



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 54.7 (2020): 958 - 973

**DOI:** <https://doi.org/10.1080/00343404.2019.1648786>

**Copyright:** © 2020 Taylor and Francis

El acceso a la versión del editor puede requerir la suscripción del recurso

Access to the published version may require subscription

# Is there an Optimal Size for Local Governments? A Spatial Panel Data Model Approach

Miriam Hortas-Rico <sup>\*</sup>and Vicente Rios <sup>†</sup>

July 21, 2019

## Abstract

The paper presents a framework for determining the optimal size of local jurisdictions and whether it varies depending on the geographical heterogeneity of the territory. To that aim, we first develop a theoretical model of cost efficiency that takes into account spatial interactions and spillover effects among neighbouring jurisdictions. The model solution leads to a Spatial Durbin panel data specification of local spending as a non-linear function of population size. The model is tested using a large local dataset over the 2003-2011 period for an aggregate measure of public spending. The empirical findings suggest a U-shaped relationship between population size and the costs of providing public services. In a second step, we investigate the role of geographical characteristics such as elevation and terrain ruggedness in the determination of the optimal jurisdiction size. Our results reveal that optimal city size decreases with elevation and increases with ruggedness.

*Keywords:* *Optimal Government Size, Spatial Panels, Spanish Municipalities, Topography.*

*JEL codes:* *C11, C23, H4, H7.*

---

<sup>\*</sup>Department of Economics and Public Finance, Universidad Autónoma de Madrid and GEN, Spain. E-mail: miriam.hortas@uam.es

<sup>†</sup>Department of Economics, Universidad Pública de Navarra, Spain. E-mail: vicente.rios@unavarra.es.

# 1 Introduction

The Great Recession has put the public sector performance at the forefront. The sharp decline in revenues combined with an increase in the expenditure needs have forced much of the public policy debate to focus on the quality and efficiency of public spending, especially at the local level. Small municipalities have been blamed for an inefficient provision of public services, as economies of scale are not maximized and spillover effects are not internalized. Amalgamation reforms have been put to use, given that larger municipalities are expected to exploit economies of scale and internalize spillovers, thus lowering costs in public goods provision. The extent to which consolidated local governments take advantage of economies of scale has promoted a heated debate among academics as the empirical evidence is, at best, mixed. Some previous empirical studies provide evidence of their success (see, Reingewertz, 2012) as long as municipalities did not exceed a critical size (Hanes, 2015), while others show no cost savings (Frere *et al.*, 2014) or even the existence of diseconomies of scale (Moisio and Uusitalo, 2013).<sup>1</sup>

The debate about the amalgamation of local governments stems from the fact that the political border of a given jurisdiction may not coincide with the economic boundary required for an efficient provision of local public goods, hence violating the principle of fiscal equivalence (Olson, 1969). Most likely, the success of amalgamations is closely related to the critical size for municipal mergers, that is, the limits of these economic boundaries, which may vary depending on (i) the public good provided, and (ii) the geographical map. On the one hand, the existence of economies of scale is often taken for granted. However, the literature offers very limited evidence on the number of services offering real potential to benefit from economies of scale. Thus, if they exist, they may depend on the public service provided and the units of measurement, such as the jurisdiction size or the size of the facility (Slack and Bird, 2012). On the other hand, factors other than economies of scale shape these economic boundaries. As noted in Dafflon (2012), the usual reference to determine the size of a jurisdiction is the number of service users; drawing a geographical landscape of the model that is flat. But this is not so in practice and, since the map is no flat, the spatial distribution of the population and physical geography also play a crucial role. Both sprawled and elevated locations might face higher public services costs than more compact or coastal-plain locations, not only because of the physical obstacles to infrastructure development but also by imposing higher provision costs on local public services, specially those based on networks. As a result, the optimal size of a jurisdiction hinges crucially on the geographical context of the jurisdiction.

While the relationship between the spatial distribution of the population and public service costs has been widely documented (see, e.g., Carruthers and Ulfarsson, 2003, 2008; Speir and Stephenson, 2002; Hortas-Rico and Solé-Ollé, 2010; Prieto *et al.*, 2015; Ihlanfeldt and Willardsen, 2018), the difficulty of obtaining data has limited the analysis of the role of geography on local public spending. The development of Geographical Information Systems (GIS) and the availability of satellite imagery and digital elevation models allow us to

construct measures that quantify different topographic features of the landscape with unprecedented detail. In order to analyse this unexplored issue, this paper uses a novel dataset on topography (i) to determine for which population levels there are (dis) economies of scale in the provision of public goods, so that a (decrease) increase in population size could lead to a decrease in local costs and, most importantly, (ii) to test whether this efficient scale varies depending on the topographic heterogeneity of the territory. In particular, the effects of geographical heterogeneity on the optimal size of municipalities are considered by means of two different exogenous variables: a terrain ruggedness index that precisely quantifies topographic irregularities of a jurisdiction's land area, and its mean elevation.

This paper brings the following contributions to the current policy debate on mergers and to the strand of research analysing the optimal size of local governments.<sup>2</sup> First, we explicitly consider the role of geography to test whether this exogenous factor affects the costs of providing local public services and, therefore, the optimal size of local jurisdictions. To that aim, we develop a theoretical model of cost efficiency that takes into account spatial interdependence in the provision of public goods among neighbouring jurisdictions (Salmon, 1987; Oates, 1999). The model builds upon similar models commonly used in the literature on local public spending (Borchding and Deacon, 1972; Bergstrom and Goodman, 1973), where local government and citizens' behaviour are jointly modelled so as to derive an estimating equation where the level of per capita spending is specified as a function of the demand for public services and their provision costs. The model solution leads to a Spatial Durbin Model (SDM) specification of local spending as a non-linear function of population size. Second, our empirical analysis employs modern Bayesian and Frequentist spatial panel data econometric techniques to validate the theoretical model and to perform inference on the relationship between population size and government spending. Importantly, our empirical strategy, which relies on the estimation of a spatial panel data model with spatial fixed effects and non-linearities in the population variable, helps us to solve some of the methodological problems that may affect the validity of the conclusions implied by previous empirical related literature (Solé-Ollé and Bosch, 2005; Hortas-Rico and Salinas, 2014).<sup>3</sup> The employment of panel data econometrics allows us to control for unobserved spatial heterogeneity by introducing municipal fixed effects, which in this context decreases the risk of obtaining biased estimation results.<sup>4</sup>

The rest of the paper proceeds as follows. The next section presents a brief review of the literature. The theoretical model is developed in the third section, while the econometric strategy and the data used in the empirical analysis are discussed in the fourth section. The fifth section presents the main empirical results of the paper. Finally, in the last section, we conclude.

## 2 Literature Review

According to the Fiscal Federalism literature, the decentralization of public services to sub-central governments allows, in the absence of externalities and economies of scale, a better adaptation of public policies to local preferences and needs (Oates, 1972), so that resources will be allocated with the greatest efficiency, accountability and responsiveness. On the one hand, sub-central governments are better informed than the central government in relation to local preferences, leading to a clear efficiency gain from delivering services in a decentralized fashion. Smaller government units may also stimulate competition between local jurisdictions for mobile residents and tax bases that will induce them to offer the best possible mix of taxes and services (Tiebout, 1956). On the other hand, citizens' participation, democratic control and the process of accountability also improve under a decentralized system, where the dominance of special-interest groups over citizens' participation are less likely to happen (Seabright, 1996; Inman and Rubinfeld, 1997; Oates, 2005; Hindriks and Lockwood, 2005). However, these positive effects might be offset by a number of negative impacts. In particular, as the size of jurisdictions decreases, economies of scale diminish and there is a greater likelihood of external effects on the provision of goods and services, leading to efficiency losses. Therefore, the optimal jurisdiction size is determined by a trade-off between scale economies and internalized spillovers on the one hand, and efficiency losses from heterogeneous preferences and reduced accountability and democracy on the other hand (Oates, 1972).

In this context, the analysis of the optimal size of local governments is especially important, as evidenced by the growing interest that has emerged in the literature since the seminal work of Oates (1972). Determining the appropriate geographic unit is key to ensuring efficiency in public provision, that is, to minimize the cost of public services. According to the Fiscal Federalism literature, this cost will mainly depend on the existence of (i) economies of scale, (ii) economies of density, and (iii) external effects.

The existence of **(i) economies of scale** can be derived from the existence of economies of scale in production, which will depend on the existence of fixed costs and technology; or consumption, which will depend on the degree of publicity and congestion costs of goods and services (Buchanan; 1965; Allen *et al.*, 1974). Following the seminal works of Borcheding and Deacon (1972) and Bergstrom and Goodman (1973), a good number of studies have investigated the effects of population size on the costs of providing local public services, without obtaining conclusive results.

The existence of **(ii) economies of density** also influences the costs of public services, as they imply a decrease in cost per user as population density increases. In the same vein, a greater dispersion of the population in the territory reduces the use of the density economies associated with public provision, thus increasing costs inefficiently (Carruthers and Ulfarsson, 2003; Hortas-Rico and Solé-Ollé, 2010), either because (i) the number of centres required to provide a certain level of service increases, or (ii) the average distance

between service users and facilities increases, rising transportation and infrastructure costs. This is the case of services with a clear spatial dimension, like those based on networks (i.e. sewerage system, public lighting, road maintenance or waste management). However, there are also factors that, at higher population densities, make it necessary to increase the level of output needed to obtain the same level of output in certain areas of expenditure, such as policing or street cleaning (Ladd, 1992).

The costs of local public services may also depend on the existence of **(iii) inter-jurisdictional spending spillovers** that occur when the benefits (or losses) of the public provision in one jurisdiction spread across its boundaries, affecting the welfare of residents in neighbouring locations (see, e.g., Brainard and Dolbear, 1967, and Pauly, 1970). As a result, the spending decisions of a local government will depend on spending policies chosen elsewhere. Another source of inter-jurisdictional strategic interaction occurs when voters use information on their neighbouring jurisdictions' public services and taxes to judge their own government's performance (Salmon, 1987). If voters consider relative performance, rational politicians will do the same and mimic their neighbours' decisions.

Therefore, the existence of economies of scale or density and the external effects on the provision of certain public services can lead to the fact that, when they are provided locally, the size of the jurisdiction is suboptimal and the provision costs are not minimized. However, the centralization of service provision can also lead to higher coordination costs, problems of governance and representation (Dur and Staal, 2008), decrease the degree of government accountability and adaptation of provision to local needs and preferences (Slack and Bird, 2012). The existing empirical evidence on the optimal size remains limited and, to some extent, controversial. Accordingly, further empirical research is required to clarify the nature of the link between population size and the costs of public services at the local level.

### 3 Theoretical Framework

We draw on the theoretical framework presented in Borcheding and Deacon (1972), and develop a model that incorporates spatial interactions and spillover effects among neighbouring jurisdictions. In this model economy, technological progress in the production of goods is assumed to be exogenous. The key distinct feature of the model with respect previous work is that includes technological externalities in the production of goods, which implies interdependence among the  $N$  municipalities denoted by  $i = 1, \dots, N$ . These municipalities have the same production possibilities but they differ because of their initial technology, spatial locations and their degree of spatial connectivity with the rest of the system.

We begin by considering the following aggregate production function of the economy

$i$ :<sup>5</sup>

$$Y_{it} = F(B_{it}, K_{it}, L_{it}) = B_{it} K_{it}^{\alpha} L_{it}^{\beta}, \alpha, \beta > 0 \quad (1)$$

where  $Y$  is the level of output,  $K$  is the level of capital,  $L$  is the level of labor,  $B$  is the level of technology and the subscript  $i$  and  $t$  denote the value of the variables for municipality  $i$  at period  $t$ . We introduce spatial correlation across municipalities' economies in the model by means of technological spillovers following Yu and Lee (2012) and Rios and Gianmoena (2018). Hence, technological advances in one municipality are allowed to have spillover effects on other economies. We specify the level of technology in the production of goods as:

$$B_{it} = B_{i0} e^{g_b t} \prod_{j \neq i}^N B_{jt}^{\psi w_{ij}} \quad (2)$$

The technology level in municipality  $i$  at period  $t$ ,  $B_{it}$ , is determined not only by its own initial level  $B_{i0}$  and the constant growth rate of technological progress  $g_b$  but also by the geometrically weighted average of the aggregate level of technology of the neighbouring economies,  $B_{jt}$ . The last term of Equation (2) captures the idea that spillovers arising from technological progress extend across jurisdictional borders with decreasing intensity because of geographical distance. This idea is in line with previous theoretical models analysing production in spatially interdependent contexts (Ertur and Koch, 2007; Yu and Lee, 2012; Rios and Gianmoena, 2018). In order to formalize this argument, the so-called spatial weight terms  $w_{ij}$  that represent the spatial interdependence between economies  $i$  and  $j$  are introduced. As is usual in the literature, these terms are assumed to be non-negative, non-stochastic and finite, with  $0 \leq w_{ij} \leq 1$  and  $w_{ij} = 0$  if  $i = j$ . Hence, we assume zero diagonal elements to exclude self-influence. It is further assumed that  $\sum_{j \neq i}^N w_{ij} = 1$  for  $i = 1, \dots, N$ , in order to avoid scale affects and ensuing explosive growth. The parameter  $\psi$ , with  $0 \leq \psi < 1$ , measures the relevance of spatial externalities in this context. Rewriting previous expression in log form and stacking over  $i$  we get that for each  $t$ :

$$\ln \mathbf{B}_t = \ln \mathbf{B}_0 + g_b t \boldsymbol{\iota}_n + \psi W \ln \mathbf{B}_t = [I_N - \psi W]^{-1} [\ln \mathbf{B}_0 + g_b t \boldsymbol{\iota}_n] \quad (3)$$

where  $\boldsymbol{\iota}_n$  is an  $N \times 1$  vector of ones and because of  $W$  is row-normalized  $[I_N - \psi W]^{-1} \boldsymbol{\iota}_n = \frac{1}{1-\psi}$ .<sup>6</sup> Therefore, the growth rate of technology in municipality  $i$  is given by  $\frac{\dot{B}_{it}}{B_{it}} = \frac{g_b}{1-\psi}$  which is greater than  $g_b$  due to the spillover effect if  $0 < \psi < 1$ .

The cost function of each municipality can be obtained after solving the minimization

cost problem below:

$$\begin{aligned} \text{Min} C_{it} &= s_{it}L_{it} + r_{it}K_{it} \\ \text{st} : Y_{it} &= F(B_{it}, K_{it}, L_{it}) \end{aligned} \quad (4)$$

where  $s_{it}$  represents the salary or price of the labour factor, and  $r_{it}$  the price of the capital factor in each municipality  $i$ .<sup>7</sup> Solving the previous cost minimization problem allows us to obtain the following cost function:

$$C_{it} = \kappa Y_{it} B_{it}^{-1} s_{it}^{\alpha} r_{it}^{\beta} \quad (5)$$

where  $\kappa = \left[ \left( \frac{\alpha}{\beta} \right)^{\beta} + \left( \frac{\alpha}{\beta} \right)^{-\alpha} \right]$ . This function measures the minimum cost of producing output given  $s_{it}$  and  $r_{it}$ . However, our interest does not rely in measuring the costs necessary to produce  $Y_{it}$ , but in those needed to produce a public service of quality,  $q_{it}$ . Therefore, we now establish how the output produced  $Y_{it}$ , is transformed into a determined level of public service provision  $q_{it}$ . It should be noted that, given the provision level of a public service, the results will not only depend on the degree of congestion in the service, but also in a variety of demographic, social, economic and geographical factors outside the control of the local governments. To take these factors into account, we further assume that quality of the service depends on the level of output  $Y_{it}$ , the population  $N_{it}$ , which measures the number of users services of the municipality and a set of exogenous cost factors  $Z_{it}$ , which would include demographic, social, economic and geographical factors. This relationship can be expressed as:

$$q_{it} = \frac{Y_{it}}{f(N_{it}) h(Z_{it})} \quad (6)$$

where  $f(\cdot)$  and  $h(\cdot)$  are functions with partial derivatives  $f'(\cdot), h(\cdot) > 0$ . A widely used formulation is the one provided by Borchding and Deacon (1972), who specify  $q_{it} = Y_{it}/f(N_{it})$  with  $f(N_{it}) = N_{it}^{\pi}$  where  $\pi$  is a positive scalar that ranges from 0 to 1 and captures the degree of congestion. But as Edwards (1990) shows, there are also many other alternative functional forms to capture the degree of publicness of local public goods. We adopt a mixed exponential-log function specification for  $f(N_{it})$  which has as a desirable property the fact that it imposes no a priori restrictions on the shape of the congestion function such that  $f(N_{it}) = e^{\pi_1 \ln N_{it} + \pi_2 \ln N_{it}^2}$  whereas  $h$  is defined as  $h(Z_{it}) = Z_{it}^{\gamma}$ .<sup>8</sup> Plugging Equation (6) into Equation (5) it is possible to obtain the cost function as:

$$C_{it} = \kappa [q_{it} f(N_{it}) h(Z_{it})] B_{it}^{-1} s_{it}^{\alpha} r_{it}^{\beta} \quad (7)$$

The empirical estimation of this cost function requires data on the quality of the provi-



sion level  $q_{it}$  and on the costs  $C_{it}$  level. Given the difficulty of having such data available, it is assumed that the level of output of services provided in each municipality is determined by the quantity demanded by the representative voter. The public goods' demand function of this representative voter is assumed to follow a Cobb-Douglas type, which is given by:

$$q_{it}^r = A p_{it}^\eta y_{it}^\delta \quad (8)$$

where  $q_{it}^r$  denotes the public good/service demand of the representative voter in municipality  $i$ ,  $p_{it}$  represents the tax-price (or the price of the public good) and  $y_{it}$  represents disposable income.  $A$  is a parameter that measures the individual preferences for the public service and  $\eta$  and  $\delta$  capture the price-elasticity and the income-elasticity of the demand, respectively. From Equation (8) it follows that the level of quality of the service demanded by the representative voter depends on the price of the public good in municipality  $i$ ,  $p_{it}$  and disposable income,  $y_{it}$ . According to the existing literature, the tax-price can be calculated from the percentage of expenditure (per service unit) paid by the representative voter, or from the tax-share ( $T_{it}^r$ ). Specifically, here, as in Downes and Pogue (1994), it is assumed that:

$$p_{it} = T_{it} \left( \frac{C_{it}}{q_{it}} \right) \quad (9)$$

This implies that we can rewrite Equation (8) as:

$$q_{it}^r = A^{\frac{1}{1+\eta}} T_{it}^{\frac{\eta}{1+\eta}} y_{it}^{\frac{\delta}{1+\eta}} C_{it}^{\frac{\eta}{1+\eta}} \quad (10)$$

Assuming that each municipality provides the amount of service demanded by the representative voter, the local public expenditure per capita  $G_{it}$  can be obtained substituting the demand function of Equation (10) in Equation (7):

$$G_{it} = \zeta B_{it}^{-1} s_{it}^\alpha r_{it}^\beta T_{it}^\eta y_{it}^\delta [f(N_{it}) h(Z_{it})]^{1+\eta} \quad (11)$$

where  $\zeta = \kappa^{1+\eta} A$ . Using  $f(N_{it}) = e^{\pi_1 \ln N_{it} + \pi_2 \ln N_{it}^2}$  and  $h(Z_{it}) = Z_{it}^\gamma$  and stacking observations over  $i$  and taking logs, the previous expression becomes:

$$\ln G_t = \varphi + \phi_1 \ln \mathbf{B}_t + \phi_2 \ln \mathbf{s}_t + \phi_3 \ln \mathbf{r}_t + \phi_4 \ln \mathbf{T}_t + \phi_5 \ln \mathbf{y}_t + \phi_6 \ln \mathbf{N}_t + \phi_7 \ln \mathbf{N}_t^2 + \phi_8 \ln \mathbf{Z}_t \quad (12)$$

where  $\varphi = \ln(\zeta)$ ,  $\phi_1 = -(1+\eta)$ ,  $\phi_2 = \alpha(1+\eta)$ ,  $\phi_3 = \beta(1+\eta)$ ,  $\phi_4 = \eta$ ,  $\phi_5 = \delta$ ,  $\phi_6 = \pi_1(1+\eta)$ ,  $\phi_7 = \pi_2(1+\eta)$ ,  $\phi_8 = \gamma(1+\eta)$ .

Stacking the parameters in  $\Phi = [\phi_2, \dots, \phi_8]$  and their corresponding covariates in  $\ln \mathbf{X}_t = [\ln \mathbf{w}_t, \dots, \ln \mathbf{Z}_t]$  and using the fact that  $\ln \mathbf{B}_t = [I_N - \psi W]^{-1} [\ln B_0 + g_b t \iota_n]$  we get:

$$\ln \mathbf{G}_t (I_N - \psi W) = \varphi (I_N - \psi W) + \phi_1 [\ln B_0 + g_b t \iota_n] + (I_N - \psi W) \ln \mathbf{X}_t \Phi \quad (13)$$

which is a Spatial Durbin Model (SDM) including spatial fixed effects and time-period fixed effects:

$$\ln \mathbf{G}_t = \tilde{\varphi} + \lambda_t \iota_N + \psi W \ln \mathbf{G}_t + \ln \mathbf{X}_t \Phi + W \ln \mathbf{X}_t \Theta \quad (14)$$

where  $\tilde{\varphi}$  is a vector of heterogeneous municipal intercepts such that  $\tilde{\varphi} = \varphi [I_N - \psi W] \iota_n + \phi_1 \ln B_{i0}$ ,  $\lambda_t = \phi_1 g_b t$  captures time-period effects (i.e, unobserved time heterogeneity) and  $\Theta$  is given by  $\Theta = -\psi \Phi$ . Finally, also note that the estimated spending parameters cannot be identified as cost function parameters (Downes and Pogue, 1994). However, cost differentials can be derived from the spending equation assuming that  $\alpha + \beta = 1$ . Then, identification of cost function parameters for the cost variables can be achieved if the reduced-form expenditure coefficients are divided by  $(1 + \eta)$ .

## 4 Empirical Model

The empirical SDM specification is given by:

$$\mathbf{Y}_t = \mu + \lambda D_t \iota_N + \rho W \mathbf{Y}_t + \mathbf{X}_t \Phi + W \mathbf{X}_t \Theta + \epsilon_t \quad (15)$$

where  $\mathbf{Y}_t$  is a  $N \times 1$  vector consisting of observations for the logarithm of government municipal spending per capita measured for every municipality  $i = 1, \dots, N$  at a particular point in time  $t = 1, \dots, T$ ,  $\mathbf{X}_t$ , is an  $N \times K$  matrix of exogenous covariates with associated response parameters  $\Phi$  contained in a  $K \times 1$  vector that are assumed to influence local government spending.  $\epsilon_t = (\epsilon_{1t}, \dots, \epsilon_{Nt})'$  is a  $N \times 1$  vector that represents the corresponding disturbance term which is assumed to be i.i.d with zero mean and finite variance  $\sigma^2$ .  $W$  is a  $N \times N$  matrix of known constants describing the spatial arrangement of the municipalities in the sample. The variable  $W \mathbf{Y}_t$  denotes contemporaneous endogenous interaction effects among the dependent variable,  $\rho$  is called the spatial auto-regressive coefficient.  $W \mathbf{X}_t$  is the matrix of exogenous regressors of neighbouring municipalities with its corresponding  $K \times 1$  vector of parameters  $\Theta$ .  $\mu = (\mu_1, \dots, \mu_n)$  is a vector of municipal fixed effects that captures cross-sectional heterogeneity due to differences in technological development and geographical location. Given that our sample spans from 2003-2011, this time interval covers two clear temporal sub-periods with their specific characteristics. The first one is

the pre-crisis period, characterized by the boom of a real estate bubble during 2003-2007. The second one is that of Great Recession caused by the international financial crisis and the crash of the housing market, from 2008-2011. Therefore, to control for the existence of time heterogeneity in spending patterns in these two periods, we include a dummy variable  $D_t$  that takes a value of 1 for the years 2008-2011 and zero otherwise. Thus,  $\lambda$  is the corresponding parameter measuring the effect of the Great Recession on spending and  $\iota_N$  is a  $N \times 1$  vector of ones. To investigate the relationship between the population and government spending by means of the Bias-Corrected Maximum Likelihood (BCML) estimator developed by Lee and Yu (2010) for spatial panels.<sup>9</sup>

As it is common in modern spatial econometrics analysis inference is based on a partial derivative interpretation and the computation of direct, indirect and total effects (LeSage and Pace, 2009; LeSage, 2014). The matrix of partial derivatives with respect to a change in a regressor  $X_k$  is given by:

$$\frac{\partial Y_t}{\partial X_t^k} = \left[ (I - \rho W)^{-1} \right] \left[ \beta^{(k)} I_N + \theta^{(k)} W \right] \quad (16)$$

In this context, *direct effects* (diagonal terms in Equation (16)) capture the effect on local government spending in  $i$  caused by a unit change in an exogenous variable  $X_k$  in  $i$ . *Indirect effects* (off-diagonal terms) can be interpreted as the effect of a change in  $X_k$  in all other municipalities  $j \neq i$  on the spending in  $i$ . Finally, *the total effect* is the sum of the direct and indirect impacts.

The model in Equation (15) can be contrasted against alternative spatial panel data model specifications which ultimately imply different spillover patterns such as the SLM, SEM and SDEMs (see Lesage, 2014). Another, relevant source of model uncertainty in spatial econometrics is the spatial weights matrix,  $W$ . Thus, we consider power distance decay with cut-offs at different thresholds of distance and a variety of k-nearest neighbour matrices with different weighting schemes. 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 following Da Silva et al. (2015) and Rios et al (2017) is employed here. This approach determines the posterior model probabilities (PMP) of the alternative spatial model specifications given a particular  $W$  (i.e,  $P(SM|W)$ ), as well as the PMP of different spatial weight matrices given a particular model specification (i.e,  $P(W|SM)$ ). Proceeding in this way, in Table (A.1) in the Appendix it is shown that the best specification is a SDM and the  $W$  is a 20 nearest neighbours matrix with weights characterized by an exponential decay of the 1% as distance increases.

## 4.1 Data

The empirical analysis conducted here is based on a large sample of 5,556 Spanish municipalities for the period 2003-2011. This eventual sample reflects the availability of budget data. It represents about the 70% of total municipalities and 87% of the whole population.

Spain is a suitable case in which to study the optimal municipal size for a variety of reasons. First, it is a highly decentralized country. It consists of three different levels of government: the central government, 17 regional governments and about 8,100 local governments, most of them with less than 1,000 inhabitants. Additionally, Spanish municipalities are responsible for delivering a huge range of public services traditionally assigned to local governments. These services include water supply, sewage and waste management, public lighting, road maintenance, local police and public transportation, among others. Second, local governments enjoy a relatively high degree of fiscal autonomy, as the services provided at the municipal level are financed mainly out of taxes and unconditional grants. Third, the Spanish local level of government is characterized by a high degree of fragmentation, which implies a structure of many independent units of government with very small populations, and limited public resources and management capacity. This fragmentation has been considered as one of the main causes of the lack of efficiency in the provision of public goods and services at the local level. In fact, the central government has promoted the reduction in the number of municipalities and the intensification of inter-municipal cooperation as a way of improving efficiency at the local level.<sup>10</sup>

As explained in the previous section, local public spending depends on both cost and demand factors. Next we briefly summarize the main variables that, according to the theoretical model and the existing empirical literature on local public spending, should be considered in our study<sup>11</sup>. Descriptive statistics, data sources and expected effects of the variables used in this study are provided in Table (A.2) in the Appendix.

### 4.1.1 Cost factors

The *cost factors* implied by the theoretical model presented in the previous section are related to population ( $f(N)$ ), input costs ( $s_{it}$ ) and other cost variables ( $Z_{it}$ ). Thus, we first consider total population, which is expected to affect local government spending through economies of scale. The population variable enters the equation with a quadratic term (*population*, *population*<sup>2</sup>), so as to account for the possible non-linear relationship between the number of inhabitants in a municipality and its per capita costs. Additionally, in Spain, the level of responsibilities of each municipality varies with population size<sup>12</sup>. Consequently, the more responsibilities the municipality assumes, the higher the local public spending should be<sup>13</sup>. To account for this effect - and to avoid that the population variable captures both its effect on costs and the effect of differences in the level of

responsibilities- we add three dummies representing the different levels of responsibility (*Responsibility1*, *Responsibility2*, *Responsibility3*) (Ladd, 1994; Solé-Ollé and Bosch, 2005; Hortas-Rico and Solé-Ollé, 2010).

Input costs are include in the model with a wage variable (*wage*), measured as the ratio between total wages and salaries paid and the number of workers. Given that wage data are not available at the municipal level, we have used provincial information.<sup>14</sup> Although these data refer to the private sector, the higher the wage in this sector, the higher the salary should be in the public sector in order to attract workers (Ladd, 1992).

Cost factors also include an additional set of variables that account for the *harshness of the environment* ( $Z_{it}$ ). First, the percentage of population over 65 (*population > 65*) and the percentage of immigrant residents (*migrants*) are also included in the model to approximate both the number of residents with special needs and the adverse conditions that may affect the level of provision necessary to maintain a certain level of service results (Ladd and Yinger, 1989; Ladd, 1992; Solé-Ollé and Bosch, 2005; Hortas-Rico and Solé-Ollé, 2010).

Second, we consider the number of population clusters per inhabitant (*pop.clusters pc*). According to the Spanish National Statistics Institute, population clusters are defined as an area encompassing ten or more buildings forming an urban layout (i.e. a grid conformed by streets, squares and other urban roads). This variable captures the spatial distribution of the population among the existing total number of clusters and is likely to affect local government spending through economies of density. As explained in Prieto *et al.*, (2015), the number of population clusters summarize best the urban sprawl characteristics of a given location by capturing tract dispersion. The expectation is that the cost of public services increases with the number of population clusters (Carruthers and Ulfarsson, 2003; Hortas-Rico and Solé-Ollé, 2010), as spatially extensive developments do not optimize on facility location of certain public services (Carruthers and Ulfarsson, 2008).

Third, the literature on local spending acknowledges the importance of *topography* in determining the cost of local public services (Slack and Bird, 2012). So far, however, previous empirical studies of the impact of geography on economic outcomes have relied, at best, on aggregated and coarse variables. The development of Geographical Information Systems (GIS) and the availability of digital elevation models allow us to construct different measures that quantify different topographic features of the landscape (i.e. altitude and ruggedness), as shown in the pioneering work by Burchfield *et al.* (2006). The presence of mountains limits accessibility, hence making basic infrastructure and public good provision more costly. In contrast, small-terrain irregularities have the opposite effect, as hillsides where public provision is more costly alternate with flat portions where public provision is less costly. Thus, two additional variables are constructed and included in the empirical specification to account for the impact of physical geography on spending. First, we introduce the mean altitude (*elevation*), to proxy the presence of mountains. Second, we include the *terrain ruggedness index*, to account for the presence of small-scale terrain

irregularities.<sup>15</sup>

#### 4.1.2 Demand factors

According to the theoretical predictions, we should also account for the effect of resources on the *demand for local public services*. On the one hand, we include the *tax-share* ( $T_{it}$ ) as a proxy of the price that residents face for public services (Downes and Pogue, 1994). Because property taxes are the major source of revenues at the local level, the tax share is specified as the property tax bill of the representative resident divided by the overall property tax revenues of the municipality. Its coefficient refers to the price elasticity of demand (parameter  $\eta$  in Equation (12)) and is hypothesized to be negative, since the higher the tax bill paid by the resident with a lower average income will be their demand for public goods and services and, therefore, the lower the municipality's level of expenditure. On the other hand, the disposable income of the representative resident ( $y_{it}$ ) includes the average per capita income of each municipality (*income pc*), whose coefficient (parameter  $\delta$  in Equation (12)) is the income elasticity of demand; and the per capita *transfers* received by each municipality (both current and capital). We expect a positive impact of these variables on local spending, since the higher the income of the representative resident, the greater the demand for public goods and, therefore, the higher the level of expenditure.

Finally, previous empirical studies suggest the need to include a last group of control variables that account for the effect of *political factors* on local spending (see, e.g., Bastida *et al*, 2013, Rios *et al*, 2017). According to the literature, the management of local public administration is the result of a combination of political factors (Astworth and Mesquita, 2006; Volkerink and de Haan, 2001). Both partisan politics and political strength influence local spending. Thus, we consider the ideology (*ideology*) and the political strength of the governing party (*gov strength*). Partisan ideology measures the impact of ideological differences on fiscal policy outputs. In Spain, after the 2003 elections, the main parties (i.e. the left-wing "Partido Socialista Obrero Español", and the right-wing "Partido Popular") hold more than half of the mayoral offices. This result held in both the 2007 and the 2011 elections. The remaining majoralties are held by other left-wing parties (Izquierda Unida) regionally-based right-and left-wing parties, as well as local parties and candidates that run as independents, mostly in small municipalities. We categorize the ideology variable for a considerable number of parties with an index ranging from 0 (left) to 10 (right) taken from the Deusto Polls database and in our own revision of electoral programmes. It is commonly argued that left-wing parties favour income redistribution and promote an active role of the Public Sector, which may increase public spending (Tellier, 2006).

The theoretical debate over the influence of government strength on the fiscal situation of public entities is grouped in two different hypotheses. While Roubini and Sachs (1989) or Borge (2005) suggest that coalition governments face higher deficits and spending levels, others have argued that divided governments have a moderating influence on fiscal policy

(Alesina and Rosenthal, 1994). We use the share of seats obtained by the ruling party in the local council.

## 5 Results

### 5.1 Baseline results

Table (1) presents the results of the SDM estimation for current spending by means of the BCML estimator using the 20-nearest neighbours with exponential decay of the 1%  $W$  matrix.<sup>16</sup> Column (1) reports the own-municipality coefficient estimates while Column (2) reports those of the neighbours. As shown in Column (1), the coefficient estimates of the spatially lagged dependent variable  $WY_t$  is both positive and significant. This result suggests that the current spending decisions of a local government are influenced by its neighbouring municipalities' spending decisions. Our model explains this positive spatial interaction through technological externalities but the positive effect obtained here is also compatible with previous findings of positive benefit spillovers and complementarity in local public goods provision (Bastida *et al*, 2013; Foucault *et al*, 2008; Hortas-Rico and Salinas, 2014; Rios *et al*, 2017). As regards our time-dummy controlling for the effects of Great Recession, we find a positive and statistically significant effect. This result can be explained by the time coverage of the study and the public policies implemented during this period. Note that the initial response of local governments to the international financial crisis and the crash of the Spanish housing bubble was to expand public spending to stimulate demand. In 2008, spending per capita increased by 8.4% whereas in 2009 the growth rate was 5.5%. In 2008, 75.1% of the municipalities in our sample increased spending per capita with respect to the level of 2007, whereas this percentage decreased to 65.4% in 2009. Therefore, the austerity measures that were implemented later on, during the period 2012-2013, are not captured by this variable.

We now turn to the interpretation of the control variables that fill out vector  $X$ . Overall, they are in accordance with the literature and display the expected signs. As mentioned in the previous section, a correct interpretation of the SDM estimates requires to look at the direct, indirect and total effects associated with changes in the set of regressors, instead of focusing on the single estimated parameters reported in Column (1). These effects are reported in Columns (3), (4) and (5), respectively.

#### INSERT TABLE (1) ABOUT HERE

We find that the direct effects are significant for several cost factors (including population, migration and population clusters), whose impact on spending is negative, whereas all demand factors apart from the political ideology exhibit a significant and positive ef-

fect on local spending. On the other hand, the indirect effects appear to be statistically significant for thirteen out of the fifteen variables considered. As refers to the cost factors, indirect effects are significant at the 1% level for all the variables, being the only exception the responsibility 3 variable that accounts for the level of responsibilities in large municipalities. As refers to the indirect effects of demand factors, we find that they matter for the income per capita, current transfers and the tax share. Taken together, these results show that the amplification phenomenon through space is particularly pronounced as in most of the cases they account for more than half of the total effect, thus corroborating the empirical relevance of spatial spillovers in this context. However, for some other variables (i.e. migrants or the number of population clusters), the indirect effects have a different sign to that of the direct effect. In fact, in these cases, the indirect effects tend to dominate the direct effect. A plausible interpretation of this result is that if all municipalities  $j \neq i$  experience a change in  $X^k$ , this will have a stronger effect in  $i$  than if only municipality  $i$  experiences a change in  $X^k$ . Note that this result is consistent with a highly interdependent and open economic environment, where changes in the rest of interacting municipalities of the system are more relevant than single municipal changes.

As mentioned above, the sum of direct and indirect effects allows us to quantify the total effect of the different control variables on government spending. When direct and indirect effects are jointly taken into account, Table (1) indicates that the total effect is statistically significant in all cases but the highest level of responsibilities and the government strength of the local council.

Let us now focus on the results for the variables of primary interest, i.e. those relating to population size and the existence of economies of scale. Our findings reveal that the relationship between population and government spending is non-linear and describes a U-shaped pattern: current spending decreases with population and increases with the square of the population. This result confirms the empirical evidence provided by Solé-Ollé and Bosch (2005) and Hortas-Rico and Salinas (2014), who also observed the existence of scale economies for municipalities with low population levels. Given the log-log specification, the estimated total effects for the linear term (-0.832) and the squared term (0.045) can be interpreted as elasticities, and the turning point  $N^*$  can be obtained after solving the condition for the minimum. Our estimates imply an optimal municipal size of  $\exp(9.29) \approx 10,865$  inhabitants.<sup>17</sup> Accordingly, the per capita spending of a given municipality may decrease due to the economies of scale up to a level of 10,865 inhabitants. Figure (1) shows the predicted relationship between the natural logarithm of population and the natural logarithm of the per capita cost in the interval ranging from 0 to 3.5 million inhabitants implied by a Monte Carlo simulation of 10,000 draws from the estimated effects. Our estimates display a higher efficient scale than the threshold of Solé-Ollé and Bosch (2005), which find that there are economies of scale for municipalities below 5,000 inhabitants, and a considerably higher threshold than that of Hortas-Rico and Salinas (2014), who quantify the optimal scale in 500 inhabitants. These discrepancies can be attributed to differences in (i) the cross-sectional and time sample and (ii) the econometric approach employed.



**INSERT FIGURE (1) ABOUT HERE**

Denote by  $\phi_1$  and  $\phi_2$  the corresponding total effects of the linear and squared terms of the log of the population ( $\tilde{N}$ ). Then, the marginal effect of a 1% increase in population can be obtained as:

$$\frac{\partial Y_t}{\partial \tilde{N}_t} = \phi_1 + 2\phi_2 \tilde{N}_t \quad (17)$$

However, in order to conduct inference on the effect of population we need to know if the estimated response given by Equation (17) is statistically distinguishable from zero. To that aim, we also need an estimate of the variance which is given by:

$$Var \left[ \frac{\partial Y_t}{\partial \tilde{N}_t} \right] = Var [\phi_1] + 4\tilde{N}_t^2 Var [\phi_2] + 4\tilde{N}_t Cov [\phi_1, \phi_2] \quad (18)$$

To derive the distribution of  $\phi_1$  and  $\phi_2$  in a first step we simulate the total effect  $\left( \frac{\partial Y_t}{\partial \tilde{N}_t} \right)$  implied by drawing  $d=1,000$  times from the variance-covariance matrix of the estimated parameters and for each draw we store the variance-covariance matrix of the average total effects  $\Sigma_{TE^{(d)}} = Varcov \left( \frac{\partial Y_t}{\partial \tilde{N}_t} \right)^{(d)}$ . In a second step, we compute the average variance-covariance matrix of the total effects over all the Monte Carlo draws  $d$ .<sup>18</sup> Following this procedure, Figure (2) shows the results of the marginal effects of a 1% increase in population. As observed, for municipalities with population sizes below the 10,865 cut-off, increasing population decreases costs per capita whereas above that threshold the effect turns positive. Overall, our findings provide evidence of the existence of economies of scale in local public goods provision as long as the municipality does not exceed a critical size. Beyond that population cut-off, diseconomies of scale arise. This result is consistent with the predicted cost pattern in the provision of local public goods observed in Figure (1).

**INSERT FIGURE (2) ABOUT HERE**

Table (1) also provides interesting information about the different control variables included in matrix  $X$ . Regarding the other *cost factors*, we first observe a positive effect of the share of migrants and the share of old population on per capita current spending. These variables measure the share of residents with special needs and the harshness of the environment (Ladd and Yinger, 1989; Ladd, 1992, 1994). Given that some services, such as health or social services, are mainly provided to this group of people, a municipality with more disadvantaged residents will spend more than other municipalities in providing the same level of these services. The positive relationship between the number of population clusters and per capita spending can be rationalized in terms of the dispersion of the population in the territory and its associated negative effects on density and scale economies

(Carruthers and Ulfarsson, 2008; Hortas-Rico and Solé-Ollé, 2010; Prieto *et al.*, 2015). Additionally, we find a positive effect of provincial wages on local spending. As suggested by Ladd (1992), this result may reflect that environments characterized by higher salaries also imply higher costs to attract public sector personnel.

With respect to the *demand factors*, some interesting findings emerge from the results. As expected, per capita income is positively related to local spending, which is in line with previous findings of Hortas-Rico and Salinas (2014), Rios *et al.* (2017) and Solé-Ollé (2006), and is consistent with the view that the higher the income of the representative resident, the greater the demand for public goods and services. Similarly, per capita current transfers have a positive effect on spending, with a coefficient that falls within the range established by the literature. In fact, our estimates suggest that the size of the implied flypaper effect is of 0.82 euros, which implies that an additional euro of current transfers leads to an increase in spending of 0.82 euros more than the stimulatory increase produced by one euro of extra income.<sup>19</sup> The price elasticity, identified as the estimated coefficient of the tax-share, is significant and around 0.67, a result that might seem counter-intuitive. However, this positive sign should be capturing the fact that higher tax collection translates into larger local public spending.

Interestingly, the effect of ideology on local spending is negative as in García-Sánchez *et al.* (2012) but clashes with the results presented in Bastida *et al.* (2013) and Rios *et al.* (2017), who find an insignificant effect. Nevertheless, this impact is theoretically well grounded given that a higher value reflects the party in government is more right-wing oriented, thus promoting a limited intervention of the government. Finally, the positive effect of a stronger government appears to be insignificant provides evidence against the view that higher electoral margins allow local politicians to expand their budget more easily (Bastida *et al.*, 2013).

## 5.2 The role of Geography

Along with the spatial distribution of the population, the physical geography (proxied here with the mean elevation and the terrain ruggedness index) is expected to have an outstanding impact on the cost of providing local public services. Even in the absence of economies of scale, the characteristics of certain public goods force us to consider these variables when determining the optimal size of jurisdictions. This is the case of services with a clear spatial dimension, like those based on networks (i.e. sewerage system, public lighting, road maintenance or waste management). For such services, both the spatial distribution of the population and the physical geography of the municipality determine the geographic contiguity of urban settlements and the connectivity of public service networks (Bel, 2011).

In this section we present a set of additional results where physical geography is in-

cluded in the model as a cost factor. The time-invariance of both variables has forced us to introduce them in the SDM model interacting with the population function. Hence, the coefficient of these interactions capture the relationship between population size and per capita local spending, conditional on the physical geography of the municipality. Figure (3) displays the model's forecast of the cost provision pattern and the marginal effects of an increase in population size for municipalities located in high and low altitudes. As expected, the findings suggest that the optimal size of municipalities hinges crucially on their topographic characteristics. On the one hand, municipalities with a high mean elevation (i.e. those above the median value) exhibit (i) a lower optimal size than those located in lower altitudes and (ii) important diseconomies of scale beyond 925 inhabitants (since costs increase considerably as population rises). It is important to notice that these locations are usually the less populated ones <sup>20</sup>. In this regard, Goerlich and Mas (2008) find that the population follows a spatial pattern of concentration on plains and coastal areas where altitude is lower. Their data show that mountain locations have experienced a depopulation process over time. As a result, the high altitude areas are characterized by a set of small, less accessible and worse connected municipalities where public service delivery becomes more costly and, therefore, their optimal size smaller. Figure (3) suggests that in low elevation municipalities the provision costs may decrease with the population once the threshold of 1,045.4 inhabitants is reached. Nevertheless, the marginal effects of an increase in population in this group of municipalities are not significant (see Figure (3d and Table A.3 in the Appendix). This precludes us from drawing any conclusions on the effect of increasing population size in municipalities with low elevation.

### INSERT FIGURE (3) ABOUT HERE

On the other hand, the terrain ruggedness has a non-negligible effect on the efficient scale of jurisdictions. The observed pattern in Figure (4a) for municipalities with a rugged terrain (i.e. those above the median value) suggests that they may largely benefit from the realisation of economies of scale, as opposed to those located in the plains (see Figure (4c)) which face diseconomies of scale above the threshold of 4,188.1 inhabitants. Thus, optimal municipal size appears to be lower in plains than in highly rugged terrains. This result is in line with Goerlich and Cantarino (2010), who find that the average size of municipalities located in less-rugged locations is below the national average.

According to previous empirical evidence (Burchfield et al, 2006; Gomez-Antonio *et al.*, 2016), small-terrain irregularities lead to more urban sprawl. That is to say, those locations with rugged terrain exhibit scattered development and lower population densities. Given that in more rugged locations natural barriers limit the territory available for new urban settlements, it is likely that the population centres tend to cluster in the space (i.e. they concentrate in those plain areas of the municipality where development is less costly)<sup>21</sup>, increasing the opportunities for agglomeration economies and economies of scale, and leading to a reduction in the costs of providing public services. On the contrary, the spatial distribution of the population is more compact in plain locations. As it is well

known, higher density typically increases public spending, as congestion costs may arise (Ladd, 1992). Thus, there is a need for a smaller optimal size of the jurisdiction that minimizes these congestion costs.

**INSERT FIGURE (4) ABOUT HERE**

A final concern is that both (i) the observed U-shaped pattern characterizing the link between population and provision costs and (ii) the estimated effects of geography on the optimal municipal size, might be contaminated by the omission of population density in the model, which could ultimately be determined by geography. In order to disentangle these effects, the previous analysis is complemented by estimating a pooled SDM dropping the municipal-fixed effects but including both the population density and the geographical variables as controls<sup>22</sup>. In addition, we investigate whether previous results on the relationship between geographical characteristics and the optimal city size remain. To that end, we estimate a set of regressions that allow us to capture nonlinear relationships and interactions between spending per capita, geography and population size<sup>23</sup>. Overall, the results of these robustness checks confirm our previous findings, giving us confidence on the robustness of our analysis. First, we find that the significant U-shaped relationship between population size and provision costs holds after controlling for population density (see Table A.5). Second, we find that the optimal size of the jurisdiction decreases with elevation (see Figure A.2 and Table A.6), which is in line with the finding that high-elevation municipalities face diseconomies of scale at a lower population threshold than low-elevation ones. Similarly, the results that state that the optimal municipal size is lower in plains than in more rugged terrains are confirmed, as we find that the optimal population size increases with the ruggedness of the terrain (see Figure A.3 and Table A.7).

## 6 Conclusions

Frequently, the high fragmentation of local governments has been considered as the main source of their inefficient performance, an argument that has become increasingly relevant after the Great Recession. Many academics advocate municipal mergers on the grounds that larger jurisdictions promote efficiency, hence reducing costs and spending. However, evidence on their success is, at best, mixed.

This paper seeks to complement previous empirical findings and contribute to our understanding of the dilemmas involved in designing the jurisdiction size of political systems. To that aim we build a theoretical model of local public spending which incorporates spatial interactions and spillover effects among neighbouring jurisdictions. The model solution leads to Spatial Durbin specification, where local spending is a non-linear function of population size, as well as other cost and demand factors. The population function

allows us to determine the existence of (dis)economies of scale in public service provision, whereas the spatial population distribution aims at capturing economies of density. Unlike previous studies, the model incorporates the physical geography among its costs variables. Both the spatial distribution of the population and the physical geography are key cost factors for certain public services, especially in those based on networks, as they determine the geographic contiguity of urban areas and the connectivity of public service networks. Hence, our analysis allows us to simultaneously test for the significance of the factors that, according to the Fiscal Federalism literature, are crucial for determining the optimal size of local governments (i.e. economies of scale, economies of density, strategic government interactions and spillover effects), and adds the role of physical geography to the debate.

The results provide evidence of a U-shaped relationship between population and local costs. The possibility to realise economies of scale exists as long as the municipality does not exceed a critical size (around 10,800 inhabitants). Beyond that point, the benefits arising from an increase in the size of municipalities are likely to decrease. Therefore, the threshold of 20,000 inhabitants suggested by the Spanish law 27/2013 Act of Rationalization and Sustainability of Local Administration could be too high and, as a consequence, mergers in municipalities between 10,000 and 20,000 inhabitants could imply efficiency losses.

In addition, the results indicate that both the spatial distribution of the population and the physical geography have a non-negligible impact on costs. On the one hand, more dispersed populations lead to diseconomies of density, hence increasing costs. On the other hand, the topography (especially the mean elevation of the municipality) is crucial in determining the optimal size of cities. According to these results, policy officials should encourage smaller jurisdictions to merge so as to reach their optimal size. Note that, given the high fragmentation of local governments in Spain, this reform would affect about 90% of municipalities. Nonetheless, a deeper analysis indicates that half of those local entities are located in the mountains, which implies a rather small efficient scale that would prevent them from merging.

Overall, these findings undermine quests for big mergers. In this context, other institutional options that have emerged as an alternative to amalgamations should be considered. These include inter-municipal cooperation for joint local service delivery, as it can be limited to certain public services where economies of scale can be achieved with fewer monitoring and political transaction costs and a minimal government restructuring (Sorensen, 2006; Bel, 2011). While the efficiency gains of mergers can be offset by a diminished quality of local democracy (Tavares, 2018), inter-municipal cooperation maintains political autonomy and is more flexible because it can have an organisational and spatial design shaped differently from case to case (i.e. for certain public services and with different neighbours, depending on the specific needs of the municipality). Overall, inter-municipal cooperation is typically found to be a more politically viable alternative than amalgamation (Feiock and Scholz, 2010) and, as such, it is an important alternative of interest to policy-makers facing local government reforms.

Finally, we would like to remark that the assumption of homogeneous public expenditure spillovers, regardless of municipal characteristics such as population or income distribution, may be too strong when analyzing local fiscal interactions. Therefore, although spatial multi-regime panel data models pose a number of econometric challenges such as their consistent estimation or the determination of the optimal number of regimes, we believe future research in the field of local finance may benefit from raising the bar in the modelling of these heterogeneities.

## 7 Acknowledgements

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

## Notes

<sup>1</sup>See Slack and Bird (2012) and Tavares (2018) for a review of the empirical literature on the existence of economies of scale, the efficiency gains and the democratic outcomes of consolidated local governments.

<sup>2</sup>Papers focusing on the optimal size of a coalition and her characteristics include Bolton and Roland, 1997; Alesina and Spolaore, 1997, 2005; Blume and Blume, 2007; Gordon and Knight, 2009; or Saarimaa and Tukiainen, 2010; or Hanes, 2015, among others.

<sup>3</sup>Unlike Solé-Ollé and Bosch (2005), where the omission of relevant spatial interaction terms could lead to bias/inconsistent and inefficient estimates (LeSage and Pace, 2009; Elhorst, 2014), we estimate a panel data version of the SDM which allows to quantify with accuracy the magnitude of spatial spillovers, thus minimizing the possibilities of overstating or under-estimating the optimal municipal size in this context. On the other hand, our analysis displays two key advantages when compared to the cross-sectional analysis based on the General Nesting Spatial Model (GNSM) of Hortas-Rico and Salinas (2014). First, unlike the GNSM, the SDM does not suffer from parameter identification issues (Elhorst, 2014).

<sup>4</sup>The empirical analysis carried out here presents some similarities with respect to Bastida *et al.* (2013) and Rios *et al.* (2017), as they also employ spatial econometric techniques to model local government spending. However, they do not focus on the link between population and spending. Rios *et al.* (2017) control for population density but does not include population and Bastida *et al.* (2013) do not consider the possibility of a non-linear effect.

<sup>5</sup>This function  $F$  is assumed to satisfy the following assumptions: (1)  $F$  is continuous, twice differentiable, positive, displays diminishing marginal products and constant returns to scale (i.e.  $\alpha + \beta = 1$ ), (2)  $F$  satisfies the Inada conditions.

<sup>6</sup>Note that  $[I_N - \psi W]^{-1} = \sum_r^\infty \psi^r W^r = \frac{1}{1-\psi}$  given that  $\sum_r^\infty W^r \psi = \psi$  and  $\sum_r^\infty \psi^r = \frac{1}{1-\psi}$  if  $|\psi| < 1$ .

<sup>7</sup> $r_{it}$  is assumed to be the constant across municipalities, due to the perfect mobility of capital.

<sup>8</sup>The signs of the first and second partial derivatives of  $f(N_{it})$  depend on the sign and relative magnitude of the population parameters.

<sup>9</sup>For an alternative approach to investigate non-linear effects in panels with large  $T$  while accounting for cross-sectional heterogeneity, cross-sectional dependence and feedback effects among the dependent and independent variables see Chudik *et al.* (2017).

<sup>10</sup>Accordingly, the recent reform of the local administration (27/2013 Act of Rationalization and Sustainability of Local Administration) establishes measures to encourage the voluntary merger of municipalities and shifting services of municipalities of less than 20,000 inhabitants upwards to the Provincial Councils. Recent examples of territorial reforms in Europe with a similar spirit can be found in Denmark, Greece, Germany or Switzerland. For a more detailed review see Steiner *et al.* (2016).

<sup>11</sup>Given that some of these control variables are not the main objective of this present study, they are discussed here only in brief. See Ladd and Yinger (1989) and Ladd (1992, 1994) for a review of arguments that justify their inclusion in the local spending model.

<sup>12</sup>Specifically, public provision is compulsory for all municipalities in services such as trash collection, street cleaning services, water supply, sewer system and street lighting, among others. Municipalities with a population greater than 5,000 inhabitants, additionally, have to provide parks, public libraries, and solid waste treatment. Municipalities with a population greater than 20,000 have to provide local police and social services. Finally, municipalities with a population higher than 50,000 inhabitants also have to provide public transport and environmental protection

<sup>13</sup>But, at least in Spain, this does not necessarily translate into different spending levels because local government tends to provide services even without explicit official responsibility. It is ultimately citizens' demands (and lack of intervention by other layers of government) that help to explain why a service is provided. Thus, the relationship between the level of responsibility and local spending might not be as evident as it might at first seem (Solé-Ollé and Bosch, 2005).

<sup>14</sup>This implies assuming perfect mobility of labour among the municipalities within the same province.

<sup>15</sup>These variables have been provided by Goerlich and Cantarino (2010). According to the authors, the inclusion of both variables is justified on the grounds that they are almost completely unrelated and both exhibit a very different spatial distribution.

<sup>16</sup>It is important to notice that these estimated parameters do not refer to the cost function but to the expenditure function. Nonetheless, cost effects can be recovered by dividing total effects by  $1/(1 + \eta)$ , where  $\eta$  is the estimated total effect of the tax-share.

<sup>17</sup>The equation implied by our estimates is given by the total effects:  $-0.832 + 0.045 * 2N^* = 0$ .

<sup>18</sup>Notice that this procedure to make inference is slightly more complex than that of linear models with interaction effects or with spatial effects but without interaction terms. The reason is that in a spatial model with interaction terms the researcher cannot directly draw from variance-covariance matrix of the parameters and the estimated parameters  $Varcov(\mu, \beta, \theta, \sigma)$  but, instead, the draws have to be taken from  $\Sigma_{TE} = Varcov\left(\frac{\partial Y_t}{\partial X_t}\right)$ .

<sup>19</sup>To calculate the size of the flypaper effect from our log-log regression and make it comparable with previous studies, elasticities are transformed as:  $\left(\frac{Y_t}{Tr_t}\right) TE_{Tr} - \left(\frac{Y_t}{I_t}\right) TE_I$  where  $TE_{Tr}$  and  $TE_I$  denote the

total effects obtained with the log-log specification for transfers and income and  $\frac{Y_t}{T r_t}$  and  $\frac{Y_t}{I_t}$  the ratio of spending to grants and income.

<sup>20</sup>According to our data, about 70 per cent of the municipalities with less than 1,000 inhabitants were included in the high-altitude subsample.

<sup>21</sup>Available data shows that the number of population clusters is higher in rugged locations but, at the same time, the average distance between the population clusters within those municipalities is lower.

<sup>22</sup>It is important to notice that the quality of the results obtained when estimating the pooled model are lower than those obtained by the fixed-effects model (R-square of 50.8% vs 89.5%) and, therefore, they should only be taken as indicative. In fact, the likelihood ratio tests on the joint significance of the fixed effects is LR = 76,824.60 (dof = 5521) with p-value = 0.00, suggesting that one should not employ the simplified model to perform inference. Obviously, the fact that in this setting we are not including spatial fixed effects precludes us to perform any comparison with the implied  $N^*$  reported in Table (1).

<sup>23</sup>In particular, we estimate alternative specifications of the following augmented SDM version:

$$\begin{aligned} Y_t = & \alpha + \lambda Crisis_t + \gamma_1 N_t + \gamma_2 N_t^2 + \gamma_3 G_t + \gamma_4 G_t^2 + \gamma_5 N_t \times G_t^2 + \gamma_6 N_t^2 \times G_t^2 + \\ & \delta_1 W N_t + \delta_2 W N_t^2 + \delta_3 W G_t + \delta_4 W G_t^2 + \delta_5 N_t \times G_t^2 + \delta_6 W N_t \times G_t^2 \\ & + \sum_k X_{t,k} \beta_k + \sum_k W X_{t,k} \theta_k + \epsilon_t \end{aligned} \quad (19)$$

where  $N$  denotes the logarithm of the population,  $G$  is the key geographical variable of the regression model interacting with  $N$  and  $X$  includes the set of previously defined demand and cost factors (including population density and the other geographical variable). Solving for  $N$  in  $\frac{\partial Y_t}{\partial N_t} = 0$  allows us to obtain the optimal population size  $N^*$  as a function of  $G$ . The simulation of the effects for elevation and terrain ruggedness are reported in the Tables A.6 and A.7 in the Appendix whereas the median forecasts of  $N^*$  stemming from the effects of models (1) to (6) in Tables A.6 and A.7 are shown in Figures A.2. and A.3., respectively.

## References

- Alesina, A.; Rosenthal, H. (1994): Partisans politics, divided governments, and the economy. New York: Cambridge University Press.
- Alesina, A., Spolaore, E. (1997): On the Number and Size of Nations, *Quarterly Journal of Economics*, 112(4), 1027-1056.
- Alesina, A., Spolaore, E. (2005): War, Peace and the Size of Countries, *Journal of Public Economics*, 89(7), 1333-1354.
- Allen, L.; Amacher, R.; Tollison, R. (1974): The economic theory of clubs: a geometric exposition. *Public Finance*, 3-4, 386-391.
- Astworth, S.; Mesquita, E. B. (2006): Delivering the goods: Legislative particularism in different electoral and institutional settings. *Journal of Politics*, 68(1), 168-179.



Bastida, F.; Guillamón, M.; Benito, B. (2013): Municipal spending in Spain: a spatial approach. *Journal of Urban Planning Development*, 139(2), 79-93.

Bel, G. (2011): Local services: Size, scale and governance. In IEB's Report on Fiscal Federalism. Barcelona: Institut d'Economia de Barcelona (IEB).

Bergstrom, T. C.; Goodman R. P. (1973): Private demand for public goods. *American Economic Review*, 63, 280-96.

Blume, L., Blume, T. (2007): The economic effects of local authority mergers: Empirical evidence for German city regions, *Annals of Regional Science*, 41, 689-713.

Borcheding, T. E., Deacon, R. T. (1972): The demand for the services of non-federal governments. *American Economic Review*, 62, 891-906.

Borge, L. (2005): Strong politicians, small deficits: Evidence from Norwegian local governments. *European Journal of Political Economy*, 21(2), 325-344.

Brainard, W.; Dolbear, F. T. (1967): The possibility of oversupply of local public goods, a critical note. *Journal of Political Economy*, 75, 86-90.

Buchanan, J. M. (1965): An economic theory of clubs. *Economica*, 32, 1-14.

Burchfield, M.; Overman, H.; Puga, D.; Turner, M. (2006): Causes of sprawl: a portrait from space, *Quarterly Journal of Economics*, 587-633.

Carruthers, J.I.; Ulfarsson, G.F. (2003): Urban sprawl and the cost of public services. *Environment and Planning B: Planning and Design*, 30, 503-522.

Carruthers, J.I. and Ulfarsson, G.F. (2008): Does Smart Growth Matter to Public Finance? Evidence from the United States, *Urban Studies*, 45 (9), 1791-1823.

Dafflon, B. (2012): Voluntary amalgamation of local governments: The Swiss debate in the European context. In Lago-Penas and Martinez-Vazquez (Eds) *The challenge of Local government size. Theoretical perspectives, International Experience and Policy reform*, Cheltenham, UK and Northampton, MA, USA: Edward Elgar.

Da Silva, D. F. C., Elhorst, J. P., da Mota Silveira Neto, R. (2015): Urban and rural population growth in a spatial panel of municipalities. *Regional Studies*, 51 (6), 894-908.

Downes, T.A.; Pogue, T.F. (1994): Adjusting school aid formulas for the higher cost of educating disadvantaged students. *National Tax Journal*, 47, 89-110.

Dur, R.; Staal, K. (2008): Local public good provision, municipal consolidation, and

national transfers, *Regional Science and Urban Economics*, 38(2), 160-173.

Edwards, J.H.Y. (1990): Congestion function specification and the publicness of local public goods, *Journal of Urban Economics*, 27, 80-96.

Elhorst, J.P. (2014): *Spatial Econometrics: From Cross-sectional Data to Spatial Panels*. Springer: Berlin New York Dordrecht London.

Ertur, C.; Koch, W. (2007): Growth, Technological Interdependence and Spatial Externalities: Theory and Evidence. *Journal of Applied Econometrics*, 22, (6), 1033-1062.

Feiock, R. C.; Scholz, J. (Eds). (2010): *Self-Organizing Federalism: Collaborative Mechanisms to Mitigate Institutional Collective Action Dilemmas*. New York: Cambridge University Press.

Foucault, M.; Madies, T.; Paty, S. (2008): Public spending interactions and local politics. Empirical evidence from French municipalities. *Public Choice*, 137, 57-80.

Frere, Q.; Leprince, M.; Paty, S. (2014): The impact of intermunicipal cooperation on local public spending, *Urban Studies* 51 (8), 1741-1760.

García-Sánchez., I.M., Mordan, N. and Prado-Lorenzo, J.M. (2012): Effect of the Political System on Local Financial Condition: Empirical Evidence for Spain's Largest Municipalities. *Public Budgeting and Finance*, 32, 40-68.

Goerlich, F.J.; Cantarino, I. (2010): Un índice de rugosidad del terreno a escala municipal a partir de modelos de elevación digital de acceso público. Documento de Trabajo 7/2010 Fundación BBVA

Goerlich, F.J.; Mas, M. (2008): Empirical evidence of population concentration in Spain, *Population-E* 63, 4, 635-650.

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.

Gordon, N., Knight, B. (2009): A spatial merger estimator with an application to school district consolidation, *Journal of Public Economics*, 93(5-6), 752-765.

Hanes, N. (2015): Amalgamation Impacts on Local Public Expenditures in Sweden, *Local Government Studies*, 41(1), 63-77.

Hindriks, J.; Lockwood, B. (2005): Centralization and Political Accountability. *CEPR Discussion Paper* 5125.

Hortas-Rico, M.; Salinas, P. (2014): Determinación de la escala mínima eficiente en la provisión de bienes públicos locales. Evidencia para los municipios españoles. *Revista de Economía Aplicada*, 66, 35-65.

Hortas-Rico, M.; Solé-Ollé, A. (2010): Does urban sprawl increase the costs of providing local public services? Evidence from Spanish municipalities. *Urban Studies*, 47, 1513-1540.

Inman, R.P.; Rubinfeld, D.L. (1997): The political economy of federalism. In D.C. Mueller (Ed.), *Perspectives on public choice*, 73-105. Cambridge: Cambridge University Press.

Ladd, H.F. (1992): Population Growth, Density and the Costs of Providing Public Services. *Urban Studies*, 29, 273-95.

Ladd, H. F. (1994): Fiscal impacts of local population growth: a conceptual and empirical analysis, *Regional Science and Urban Economics*, 24, pp. 661-686.

Ladd, H. F.; Yinger, J. (1989): *Americas Ailing Cities: Fiscal Health and the Design of Urban Policy*. Baltimore, MD: Johns Hopkins University Press.

Lee, L.F. and Yu, J. (2010): Estimation of Spatial Autoregressive Panel Data Models with Fixed Effects. *Journal of Econometrics*, 154, 165-185.

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

LeSage, J.P. (2014): Spatial Econometric Panel Data Model Specification: A Bayesian Approach. *Spatial Statistics*, 9, 122-145

Moisio, A.; Uusitalo, R. (2013): The impact of municipal mergers on Public Expenditures in Finland, *Public Finance and Management*, 13 (3), 148-166.

Oates, W.E. (1972): *Fiscal Federalism*. Harcourt Brace Jovanovich, New York.

Oates, W.E. (1999): An essay on Fiscal Federalism, *Journal of Economic Literature*, 37, 1120-49.

Oates, W.E. (2005): Toward a second-generation theory of fiscal federalism. *International Tax and Public Finance*, 12, 349-373.

Olson, M. (1969): The principle of fiscal equivalence: the division of responsibilities among different levels of government. *American Economic Review*, 59, 479-487.

Pauly, M. (1970): Optimality, public goods and local governments, a general theoretical

analysis. *Journal of Political Economy*, 78, 572-585.

Prieto, A.M.; Zofio, J.L.; Alvarez, I. (2015): Cost economies, urban patterns and population density: the case of public infrastructure for basic utilities, *Papers in Regional Science*, 94(4), 795-817.

Reingewertz, Y. (2012): Do municipal amalgamations work? Evidence from municipalities in Israel. *Journal of Urban Economics* 72, 240-251.

Rios, V.; Pascual, P.; Cabasés, F.(2017): What drives local government spending in Spain? A dynamic spatial panel approach, *Spatial Economic Analysis*, 12, 2-3, 230-250.

Rios, V.; Gianmoena, L. (2018): Convergence in CO2 emissions: A Spatial Economic Analysis with Cross-Country Interactions. *Energy Economics*, 75, 222-238.

Roubini, N.; Sachs, J. D. (1989): Political and economic determinants of budget deficits in the industrial democracies. *European Economic Review*, 33(5), 903-933.

Saarimaa, T., Tukiainen, J. (2010): Coalition formation and political decision-making: Evidence from Finnish municipal mergers. mimeo.

Salmon, P. (1987): Decentralization as an incentive scheme. *Oxford Review of Economic Policy*, 3, 24-43.

Steiner, R., Kaiser, C., and Eythorsson, G. T. (2016): *A comparative analysis of amalgamation reforms in selected european countries*. In Kuhlmann, S. and Bouckaert, G., editors, *Local Public Sector Reforms in Times of Crisis: National Trajectories and International Comparisons*, pages 23-42. Palgrave Macmillan UK, London.

Seabright, P. (1996): Accountability and decentralisation in government: An incomplete contracts model. *European Economic Review*, 40, 61-89.

Slack, E.; Bird, R. (2012): Merging municipalities: is bigger better?, in Moisio, A. (ed.) *Rethinking local government: Essays on municipal reform*, VATT Publications 61:83-123, Government Institute for Economic Research, Helsinki 2012.

Solé-Ollé, A. (2006): Expenditure Spillovers and Fiscal Interactions: Empirical Evidence from Local Governments in Spain, *Journal of Urban Economics*, 59, 32-53.

Solé-Ollé, A. and Bosch, N. (2005): On the relationship between authority size and the costs of providing local services: lessons for the design of intergovernmental transfers in Spain, *Public Finance Review*, 33(3), pp. 343-384.

Sorensen, R.J. (2006): Local government consolidations: the impact of political trans-

action costs, *Public Choice*, 127, 75-95.

Speir, C.; Stephenson, K. (2002): Does sprawl cost us all? Isolating the effects of housing patterns on public water and sewer costs, *Journal of the American Planning Association*, 68(1), 56-70.

Tavares, A. (2018): Municipal amalgamations and their effects: a literature review, *Regional Studies on Development*, 22(1), 5-15.

Tellier, G. (2006): Public expenditures in Canadian provinces: An empirical study of politico-economic interactions. *Public Choice*, 126(3-4), 367-385.

Tiebout, C. (1956): A pure theory of local expenditures, *Journal of Political Economy* 64 (5), 416-424.

Volkerink, B.; de Haan, J. (2001): Fragmented governments effect on fiscal policy: New evidence. *Public Choice*, 109(3), 221-242.

Yu, J. and Lee, L.F. (2012): Convergence: A Spatial Dynamic Panel Data Approach. *Global Journal of Economics*, 1 (1), 1-36.

## Table and Figures

Table 1: Baseline Estimation Results and Effect Decomposition.

Variable	Own Coefficient (1)	Neighbour's Coefficient (2)	Direct Effects (3)	Indirect Effects (4)	Total Effects (5)
Ln Population	-0.506*** (-13.84)	0.093 (1.10)	-0.511*** (-14.02)	-0.321** (-2.04)	-0.832*** (-5.28)
(Ln Population) <sup>2</sup>	-0.006** (-2.07)	0.028*** (4.67)	-0.005* (-1.81)	0.050*** (4.51)	0.045*** (4.08)
Responsibility 1	0.284*** (19.99)	0.203*** (3.69)	0.296*** (19.99)	0.679*** (6.51)	0.974*** (9.05)
Responsibility 2	0.197*** (14.72)	0.109* (1.93)	0.204*** (14.57)	0.408*** (3.79)	0.612*** (5.52)
Responsibility 3	0.121*** (8.75)	-0.003 (-0.05)	0.123*** (8.26)	0.116 (0.90)	0.239* (1.79)
Ln Pop. clusters pc	-0.010*** (-4.55)	0.025*** (3.71)	-0.010*** (-4.17)	0.039*** (3.05)	0.029** (2.25)
Migration (%)	-0.001*** (-3.98)	0.006*** (8.99)	-0.001*** (-3.35)	0.011*** (8.17)	0.010*** (7.43)
Population < 65	0.002*** (10.91)	0.001** (2.45)	0.002*** (11.01)	0.005*** (4.79)	0.007*** (7.01)
Ln Wages	0.044 (0.78)	0.117* (1.94)	0.050 (0.88)	0.272*** (4.07)	0.322*** (8.98)
Ln Income pc	0.015** (2.33)	0.241*** (17.90)	0.023*** (3.72)	0.486*** (20.54)	0.509*** (21.59)
Ln Current Transfers pc	0.100*** (45.78)	0.048*** (7.37)	0.103*** (47.33)	0.192*** (16.99)	0.296*** (25.51)
Ln Capital Transfers pc	0.007*** (8.37)	0.000 (-0.10)	0.007*** (8.50)	0.006* (1.77)	0.014*** (3.70)
Tax-share	0.117*** (19.98)	0.220*** (11.94)	0.126*** (20.72)	0.546*** (16.30)	0.672*** (19.59)
Ln Ideology	-0.001 (-0.37)	-0.009 (-1.57)	-0.001 (-0.51)	-0.018* (-1.76)	-0.019* (-1.88)
Government strength	0.022*** (2.63)	-0.011 (-0.36)	0.022*** (2.59)	0.000 (0.00)	0.022 (0.34)
Crisis	0.032*** (9.75)				
WY <sub>t</sub>	0.499*** (58.96)				
R-squared	0.895				
Log-Likelihood	25942.4				
$\hat{\sigma}_\epsilon$	0.023				

Dependent variable: Log of per capita current spending. Notes: t-stats in parenthesis, \* significant at 10% level, \*\* significant at 5% level, \*\*\* significant at 1% level. Inferences regarding the statistical significance of these effects are based on the variation of 1,000 simulated parameter combinations drawn from the variance-covariance matrix implied by the BCML estimates. The results are obtained using the 20 nearest neighbour's spatial weights matrix with exponential decay of the 1%.

Figure 1: Optimal Local Government Size

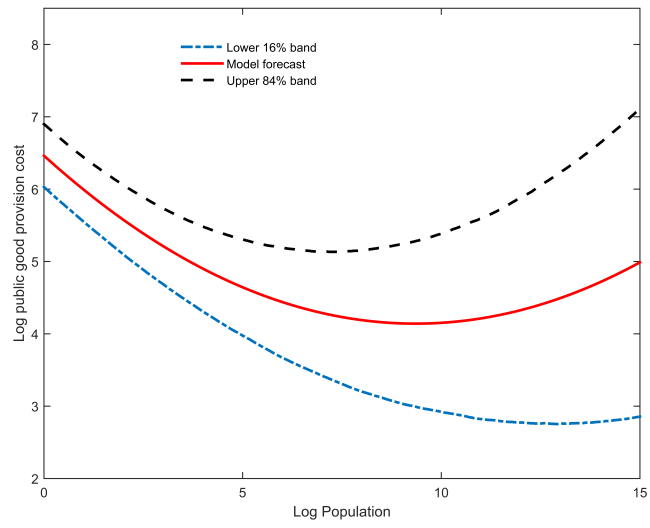




Figure 2: Population effects on costs

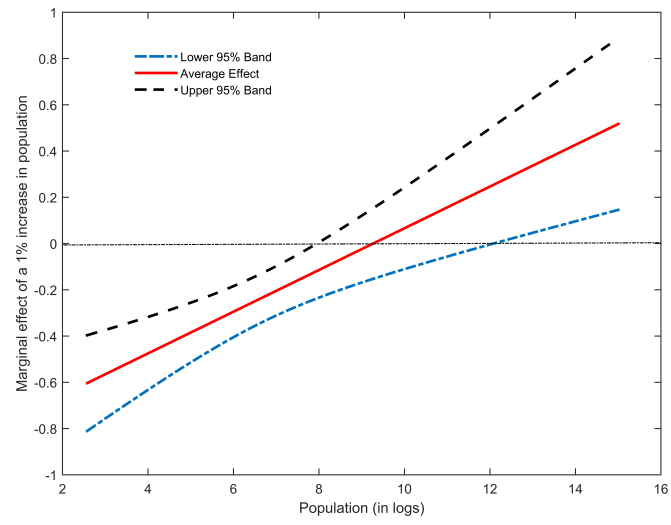
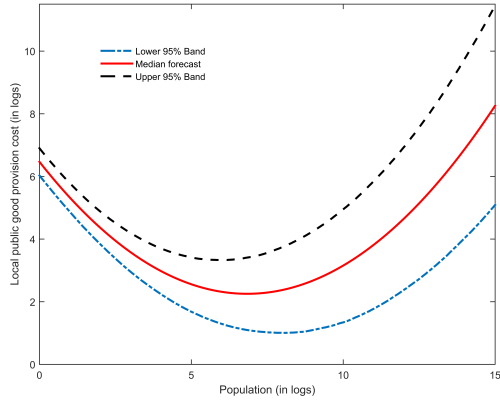
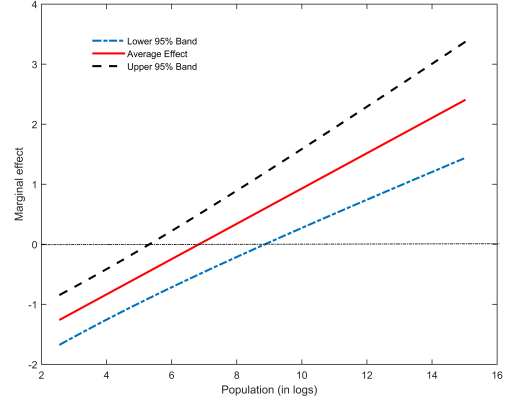


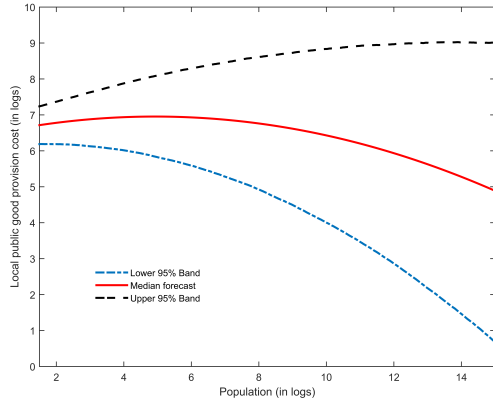
Figure 3: The effect of Geography (I): Terrain Elevation



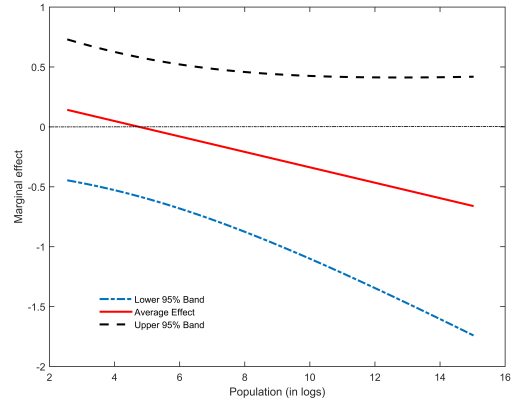
(a) Provision costs (high elevation)



(b) Marginal effects (high elevation)

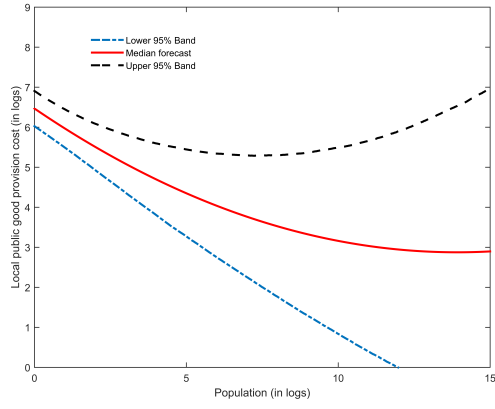


(c) Provision costs (low elevation)

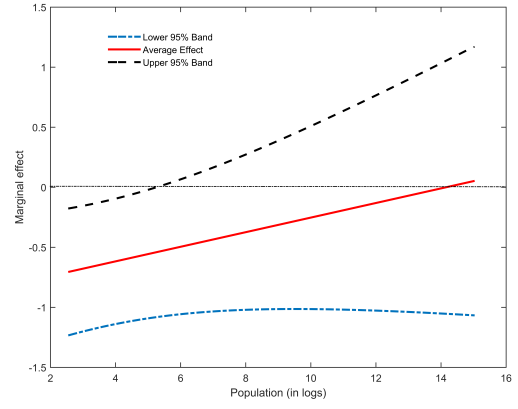


(d) Marginal effects (low elevation)

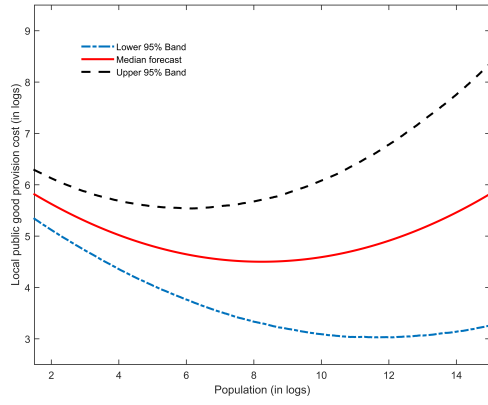
Figure 4: The effect of Geography (II): Terrain Ruggedness



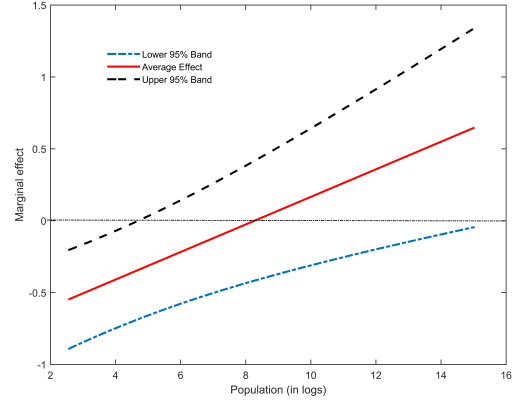
(a) Provision costs (high ruggedness)



(b) Marginal effects (high ruggedness)



(c) Provision costs (low ruggedness)



(d) Marginal effects (low ruggedness)