. doi:10.1111/j.1435-5957.2011.00383.x

Wheeler, C.H. (2008). Urban decentralization and income inequality: Is sprawl associated with rising income segregation across neighborhoods? *Regional Economic Development, Federal Reserve Bank of St. Louis*, 4(1), 41-57.

Figure 1: Spatial Patterns of Income Inequality in Spain



(a) Gini Index



(b) The role of geographical location

Table 1: Main Results: Bayesian Spatial Error Model Averaged Estimates.

Variable	PIP	Lower 5%	Cond Post. Mean	Cond Post. Std	Upper 95%	T-Stat > 1.96	Sign Pos.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Creative class	0.9999	0.2811	0.3599	0.0613	0.4623	1.00	1.00
Services	0.9999	-0.2406	-0.1971	0.0390	-0.1461	1.00	0.00
Manufacturing	0.9986	-0.1129	-0.0890	0.0211	-0.0719	1.00	0.00
Housing values	0.9850	0.0063	0.0107	0.0050	0.0188	1.00	1.00
Average Income	0.9482	0.0396	0.0583	0.0166	0.0799	1.00	1.00
Unemployment rate	0.8478	-0.1550	-0.1126	0.0349	-0.0776	1.00	0.00
Public spending	0.7481	-0.0741	-0.0612	0.0209	-0.0450	1.00	0.00
Educational Inequality	0.5610	0.0874	0.1237	0.0515	0.1540	0.98	1.00
Corruption	0.5517	0.0001	0.0002	0.0001	0.0003	0.90	1.00
Population +65	0.5153	0.0006	0.0015	0.0007	0.0022	0.86	1.00
Left-wing government	0.3795	-0.0375	-0.0285	0.0126	-0.0192	0.91	0.00
Population size	0.3384	-0.0163	-0.0104	0.0052	0.0051	0.88	0.10
Female participation	0.2091	0.0226	0.0629	0.0299	0.0958	0.68	1.00
Distance to Coast	0.0937	0.0000	0.0000	0.0000	0.0000	0.43	0.00
Household size	0.0802	-0.0577	-0.0390	0.0128	-0.0105	0.42	0.02
Low-skilled workers	0.0640	-0.0820	-0.0527	0.0165	-0.0201	0.27	0.03
Regional Alignment	0.0331	0.0033	0.0041	0.0010	0.0053	0.00	1.00
Taxation	0.0251	-0.0052	0.0041	0.0015	0.0151	0.19	0.74
Bars and restaurants	0.0239	-0.0032	0.0041	0.0011	0.0119	0.18	0.82
Construction	0.0235	-0.0628	-0.0353	0.0079	-0.0044	0.01	0.03
Crime rate	0.0227	-0.0587	0.1686	0.0339	0.4206	0.28	0.87
Agriculture	0.0225	-0.0365	-0.0014	0.0053	0.0579	0.11	0.40
National Alignment	0.0216	-0.0030	-0.0001	0.0007	0.0046	0.00	0.29
Population density	0.0195	-0.0020	-0.0004	0.0003	0.0017	0.00	0.35
Urban blight	0.0192	-0.0306	-0.0136	0.0036	0.0032	0.00	0.09
Homeownership rate	0.0190	0.0000	0.0001	0.0000	0.0003	0.00	0.88
Government strength	0.0188	-0.9443	-0.5315	0.2144	-0.1397	0.00	0.01
Transfers	0.0183	-0.0210	-0.0082	0.0025	-0.0028	0.00	0.02
Road accessibility	0.0182	0.0000	0.0000	0.0000	0.0000	0.02	0.96
Immigration	0.0181	0.0000	0.0003	0.0001	0.0007	0.02	0.91
Urban Sprawl	0.0178	-0.0115	-0.0010	0.0016	0.0048	0.00	0.51
Urban area	0.0177	0.0006	0.0014	0.0005	0.0024	0.00	1.00
Relative income	0.0174	-0.0015	0.0009	0.0005	0.0036	0.00	0.68
Spatial Autocorr (λ)	1.0000	0.1930	0.2772	0.0508	0.3596	1.00	1.00

Notes: The dependent variable in all regressions is the Gini index of income in 2006. All the results reported here correspond to the estimation of the top 10,000 models from the 86,899 million possible regressions including any combination of the 33 variables. Prior mean model size is 16.5. Variables are ranked by Column (1), the posterior inclusion probability. Columns (2) to (5) reflect the lower 5% bound, the posterior mean, standard deviations and upper 95% bound for the linear marginal effect of the variable conditional on inclusion in the model, respectively. Column (6) is the fraction of regressions in which the coefficient has a classical t-test greater than 1.96, with all regressions having equal sampling probability. The last column denotes the sign certainty probability, a measure of our posterior confidence in the sign of the coefficient.

Table 2: Model Averaged BGWR Estimates

	Minimum β_i (1)	Median β_i (2)	Maximum β_i (3)	Coefficient of variation (4)	Spatial Stationarity (5)	Sign Positivity (6)
Creative class	0.2456	0.3317	0.4949	0.2164	8.88 [0.00]	1.000
Services	-0.0388	-0.0141	0.13	5.8333	2.97 [0.04]	0.415
Manufacturing	-0.1963	-0.0603	-0.0213	0.6812	4.73 [0.00]	0.000
Housing values	0.0088	0.0118	0.0158	0.1675	1.21 [0.30]	1.000
Average income	0.0319	0.0781	0.1394	0.4145	4.25 [0.00]	1.000
Unemployment rate	-0.2758	-0.1252	-0.0606	0.3579	1.05 [0.35]	0.000
Public Spending	-0.0732	-0.0394	0.0014	0.4201	1.67 [0.17]	0.035
Educational Inequality	0.118	0.2151	0.3461	0.2572	3.47 [0.02]	1.000
Corruption	0	0.0002	0.0005	0.6649	5.46 [0.00]	0.989
Population +65	-0.0004	0.0009	0.0012	0.62	0.94 [0.41]	0.836
Left-wing government	-0.068	-0.0527	-0.0276	0.1884	0.76 [0.50]	0.000
Population size	-0.0058	-0.001	0.0055	3.547	1.18 [0.31]	0.396

Note: The sign positivity is calculated as the fraction of municipalities for which the estimated parameter is positive. The parameter stationarity test is calculated following Fotheringham et al (2002). The null H_0 corresponds to spatial stationarity. F-test values and p-values (in brackets) are obtained computing the median of the distribution over 3,000 model draws, where the X's forming the models are drawn from a Binomial distribution with an expected model size given by $\Xi \sim \text{Bin}(k, \phi)$ with $k = 33$ and $\phi = 0.3$.

APPENDIX

Table A1 : Definitions, sources and descriptive statistics of the explanatory variables.

Variable	Definition	Source	Mean	Std	Expected Effect
Outcome variable					
Income inequality	Gini index of income 2006	TAO	0.39	0.06	
<i>(i) Human capital</i>					
Creative class ⁽¹⁾ .	Share of professional and knowledge-work occupations	INE	0.17	0.07	+
Low-skilled workers	Share of workforce employed in low-skilled job	INE	0.14	0.08	?
Educational inequality ⁽²⁾	Gini index of educational attainment	INE	0.41	0.06	+
<i>(ii) Economic factors</i>					
Average income	Per capita income 2002 (logs)	TAO	9.51	0.26	?
Unemployment rate	Share of the 16 years or older population that is unemployed	INE	0.14	0.07	?
Manufacturing	Share of workforce employed in manufacturing	INE	0.20	0.12	-
Construction	Share of workforce employed in construction	INE	0.14	0.06	-
Services	Share of workforce employed in services	INE	0.33	0.11	-
Agriculture	Share of workforce employed in agriculture	INE	0.09	0.11	+
Female participation	Female share of the labor force	INE	0.42	0.08	?
Housing values	Average housing value within each municipality (logs)	PAO	12.30	1.19	+
Homeownership rate	Share of houses occupied by owner	INE	0.83	0.06	?
Relative income	Share of municipal GDP over total GDP (%)	INE	0.10	0.51	?
<i>(iii) Demographic factors</i>					
Urban area	Dummy variable, 1 if the municipality belongs to an urban area, 0 otherwise	Boix et al (2012)	0.59	0.49	+
Immigration	Share of immigrants in the resident population	INE	0.04	0.05	+
Population +65	Share of total resident population that is 65 years or older	INE	0.16	0.04	+
Household size	Average number of individuals per household (logs)	INE	1.08	0.09	-
Population size	Total resident population in the municipality (logs)	INE	9.55	0.942	?
Population density	Number of residents per km^2	INE	5.38	1.47	?

Table A1 : Definitions, sources and descriptive statistics of the variables (Continued)

Variable	Definition	Source	Mean	Std	Expected Effect
<i>(iv) Fiscal policy</i>					
Taxation	Per capita regional government revenues from the personal income tax (logs)	SMF	5.54	0.47	-
Public Spending	Per capita regional government total expenditure (logs)	SMF	7.81	0.14	-
Transfers	Per capita total local revenues from transfers (logs)	SMF	5.46	0.33	-
<i>(v) Politics</i>					
Left-wing governments ⁽³⁾	Vote share of the left-wing parties in the municipal council.	MHA	0.54	0.18	-
Government strength	Share of seats held by the major's party in the municipal council	MHA	0.52	0.11	?
Regional alignment ⁽⁴⁾	Dummy variable, 1 if regional and local governments are aligned, 0 otherwise	MHA	0.27	0.44	?
National alignment ⁽⁴⁾	Dummy variable, 1 if national and local governments are aligned, 0 otherwise	MHA	0.43	0.49	?
Corruption ⁽⁵⁾	Combined max-min normalized local-regional corruption index	Press reports	0.55	0.43	+
<i>(vi) Local amenities</i>					
Urban blight ⁽⁶⁾	Share of total houses with problems	INE	0.23	0.09	?
Crime rate	Number of crimes and misdemeanors per 1,000 inhabitants (logs)	INE	0.04	0.02	?
Bars and restaurants	Number of bars and restaurants per 1,000 inhabitants (logs)	INE	1.77	0.40	?
Road accessibility	Kilometers of road and rail network (logs)	NGI	8.03?	1.22?	?
Urban sprawl ⁽⁸⁾	% of undeveloped land around residential land within the immediate neighborhood	CLC and GIS	0.63	0.17	?
Distance to coast	Distance from the municipality centroid to the coast	NGI and GIS	72.35	92.79	?

Notes: Unless specified all control variables are taken at year 2001 and at the level of municipality, except for *taxation* and *total spending* which are measured at the regional level. The income variable used is pretax gross income, and it is defined as the sum of salary and wage income, retirement pensions, general unemployment subsidies, some non-exempt welfare payments and some disability pensions, net self-employment income, interest, dividends, royalty income, survivor annuities, net rental and income from other estates including imputed rent for second dwellings homeowners, and realized capital gains (except those from reinvesting in the customary dwelling). Distance-based variables have been calculated using Geographical Information Systems (GIS). INE denotes the National Statistics Institute. Most of INE data is comes from the Census of Population and Housing. PAO denotes Property Assessment Office, TAO denotes Tax Administration Office, SMF is the Spanish Ministry of Finance, and MHA is the Ministry of Home Affairs. NGI denotes National Geographic Institute and CLC denotes Corine Land Cover data. (1) Includes management occupations, business and financial operations, scientists and intellectuals. (2) Own calculations as in Thomas et al (2001), using the proportion of resident population without studies, with primary, secondary and tertiary studies. In particular, we have used an indirect method based on the Lorenz curve, with the cumulative percentage of the schooling years on the vertical axis and the cumulative percentage of the population on the horizontal axis. To that aim, we use data on the proportion of population with various levels of educational attainment and the years of schooling for each level, according to the structure of the educational system in Spain. We have divided the population of each municipality into seven categories: proportion of population without studies or illiterate, proportion of population with primary education, proportion of population with secondary education (compulsory, until 16 years old), proportion of population with tertiary education (non compulsory, from 16 to 18 years old), proportion of population with tertiary education (3-years college degree), proportion of population with tertiary education (5-years college degree), proportion of population with tertiary education (PhD). The seven groups are both mutually exclusive and collectively inclusive for the concerned population.(3) Parties on the left are: PSOE, PCE, IC, and several left regionalist parties. (4) Regional or national and local governments are aligned if they are controlled by the same party. (5) The index is constructed as: $CI = \left(\frac{C_i - C_{min}}{C_{max} - C_{min}} \right)$ where $C_i = 0.5CC_i + 0.5CR_i$ with CC_i a dummy variable indicating whether there have been corruption scandals in the municipality and CR a continuous variable with the number of corruption scandals in the region. (6) The type of problems considered are noise, dirty, pollution or lack of green space. (7) See Gomez-Antonio et al. (2016) for further details on the definition of this variable.

Table A2: Spatial Bayesian Model Selection

Spatial Weight Matrix	Posterior Model Probabilities For Spatial Models $P(SM W, X)$				Posterior Model Probabilities For W Matrices $P(W SM, X)$			
	SLM (1)	SDM (2)	SEM (3)	SDM (4)	SLM (5)	SDM (6)	SEM (7)	SDM (8)
5-nearest neighbours	0.001	0.000	0.745	0.254	0.0198	0.0059	0.0038	0.0046
6-nearest neighbours	0.000	0.000	0.810	0.190	0.4505	0.1434	0.1844	0.1212
7-nearest neighbours	0.000	0.000	0.856	0.143	0.3177	0.1020	0.2520	0.0815
8-nearest neighbours	0.001	0.000	0.876	0.123	0.0745	0.0559	0.0840	0.0483
9-nearest neighbours	0.000	0.000	0.903	0.097	0.0260	0.0307	0.0510	0.0237
10-nearest neighbours	0.000	0.000	0.904	0.095	0.0347	0.0269	0.0765	0.0200
11-nearest neighbours	0.001	0.000	0.928	0.071	0.0445	0.0349	0.1768	0.0408
12-nearest neighbours	0.002	0.000	0.969	0.029	0.0023	0.0260	0.0453	0.0288
13-nearest neighbours	0.004	0.000	0.991	0.005	0.0004	0.0117	0.0399	0.0167
14-nearest neighbours	0.006	0.000	0.991	0.003	0.0001	0.0070	0.0100	0.0123
15-nearest neighbours	0.002	0.000	0.884	0.114	0.0043	0.0581	0.0073	0.0554
20-nearest neighbours	0.005	0.001	0.916	0.078	0.0034	0.1530	0.0184	0.1807
25-nearest neighbours	0.003	0.000	0.916	0.081	0.0037	0.1416	0.0215	0.1405
30-nearest neighbours	0.003	0.000	0.911	0.086	0.0054	0.2030	0.0266	0.2184
Contiguity	0.003	0.000	0.697	0.300	0.0127	0.0000	0.0024	0.0070

Notes: Columns (1) to (4) show the probability of each spatial model SM conditional on W and X ($P(SM|W, X)$), while Columns (5) to (8) show probability of the W conditional on the spatial model SM and X ($P(W|SM, X)$). The results are obtained averaging over 1,000 subsets of X of size $\Xi \sim \text{Bin}(k, \phi)$ with $\phi = 0.5$. For each randomly generated subset regressors X_i the Bayesian estimation of the PMP is based on 1,000 draws with a burn-in sample of 100 draws. To derive individual PMPs we employ a normal-gamma conjugate prior for $\delta = [\alpha, \beta]$ and σ and a beta prior for λ :

$$p(\delta) \sim N(c, \Sigma)$$

$$p\left(\frac{1}{\sigma^2}\right) \sim \Gamma(d, v)$$

$$p(\lambda) \sim B(a_0, a_0)$$

To avoid situations where the conclusions depend heavily on subjective prior information we rely on diffuse or non-informative prior distributions. Parameter c is set to zero and Σ to a very large number ($1e+12$) which results in a diffuse prior for δ . The diffuse priors for σ and λ (in the case of the SLM/SDM ρ), are obtained setting $d = 0$ and $v = 0$ and $a_0 = 1.01$.

Table A3 : Bayesian Heteroscedastic Spatial Error Model Averaged Estimates

Variable	PIP	Lower 5%	Cond Post. Mean	Cond Post. Std	Upper 95%	T-Stat > 1.96	Sign Pos.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Creative class	0.9999	0.2800	0.3666	0.0625	0.4749	1.00	1.00
Services	0.9999	-0.2590	-0.2140	0.0379	-0.1582	1.00	0.00
Manufacturing	0.9986	-0.1140	-0.0935	0.0203	-0.0753	1.00	0.00
Housing values	0.9845	0.0069	0.0111	0.0049	0.0194	1.00	1.00
Average Income	0.9497	0.0430	0.0637	0.0176	0.0884	1.00	1.00
Unemployment rate	0.8489	-0.1628	-0.1175	0.0334	-0.0813	1.00	0.00
Public spending	0.7468	-0.0628	-0.0489	0.0188	-0.0307	0.95	0.00
Educational inequality	0.5631	0.0965	0.1342	0.0530	0.1695	1.00	1.00
Corruption	0.5617	0.0001	0.0002	0.0001	0.0002	0.94	1.00
Population +65	0.5169	0.0005	0.0014	0.0007	0.0021	0.84	0.99
Left-wing government	0.3933	-0.0374	-0.0290	0.0126	-0.0194	0.95	0.00
Population size	0.3333	-0.0166	-0.0102	0.0051	0.0057	0.94	0.11
Female participation	0.2008	0.0298	0.0679	0.0299	0.1030	0.74	1.00
Distance to Coast	0.0908	0.0000	0.0000	0.0000	0.0000	0.68	0.00
Household size	0.0806	-0.0616	-0.0434	0.0131	-0.0172	0.64	0.00
Low-skilled workers	0.0668	-0.0758	-0.0416	0.0143	-0.0042	0.24	0.04
Regional Alignment	0.0352	0.0013	0.0021	0.0008	0.0031	0.00	1.00
Taxation	0.0271	-0.0080	0.0035	0.0015	0.0138	0.20	0.73
Bars and restaurants	0.0262	-0.0029	0.0039	0.0010	0.0120	0.19	0.80
Crime rate	0.0245	-0.0818	0.1173	0.0293	0.3555	0.20	0.82
National Alignment	0.0236	-0.0046	-0.0017	0.0007	0.0028	0.00	0.29
Construction	0.0227	-0.0754	-0.0480	0.0091	-0.0234	0.05	0.02
Agriculture	0.0217	-0.0195	0.0128	0.0046	0.0648	0.08	0.70
Population density	0.0200	-0.0017	-0.0003	0.0003	0.0019	0.00	0.34
Urban blight	0.0193	-0.0245	-0.0056	0.0031	0.0152	0.00	0.29
Urban area	0.0191	-0.0003	0.0007	0.0004	0.0019	0.00	0.85
Road Accessibility	0.0190	0.0000	0.0000	0.0000	0.0000	0.09	1.00
Homeownership rate	0.0189	0.0000	0.0001	0.0000	0.0003	0.00	0.89
Immigration	0.0189	0.0000	0.0004	0.0001	0.0007	0.05	0.95
Government strength	0.0185	-0.7772	-0.3591	0.1985	0.0542	0.00	0.07
Transfers	0.0185	-0.0217	-0.0083	0.0024	-0.0024	0.00	0.01
Relative income	0.0178	-0.0012	0.0013	0.0005	0.0049	0.02	0.73
Urban Sprawl	0.0172	-0.0150	-0.0038	0.0015	0.0024	0.00	0.24
Spatial Autocorr (λ)	1.0000	0.1442	0.2274	0.0499	0.3080	1.00	1.00

Notes: The results reported here correspond to the estimation of the heteroscedastic SEM for the top 10,000 models out of the 17,202 generated by the MC^3 sampler. To obtain Bayesian heteroscedastic SEM model averaged estimates we introduce a set of variance scalars (v_1, \dots, v_2) in the estimations of Equation (3) and we now assume $v \sim N[0, \sigma^2 V]$, where $V = \text{diag}(v_1, \dots, v_2)$. Therefore, the computation of the key statistics in Equations (9), (10) and (11) are based on the following variant of the SEM model:

$$\begin{aligned}
y &= \alpha \iota_n + X\beta + \epsilon \\
\epsilon &= \lambda W\epsilon + v \\
v &\sim N[0, \sigma^2 V] \\
V &= \text{diag}(v_1, \dots, v_2) \\
\pi(\delta) &\sim N[c, \Sigma] \\
\pi(1/\sigma) &\sim \Gamma(v_0, d_0) \\
\pi\left(\frac{r}{v_i}\right) &\sim IID\chi^2(r)
\end{aligned}$$

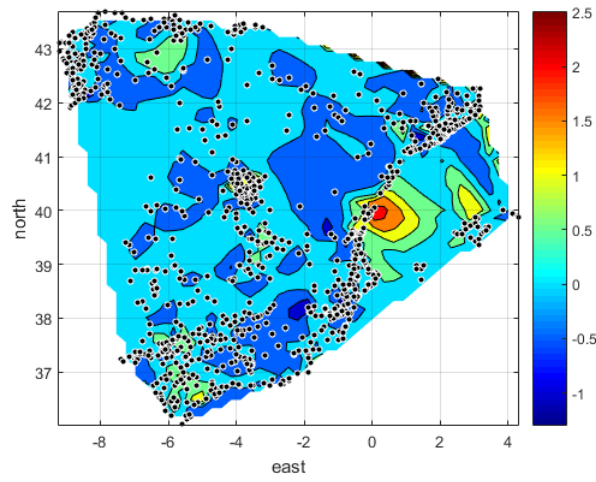
Table A4 : Top 50 Biggest Cities Excluded

Variable	PIP	Lower 5%	Cond Post. Mean	Cond Post. Std	Upper 95%	T-Stat > 1.96	Sign Pos.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Creative class	0.9997	0.2577	0.3457	0.0655	0.4522	1.00	1.00
Services	0.9982	-0.2406	-0.1935	0.0396	-0.1366	1.00	0.00
Manufacturing	0.9949	-0.1112	-0.0904	0.0204	-0.0714	1.00	0.00
Housing values	0.9925	0.0084	0.0122	0.0038	0.0204	1.00	1.00
Average Income	0.9347	0.0405	0.0616	0.0185	0.0859	1.00	1.00
Public spending	0.8277	-0.0681	-0.0535	0.0189	-0.0365	0.97	0.00
Unemployment rate	0.7684	-0.1593	-0.1154	0.0362	-0.0795	1.00	0.00
Educational inequality	0.6153	0.1023	0.1382	0.0522	0.1684	1.00	1.00
Population +65	0.4875	0.0004	0.0014	0.0007	0.0021	0.85	1.00
Corruption	0.4549	0.0001	0.0002	0.0001	0.0002	0.91	1.00
Left-wing government	0.3572	-0.0383	-0.0295	0.0128	-0.0195	0.94	0.00
Female participation	0.1530	0.0267	0.0648	0.0264	0.0987	0.72	1.00
Population size	0.1517	-0.0184	-0.0108	0.0041	0.0065	0.81	0.08
Relative income	0.1198	-0.1294	-0.0856	0.0341	-0.0083	0.42	0.05
Low-Skilled workers	0.0595	-0.0790	-0.0375	0.0134	0.0215	0.21	0.07
Distance to Coast	0.0586	0.0000	0.0000	0.0000	0.0000	0.51	0.00
Household size	0.0570	-0.0557	-0.0375	0.0104	-0.0131	0.43	0.01
Population density	0.0401	0.0000	0.0000	0.0000	0.0000	0.03	0.00
Regional Alignment	0.0395	0.0019	0.0028	0.0009	0.0039	0.00	1.00
Taxation	0.0338	-0.0020	0.0077	0.0019	0.0165	0.35	0.91
Construction	0.0248	-0.0789	-0.0493	0.0098	0.0157	0.08	0.06
Crime rate	0.0235	-0.0590	0.1402	0.0322	0.3864	0.19	0.85
Agriculture	0.0230	-0.0078	0.0437	0.0063	0.0971	0.41	0.91
Urban blight	0.0219	-0.0260	-0.0075	0.0034	0.0089	0.00	0.22
Homeownership rate	0.0195	-0.0001	0.0001	0.0000	0.0003	0.00	0.88
National Alignment	0.0194	-0.0048	-0.0021	0.0006	0.0023	0.00	0.24
Bars and restaurants	0.0192	-0.0052	0.0011	0.0008	0.0104	0.10	0.54
Transfers	0.0182	-0.0250	-0.0103	0.0026	-0.0046	0.00	0.02
Immigration	0.0180	0.0000	0.0003	0.0001	0.0008	0.07	0.94
Road accessibility	0.0175	0.0000	0.0000	0.0000	0.0000	0.00	0.98
Government strength	0.0171	-0.5682	-0.1040	0.1937	0.3314	0.00	0.37
Urban Sprawl	0.0163	-0.0156	-0.0062	0.0016	-0.0003	0.00	0.05
Urban area	0.0150	-0.0014	-0.0006	0.0004	0.0004	0.00	0.12
Spatial Autocorr (λ)	1.0000	0.1930	0.2786	0.0516	0.3636	1.00	1.00

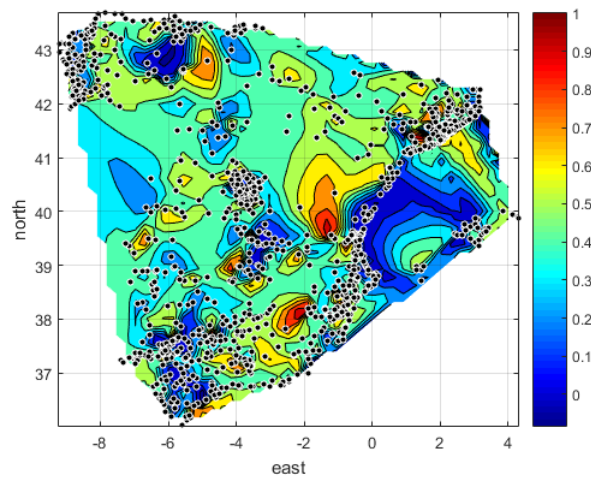
Table A5 : Top 1% Inequality Outliers Excluded

Variable	PIP	Lower 5%	Cond Post. Mean	Cond Post. Std	Upper 95%	T-Stat > 1.96	Sign Pos.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Creative Class	0.9996	0.2579	0.3136	0.0513	0.3890	1.00	1.00
Services	0.9996	-0.2483	-0.2133	0.0337	-0.1674	1.00	0.00
Manufacturing	0.9996	-0.1121	-0.0907	0.0188	-0.0720	1.00	0.00
Housing Values	0.9991	0.0071	0.0122	0.0050	0.0192	1.00	1.00
Unemployment rate	0.9837	-0.1430	-0.1088	0.0282	-0.0781	1.00	0.00
Educational inequality	0.9741	0.1010	0.1316	0.0376	0.1637	1.00	1.00
Average Income	0.9482	0.0375	0.0543	0.0154	0.0720	1.00	1.00
Left-wing government	0.8294	-0.0360	-0.0291	0.0100	-0.0215	1.00	0.00
Corruption	0.8109	0.0001	0.0002	0.0001	0.0002	0.97	1.00
Population +65	0.7196	0.0007	0.0015	0.0006	0.0022	0.89	1.00
Female participation	0.6680	0.0393	0.0760	0.0316	0.1056	0.86	1.00
Population size	0.5556	-0.0145	-0.0109	0.0046	-0.0075	0.95	0.02
Public spending	0.3740	-0.0549	-0.0403	0.0175	-0.0230	0.87	0.00
Distance to Coast	0.2054	0.0000	0.0000	0.0000	0.0000	0.81	0.00
Household size	0.1991	-0.0620	-0.0457	0.0199	-0.0164	0.72	0.00
Low-Skilled workers	0.0888	-0.0752	-0.0445	0.0164	-0.0131	0.22	0.01
Agriculture	0.0474	-0.0295	0.0031	0.0067	0.0377	0.02	0.51
Transfers	0.0430	-0.0156	-0.0106	0.0038	-0.0065	0.00	0.00
Taxation	0.0365	-0.0091	0.0026	0.0016	0.0137	0.21	0.68
Homeownership rate	0.0344	0.0000	0.0002	0.0001	0.0003	0.00	0.98
National Alignment	0.0281	-0.0051	-0.0014	0.0007	0.0018	0.01	0.45
Regional Alignment	0.0279	0.0011	0.0020	0.0006	0.0027	0.00	1.00
Construction	0.0259	-0.0629	-0.0378	0.0079	-0.0105	0.02	0.00
Bars and restaurants	0.0256	-0.0025	0.0024	0.0009	0.0068	0.03	0.78
Urban area	0.0243	-0.0014	-0.0004	0.0005	0.0008	0.00	0.28
Road accessibility	0.0240	0.0000	0.0000	0.0000	0.0000	0.01	1.00
Urban blight	0.0239	-0.0111	0.0032	0.0032	0.0170	0.00	0.63
Urban Sprawl	0.0235	-0.0121	-0.0053	0.0017	-0.0004	0.00	0.04
Crime rate	0.0230	-0.0559	0.0893	0.0239	0.2967	0.16	0.80
Immigration	0.0230	-0.0001	0.0002	0.0001	0.0005	0.01	0.83
Government Strength	0.0221	-0.7344	-0.3686	0.2031	-0.0460	0.00	0.03
Relative income	0.0218	-0.0011	0.0009	0.0005	0.0031	0.00	0.72
Population density	0.0216	-0.0015	-0.0002	0.0002	0.0012	0.00	0.37
Spatial Autocorr (λ)	1.0000	0.1367	0.2251	0.0534	0.3129	1.00	1.00

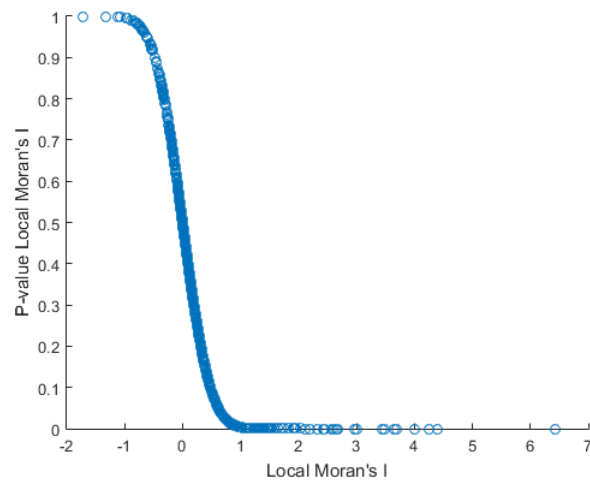
Figure A1: Spatial Heterogeneity Diagnostics



Contour map of the Local Moran's I

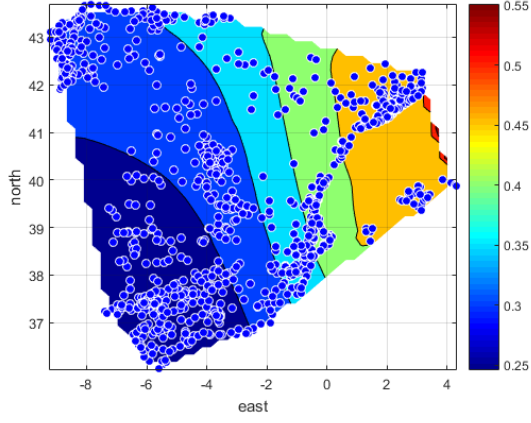


Contour map of the P-values of the Local Moran's I

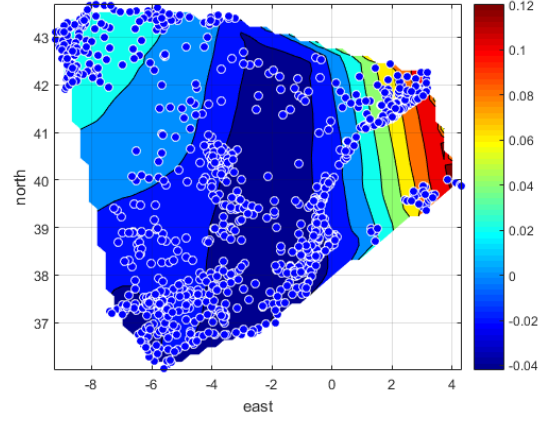


P-values and Local Moran's I

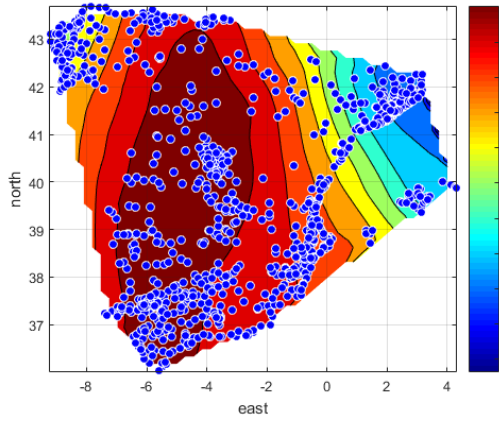
Figure A2 : Heterogeneous Effects on Income Inequality



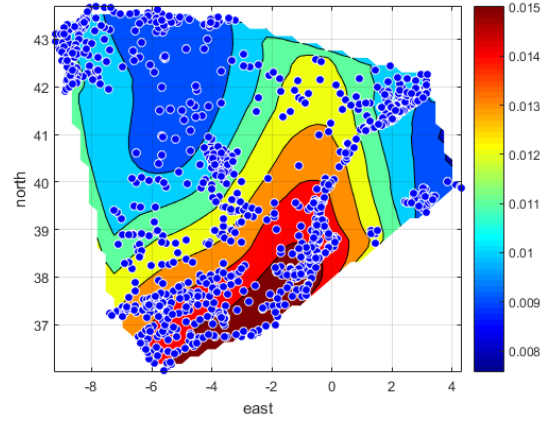
(a) Creative Class



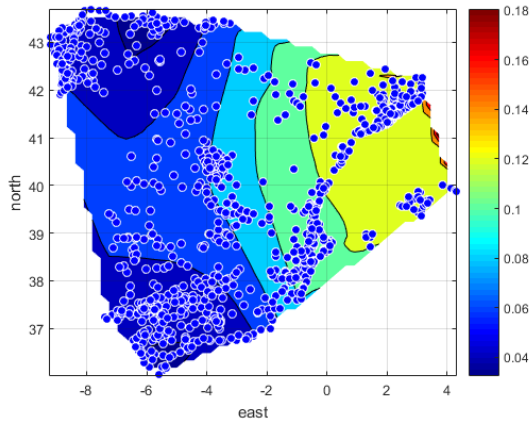
(b) Employment in services



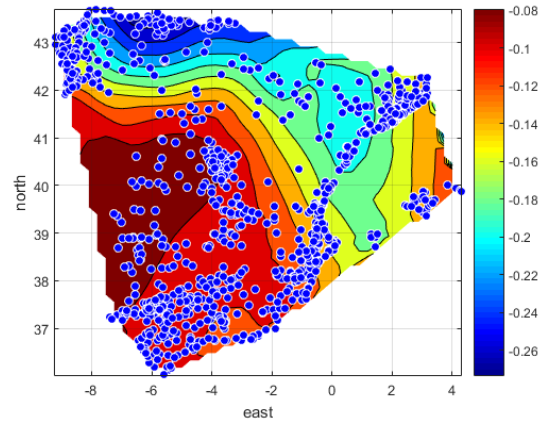
(c) Employment in manufacturing



(d) Housing values

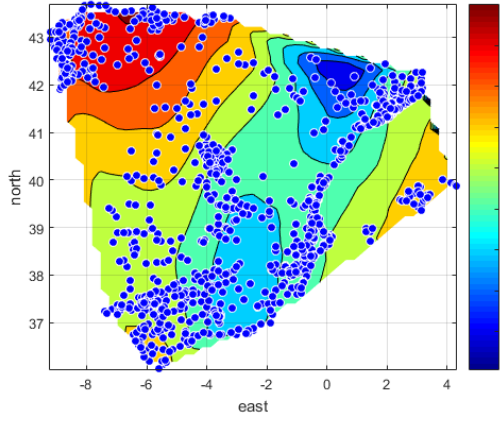


(e) Average Income

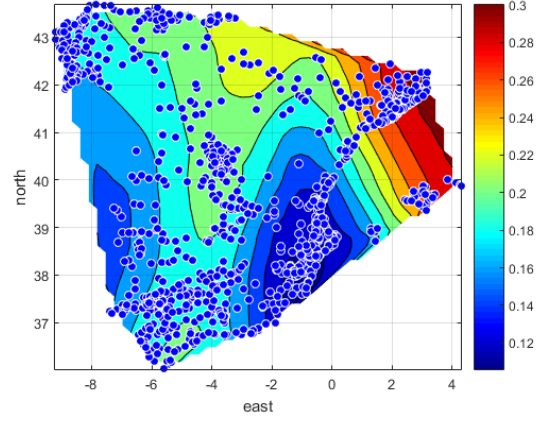


(f) Unemployment rate

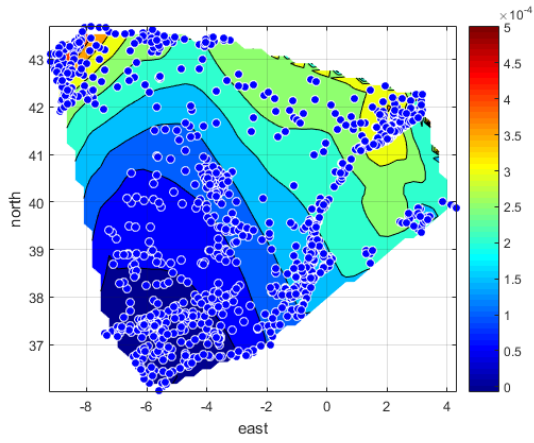
Figure A3 : Heterogeneous Effects on Income Inequality



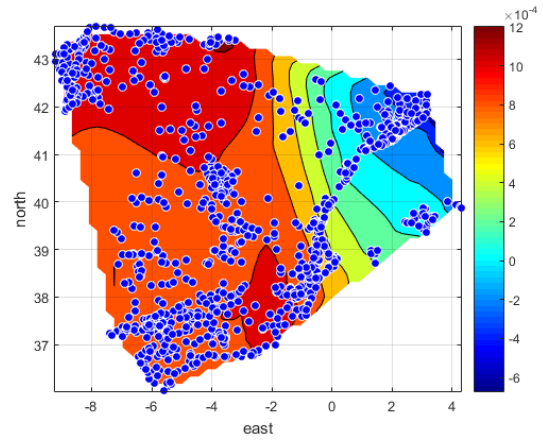
(a) Public spending



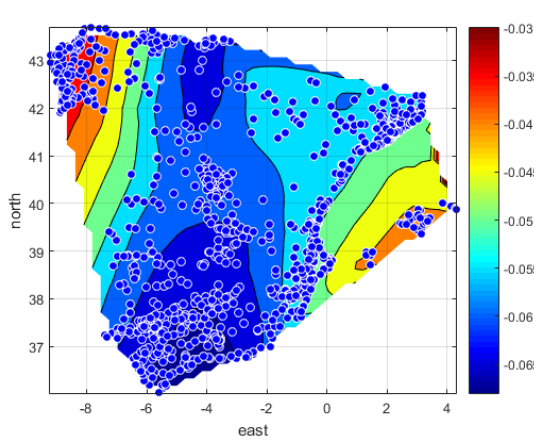
(b) Educational Inequality



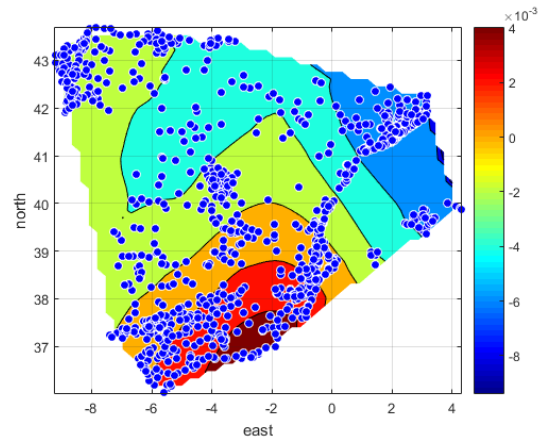
(c) Corruption



(d) Population +65



(e) Left-wing Government



(f) Population size

REFERENCES

Gomez-Antonio, M., Hortas-Rico, M., & Li, L. (2016). The causes of urban sprawl in Spanish urban areas: a spatial approach. *Spatial Economic Analysis*, 11(2), 219-247. doi:10.1080/17421772.2016.1126674

Thomas, V., Wang, Y., & Fan, X. (2001). Measuring Education Inequality; Gini coefficients of education. *The World Bank Policy Research Working Paper 2525*, 1-37. <http://documents.worldbank.org/curated/en/361761468761690314/Measuring-education-inequality-Gini-coefficients-of-education>