A FRAMEWORK FOR MEASURING POPULATION DISPERSION IN SPAIN:
METHODS AND SOURCES

Ángela Blanco Moreno*

Ana Santos Rodríguez**

D-2021-01

June 2021

The authors acknowledge the support from the IGAE (*Departmento de Informatica Presupuestaria*) for providing additional infrastructure to deal with large files. Also we thank the *Instituto Nacional de Estadística* for its assistance with the population database. In addition, we are grateful to Microsoft for granting free access to Bing Maps Distance Matrix API. Finally, the following persons have contributed to this Project: Margarita González Gómez (*DG Presupuestos*) by converting textual data from the Sistema de Información Urbana (SIU) to Excel. José Luis Matas Wic (IGAE) by contributing to the development of VBA code for the functions that take travel distances and travel durations from Bing Maps Distance Matrix API to Excel. José Luis Ares Martín, Alejandro Arriaga Piñeiro, and Juan Luis Serrano Martín de Blas (IGAE); and José Antonio Rodríguez and Juan Pedro Villarroel (*DG Presupuestos*) for their helpful assistance and support with Information and Communication Technologies issues surfaced throughout this Project. José García-Blanch De Benito has taken the time to read the Paper and contributed with valuable comments.

* Directorate General for Budget, Spain.
Correspondence: Angela Blanco Moreno
Subdirección General de Análisis y Programación Económica
Dirección General de Presupuestos
Avda. de Alberto Alcocer, 2, 28071 Madrid
e-mail: ablanco@sepg.hacienda.gob.es
Phone: 91 583 49 48. Fax: 91 583 50 93

** AIREF (previously Directorate General for Budget), Spain.

This document is available at: https://www.sepg.pap.hacienda.gob.es/sitios/sepg/es-ES/Presupuestos/DocumentacionEstadisticas/Documentacion/paginas/documentosdetrabajo.aspx

Abstract

The objective of this paper is to present a methodology to measure population dispersion in Spain's Regions. The ultimate goal is to provide a flexible tool for policy decision-making concerning the budgetary sustainability of fundamental public services: education, health and essential social services. Until now, it had not been sufficiently explored in Spain, unlike other spending drivers have been, such as population ageing with which it has clear interaction. We expect that the tool presented in this Paper will contribute to the functioning of the territorial administrations, which are required to maintain the full exercise of their autonomy within a framework of budgetary stability. This Paper is one of an exploratory nature. It analyses alternative ways of measuring dispersion through various primary components, as determined by the literature and in accordance with the availability of reliable statistical data. An additional paper will follow shortly as an application of the designed methodology, where we will use the identified primary components to quantify the designed indicators.

Keywords: Budgetary stability; fiscal decentralisation; population dispersion; indicators methodology.

JEL Codes: C43, H71, H77, R14.

CONTENT

1. INTRODUCTION

Population dispersion is one of the spending drivers in fundamental public services: education, health and essential social services (FPS),[1] and thus influences the sustainability of public finances. Geographical areas where population is highly dispersed would need to offer services at higher rates of intensity of resources, to ensure equal access. As a driver of public expenditure, it has not yet been explored in Spain as much as other drivers have been, such as population ageing, with which it interacts. By way of example, ageing interacts with the progressive depopulation of rural areas (the "*depopulated Spain*"), which relates, in turn, with the evolution of dispersion.

We have developed this work in the context of the analysis of budgetary stability in Spain. Due to its influence in financial sustainability, population dispersion should be considered in the decision-making process regarding the budgeting and planning of FPS. Considering the *de facto* federal structure of Spain, which provides that FPS are mainly managed by Spain's Regions,[2] the sustainability of fundamental public services at the national level is determined by the ability of Regional governments to comply with fiscal stability requirements. Indeed, according to Delgado, M. et al. (2016), "*a larger share of regions' spending on said fundamental public services limits regions' ability to adjust and comply with fiscal targets once their revenue-raising capacity is taken into account.*" Over the last four decades, regional governments have become accountable for delivering more than ⅔ of these fundamental public services. Thus, in Spain, addressing the sustainability of public spending in fundamental public services is dependent upon autonomous allocation decisions of Regional administrations within the mandatory framework of budgetary stability.

We claim that the sustainability of public spending requires disruptive innovative solutions[3] to address the provision of essential public services in geographical areas with high population dispersion. The first step to integrate population dispersion into decision-making processes

[1] The notion "fundamental public services" set in the Spanish Constitution constitutes an indeterminate legal concept that allows the legislator a very wide freedom of configuration. There have been extensive controversy regarding the content of such services. Our analysis will be at the national level and focuses on welfare state fundamental public services whose cost of provision is linked to population dispersion. Thus we focus on education, health and essential social services. We will discuss this issue in greater depth in a second paper on population dispersion coming shortly.
[2] According to article 15.1 of the Organic Law of Financing of the Autonomous Communities (LOFCA).
[3] https://ec.europa.eu/health/sites/health/files/expert_panel/docs/012_disruptive_innovation_en.pdf

would be to ensure the availability of valid indicators, to provide evidence-based choices. An objective of this Paper is to present a methodology through which valid indicators are used to quantify and measure population dispersion in Spain's Regions. A further objective is to provide a flexible tool for policy decision-making, integrating population dispersion in Spain's SPF sustainability analysis.

Against this backdrop, based on a literature review, we have identified a definition of population dispersion, as well as algorithms to formulate indicators and the primary components that compose them. We have designed dispersion indicators with a bottom up approach taking into account two leading vectors: the territorial vector and the population vector.

As for the territorial vector, our basic geographical units will be Spain's singular population entities and municipalities. They will be the reference to measure the indicators' primary components that we will present in this Paper. By aggregation, we will calculate the indicators' primary components at the provincial level. On its side, the algorithms used to formulate population dispersion indicators that we propose will yield provincial values. Once again, by aggregation, we will calculate the indicators at the regional and national levels.

As for the population vector, we establish two leading elements for modelling population dispersion: the people and the locations where the people reside. It gives rise to two approaches: "dispersion of people" versus "dispersion of locations." By creating this distinction, relevant issues surface concerning organizing the provision of FPS: On the one hand, less dispersion of people will trigger economies of scale when providing FPS (including Reference Services, especially when centrality is high). On the other hand, even if the dispersion of people is less than the dispersion of locations, the need to guarantee universal access to FPS implies decreasing productivity in the supply of services to the population, depending on the province's zone, which results in losses of economies of scale. Thus, and regarding decision-making, even if efficiency reasons would advise focusing on the dispersion of people, both people-based and location-based indicators should be combined for the set of indicators to be adequate from the perspective of equal access. In addition, population-based indicators would allow the replication of this model for specific population or age groups. While this work focuses on total population indicators, we believe it paves the way for further research in population dispersion dissimilarities according to age.

The Paper begins with a literature review in the search for a conceptual framework to define population dispersion and its measurement. We acknowledge that population dispersion is a multidimensional concept. Therefore, we select a set of indicators to gauge each of these different dimensions and formulate them adjusting the definitions found in the literature to the scope of our analysis: Spain's Regions.

It is organised as follows. After the introduction, we briefly describe the literature review. Then, we present our findings and propose a definition for population dispersion. Afterwards, we detail the formulation of the selected indicators adjusted to Spain's Regions. Finally, we summarise our conclusions and provide some policy implications. In addition, we provide five annexes to support our conclusions: Annex I lists the selected papers in the literature review; Annex II describes the nomenclature that we use in this paper; Annex III presents the tables with basic descriptive statistics on the primary components of the indicators; Annex IV provides the technical details of the indicators' formulation. Annex V provides additional technical details regarding our approximation for the maximum spatial separation attainable between population locations within a province. References are included at the end of the Paper.

2. LITERATURE REVIEW

By conducting a literature review, we have built a conceptual framework to define population dispersion and its measurement in Spain. The ultimate objective is to obtain indicators in accordance with the *state of the art* and adjusted to Spain's reality and regional structure. The technical characteristics of the search plan are found below. The list of selected papers is provided in Annex I. It also includes further references identified during the evaluation process of the literature review.

Search engines: Google

Search terms in Google:

English

- "Dispersion indicators" population OR geographical OR regional OR spatial
- "Population dispersion" OR "regional dispersion" OR "geographical dispersion" measure OR index OR indicator
- "Population concentration" OR "regional concentration" OR "geographical concentration" measure OR index OR indicator.

Spanish

- "Indicadores de dispersión" poblacional OR geográfica OR regional OR espacial
- "Dispersión poblacional" OR "dispersión de la población" OR "dispersión regional" OR "Dispersión geográfica" medida OR índice OR indicador
- "Concentración poblacional" OR "concentración de la población" OR "concentración regional" OR "concentración geográfica" medida OR índice OR indicador
- "Medidas de concentración y competencia."

Selection criteria:

- Scientific journals
- Institutional papers and documents from Universities, research institutes and public institutions that provided detail on:
  - Methodological criteria to define and measure the population dispersion or concentration and related indicators
  - The calculation of the dispersion indicators
  - The interpretation of the dispersion indicators
  - The benchmarking of the dispersion indicators
  - Free online availability.

Search period: June 2018 to August 2018.

3. DEFINITION OF POPULATION DISPERSION

In the literature review, we have found one indicator already in use in Spain for modelling population dispersion. Specifically, in the financing model for Spanish Regions. This indicator is the number of singular population entities (hereinafter, singular entities –SE–), a type of population unit defined for statistical purposes. According to the INE, "*A* ***singular population entity*** *is understood to be any habitable area of the municipal terminality, inhabited or exceptionally inhabited, clearly differentiated within the same and which is known by a specific denomination that identifies it without possibility of confusion*."[4]

Table 1 shows the SE distribution in Spain in 2019. According to this indicator, the Region with the highest spending needs due to population dispersion would be Galicia (49.09%), followed by Asturias (11.25%), Castilla y León (9.99%), Cataluña (6.32%) and Andalucía (4.55%). The rest of regions present weights between 0.42% and 2.77% (Figure 1).

The *Report of the Committee of Experts*[5] *for the review of the regional financing model* of July 2017 describes and assesses the number of SE, as it is one of the indicators included in the model's algorithm to calculate the "adjusted population": the standard need unit in the financing model of Spain's regions (see Tables 2 and 3).[6]

In our view, the number of singular entities existing in each Region in a given year is not a suitable indicator to capture the additional costs of providing fundamental public services because of population dispersion. The main reasons being that this indicator includes all SE, even if they are not inhabited. In addition, it does not take into account some relevant associated cost drivers, such as the distance between land uses, the distance to the capital of the province, and the extent to which most of the population is concentrated in locations closer to each other than the entire set of locations. Moreover, the current indicator is not normalised by the size of the province, while its extension is one of the cost drivers specifically considered in the allocation model. Finally yet importantly, we have verified through simulation techniques that the maximum spatial separation attainable by population entities within a province is unrelated to the number of entities, but only to the size of the province itself (Annex V).

---

[4] INE's definition as in http://www.ine.es/nomen2/Metodologia.do?L=1 and https://www.boe.es/boe/dias/2015/03/24/pdfs/BOE-A-2015-3109.pdf

[5] Hereinafter, CE.

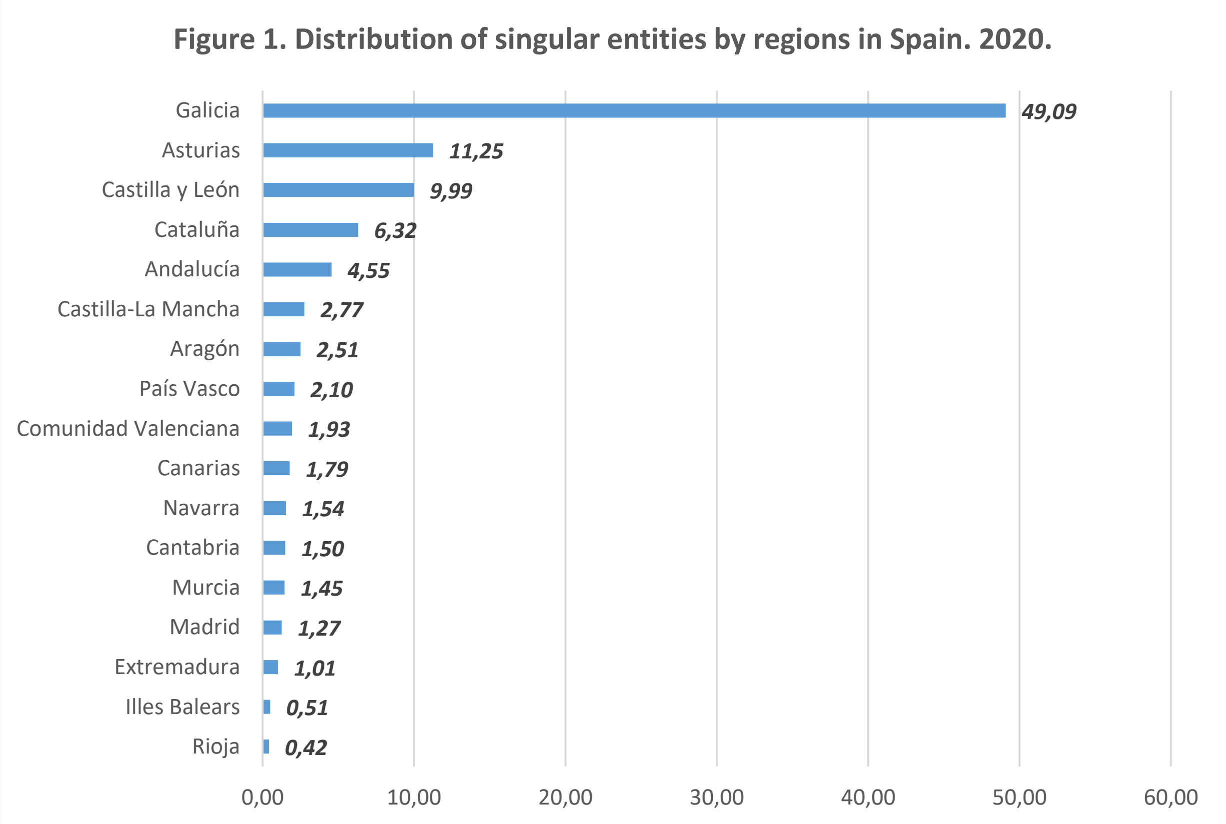
[6] https://www.hacienda.gob.es/CDI/sist%20financiacion%20y%20deuda/informaci%C3%B3nccaa/informe_final_comisi%C3%B3n_reforma_sfa.pdf

Table 1. Distribution of singular entities by regions in Spain. 2019

Regions	Number	Percentage
Total	**61,835**	**100**
Andalucía	2,815	4.55
Aragón	1,554	2.51
Asturias	6,955	11.25
Illes Balears	317	0.51
Canarias	1,108	1.79
Cantabria	930	1.5
Castilla y León	6,177	9.99
Castilla-La Mancha	1,713	2.77
Cataluña	3,903	6.32
Comunidad Valenciana	1,192	1.93
Extremadura	622	1.01
Galicia	30,362	49.09
Madrid	786	1.27
Murcia	896	1.45
Navarra	950	1.54
País Vasco	1,297	2.1
La Rioja	258	0.42

Source: INE and author's own work.

Figure 1. Distribution of singular entities by regions in Spain. 2019.



Source: INE and author's own work.

TABLE 2: Variables used for the calculation of the adjusted population or standard need units in the financing model of Spain's regions Current system vs. The proposal of the Committee of Experts (CE)	
System 2009	***CE's proposal***
Basic indicators of demand:	
Adjusted covered population in terms of health expenditure	Adjusted covered population in terms of health expenditure
School-age population (0-16)	School-age population (0-18) + Enrolment in higher training cycles*
	University students enrolled in public centres (degree and master)*
Population over 65 years	Population over 65 years of age weighted by age groups
	Population at risk of poverty or exclusion
Total population	Total population
Corrective variables:	
Surface area	Surface area
Dispersion	**Dispersion (**)**
Insularity	Insularity
	Fixed costs
New variables to incorporate or at least to study:	
	Price level
	Per capita income
	Floating and linked population
	Orography
	Improved dispersion indicator (**)
	Diseconomies of scale in health

(*)The CE also left open other possibilities for students aged from 0 to 3 and university students. In both cases, we could use the total resident population as well as the actual enrolled one. In the case of the university, if we chose the first option, we would need a compensating mechanism for the net receiving communities regarding the expenses generated by non-residents students.
(**) The dispersion indicator is the number of singular entities existing in each Region in the current year. As regards the dispersion of the population, the CE considers that the current indicator, based on the number of singular entities, could be improved and recommends considering the advantages and disadvantages of other possible alternatives that take into account the distances between urban centres and the geographical distribution of the population.
Source: Retrieved from De la Fuente, A. (2017).

TABLE 3: Variables used for the calculation of the adjusted population or standard need units in the financing model of Spain's regions	
VARIABLES	**WEIGHTs**
Population	30%
Surface area	1.8%
Dispersion	0.6%
Insularity	0.6%
Adjusted covered population in terms of health expenditure	38%
Population aged over 65	8.5%
Population aged 0 to 16	20.5%

Source: Ministry of Finance of Spain. Retrieved from:
https://www.hacienda.gob.es/es-ES/Areas%20Tematicas/Financiacion%20Autonomica/Paginas/Regimen%20comun.aspx

Indeed, according to the CE's report, "*the current indicator, based on the number of singular population entities, could be improved and recommends considering the advantages and disadvantages of other possible alternatives that take into account the distances between urban centres and the geographical distribution of the population*."

In the literature review, we have not found a definition of population dispersion commonly accepted. From the analysis of the selected papers, we have identified several approaches to dispersion in different fields, such as urban sprawl, industrial concentration, population concentration, residential segregation, etc.

Borrowing from them, mainly from urban sprawl literature, we understand that population dispersion reflects a stylised judgment about a general pattern of land development factored by multiple dimensions. In addition, the judgement can be either static (at a specific date) or dynamic (for a period).

In general, different agents or observation units may use the territory for alternative developments, with residential and not residential purposes: living (people, residences), working (jobs), economic activity (companies), etc. Therefore, alternative patterns of land development may be considered depending on the specific type of land use under consideration (population residences, employment, companies, etc.).

The definition of dispersion according to our objective is in relation with residential purposes, the people who use the land of singular entities (or municipalities –MUN–) being the observation units. Thus, we identify land uses with parcels of land, whose boundaries are as per the existing territorial units (administrative or political) into which Spain is organised. Specifically, we use the basic local entities of territorial division in Spain: singular entities or municipalities. The terminology used for the basic territorial elements integrating the definition and modelling of population dispersion is provided in Annex II.

We could have also focussed on land uses hosting facilities for EPS across Spanish provinces. This is not the approach used in this Paper because our goal is to identify spending drivers independently from prior public investment decisions (exogeneity). This will facilitate further analyses regarding the quantification of the effect of population dispersion on per capita FPS spending.

Based on the above, **we define population dispersion as a multidimensional concept representing a specific pattern of land use by the population for residential purposes that exhibits low levels of some combination of six distinct dimensions**:[7]

- Proximity
- Centrality
- Nuclearity
- Density
- Concentration
- Continuity.[8]

We now define the six dimensions of dispersion based on the literature review and the scope of our analysis.

Proximity[9]

Proximity is the degree to which pairs of land uses are close to each other within a given geographical area. Low proximity is associated with high dispersion.

In our context, it is the degree to which SE (or municipalities) within the same province are close to each other. An additional nuance can be made, depending on whether the focus is on the physical location ("**geographical proximity"**) or the people that inhabit them ("**population proximity**").

Centrality[10]

Centrality is the degree to which land uses are located close to a point taken to be the centre of the geographical unit: the so-called Central Business District (CBD).[11] Low centrality is associated with high dispersion.

[7] Galster, G. et al. (2001) use a more elaborate definition that accounts for additional factors, such as clustering and mixed uses. Regarding clustering, we have excluded this dimension because of the difficulty to operationalise this concept with the available data for Spain. As for mixed uses, this concept compares different types of use of land and we focus only on residential use (population).

[8] We have included continuity although this concept is difficult to interpret in the context of our analysis where we are unable to incorporate the fact that there are vacant land areas within municipalities. Some authors have calculated some density indicators while accounting for the presence of vacant land areas, but undertaking such an exercise would require resources beyond our present ones. Nonetheless, we propose some indicators for continuity that would provide a rough insight of that dimension within a province.

[9] The main papers that address the **proximity** dimension are the following (please, refer to Annex I):
Massey, D. et al. (1988); Midelfart-Knarvik K.H. et al. (2000); Galster, G. et al. (2001); Iceland, J. et al. (2002); Lage de Sousa, F. (2002); Midelfart-Knarvik K.H. et al. (2002); Wassmer, R. et al. (2005); Franz, G. et al. (2006); Dominicis, L. et al. (2007); Gindle, G.A. (2010); OTN (2011); Folch, D. (2012); Nuñez, G. (2014); Boontore, A. (2014); Alvarez, C. et al. (2015); Pereira, R.H.M. et al. (2013); Pereira, R.H.M. et al. (2015).

[10] The main papers that address the **centrality** dimension are the following: Massey, D. et al. (1988); Anas, A. et al. (1998); Torrens, P.M. et al. (2000); Galster, G. et al. (2001); Malpezzi, S. et al. (2001); Glaeser, G. et al. (2004); Quinn, L.M. (2004); Lee, B. (2006); Monkkonen, P. (2010); Folch, D.C. (2012); Pereira, R.H.M. et al. (2013); Muñiz, I. et al. (2013); Boontore, A. (2014); Lee, S. (2015); Pereira, R.H.M. et al. (2015); Gandhi, S.R. et al. (2016); Kavanagh, L. et al. (2016); Torres, T. (2017); Ottensmann, J.R. (2017(a)); Ottensmann, J.R. (2017(b)).

[11] This is according to the terminology in urban sprawl. Please refer to Annex II on nomenclature.

In our context, it is the degree to which SE (or municipalities) within the same province are close to its CBD. An additional nuance can be made, depending on whether the focus is on the closeness to the CBD of the physical locations (geographical centrality) or of the people that inhabit them (population centrality).

The centrality of a geographical area increases as the radius from the CBD within which the greater proportion of population is located shortens. Conversely, an area will exhibit lower centrality where greater distances from the centre are required to contain the same proportion of people. In a sense, centralisation and concentration are synonymous; however, a highly concentrated distribution of population is not a highly centralised one unless people cluster at the CBD; while all highly centralised distributions are by definition also highly concentrated.

Nuclearity[12]

Nuclearity is the extent to which a geographical area is characterised by a mononuclear pattern of development. Low nuclearity is associated with high dispersion.

In our context, nuclearity is the degree to which a province is characterised by a mononuclear pattern of residential development. Nuclearity is maximised if a province has a mononuclear development, which happens if the CBD is the only locus of intense development. If the population is located over several intensely developed places and each represents a substantial proportion of the total, it is polynuclear. Centrality is a characteristic best suited to mononuclear areas.

Density[13]

Density is the average number of population per km^2 within a geographical area (population density). It is a widely used indicator of population dispersion, associating low densities to high population dispersion. Nonetheless, density is only one dimension of dispersion that, as said,

[12] The main papers that address the **nuclearity** dimension are the following (please, refer to Annex I): Galster, G. et al. (2001); Franz, G. et al. (2006)

[13] The main papers that address the **density** dimension are the following (please, refer to Annex I): Clark, C. (1951); Torrens, P.M. *et al.*; (2000); Galster, G. *et al.* (2001); Malpezzi, S. *et al.* (2001); Ewing, R. *et al.* (2002); Cosby, K.L. (2004); Glaeser, E.L. *et al.* (2004); Song, Y. et al. (2004); Tsai, Y.H. *et al.* (2005); Franz, G. et al. (2006); Goerlich, F.J. (2006); Lee, B. (2006); Goerlich. F.J. *et al.* (2008); Martinez-López, M. (2010); Folch, D.C. (2012); Santos, J.M. *et al.* (2012); Angulo, A.M. *et al.* (2013); Muñiz, I. *et al.* (2013); Bertaud, A. (2014); Boontore, A. (2014); Alvarez, C. *et al.* (2015); Lee, S. (2015); Gandhi, S.R. *et al.* (2016); Tian, S. *et al.* (2017).

needs further elements to be properly captured. In our context, density is the degree to which Spain's provinces are thickly populated throughout their territory. The lower the population density the higher the population dispersion.

Concentration[14]

Concentration is the degree to which population is located disproportionately in a relatively small area rather than evenly spread throughout the geographical area of analysis. Low concentration is associated with high dispersion.

In our context, it is the degree to which population is located disproportionately in a relatively small area rather than spread evenly throughout the province.

Continuity[15]

In the urban sprawl context, "*Continuity is the degree to which developable land has been built upon at urban densities in an unbroken fashion*"; "*Continuous development may occur at any level of density, although the steady outward march of low-density development in concentric rings from the urban centre or core is commonly characterised as sprawl*" (Galster, G. et al. (2001)). For the purpose of this work, continuity is the degree to which the land area of a province is developed for residential purposes in an unbroken fashion. Thus, continuity indicates whether land area is developed for residential purposes uniformly across the province or in a few locations.

[14] The main papers that address the **concentration** dimension are the following (please, refer to Annex I): Clark, C. (1951); Stephan, G.E. (1977); Massey, D. et al. (1988); Fosset, M. (1990); Ellison, G. et al. (1997); Ferreira, E. et al (1997); Otterstrom, S.M. (1997); Anas, A. et al. (1998); Callejon, M. (1998); Bertraud, A. et al. (1999); Mayor, M. et al. (2000); Torrens, P.M. et al. (2000); Galster, G. *et al.* (2001); Malpezzi, S. et al. (2001); Martori, J.C. et al. (2001); Nuñez, S. et al. (2001); Aiginger, K. et al. (2002); Iceland, J. et al. (2002); Lage de Sousa, F. (2002); Sanchis, A. et al. (2002); Alonso, O. et al. (2003); Otterstrom, S.M. et al. (2003); Aiginger, K. et al. (2004); Bárcena, E. et al. (2004); Cosby, K.L. (2004); Lemelin, A. (2004); Paluzi, E. et al. (2004); Park, S. et al. (2004); Ayuda, M.I. et al. (2005); Bertinelli, L. (2005); Coelho Avila, P. (2005); Martí-Henneberg, J. (2005); Santa Maria, M.J. et al. (2005); Tsai, Y. (2005); Alonso, O. et al. (2006); Franz, G. et al. (2006); Goerlich, F.J. et al. (2006); Lee, B. (2006); Picchizzolu, R. (2006); Castañeda, C. (2007); Do, Q.A. et al. (2007); Dominicis, L. (2007) et al.; Aso, Y. (2008); Combes, P.P. et al. (2008); Cutrini, E. (2008); Goerlich. F.J. *et al.* (2008); He, C. et al. (2008); PUCC (2009); Ayuda, M.I. et al. (2010); Monkkonen, P. (2010); Otto, A. et al. (2010); Salazar, X. et al. (2010); Atienza, M. et al. (2012); Campos, C. (2012); Santos, J.M., et al. (2012); Jurado, I. et al. (2013); Lis-Gutiérrez, J.P. (2013); Liu, Z. (2013); Malpezzi, S. (2013); Muñiz, I. et al*.* (2013); Bertraud, A. et al. (2014); Boontore, A. (2014); Lemelin, A. (2014); Nuñez, G. (2014); Santic, D. (2014); Zurita, J. (2014); Allen, R. et al. (2015); Dauth, W. et al. (2015); Lee, S. (2015); Pereira, R.H.M. et al. (2015); Rastvortseva, S. et al. (2015); Van Egeraat, C. (2015); Centurion, I. (2016); Gandhi, S.R. et al. (2016); Martori, J.C. et al. (2016); OECD (2016); Sobrino, J. (2016); Van Egeraat, C. (2016); EU (2017); Gude, A. et al. (2017); Khoirunurrofik (2017); Maslikhina, V. (2017); Ottensmann, J.R. (2017 (b)); Torres, T. (2017); OECD (2018 (a)); OECD (2018 (b)).

[15] The main papers that address the **continuity** dimension are the following (please, refer to Annex I): Galster, G. et al. (2001); Malpezzi, S. et al. (2001); Tsai, Y.H. (2005); Wassmer R.W. et al. (2005); Franz, G. et al. (2006); Muñiz I., et al. (2013); Ghandi, S.R. et al. (2016); OECD (2018(b)).

4. MODELING POPULATION DISPERSION INDICATORS: METHOD AND SOURCES

This work analyses alternative ways of measuring dispersion through various primary components, as determined by the literature and in accordance with the availability of reliable statistical data. Based on the literature review, we have identified the algorithms to build dispersion's indicators, their primary components and the data sources to measure them. At this point, we detail all of them and formulate dispersion indicators.

Each dispersion dimension captures a specific aspect of dispersion and thus requires specific primary elements for its measurement. The primary components we use are:

- **Population** in land uses
- **Distances** between any two land uses within the same province[16]
- **Breadth of the province**
- **Nuclei**
- **Land area** of land uses (municipalities)[17]
- **Crude population density** (municipalities).

By way of introduction, we summarise next the role of these primary components in the modelling of dispersion indicators. Within a province, we begin with the ***population*** that uses the territory (land uses) for residential purposes, and we gauge the spatial separation between land uses' locations through the ***distances*** between their centroids. For each singular entity, we define its centroid as its geographical coordinates provided by the National Geographical Institute (IGN). For each municipality, we define its centroid as the centroid of the singular entity that holds the capital as indicated by the IGN. To improve comparability between provinces of different size and shape we rescale distances within a province via the ***breadth of the province***. In addition to the spatial separation between land uses, we analyse the spatial separation to the centre of the province. To this end, we use again distances and complement with measures to gauge the extent to which population locates close to the centre through the comparison of the accumulation of population and the accumulation of land area around it.

[16] Please, notice that we always refer to distances between any two land uses within the same province. No distance between two SE or municipalities of different provinces is involved in the calculations.
[17] We lack information on land area for singular entities.

Afterwards, via the number of ***nuclei*** (loci of intense residential development), we look at the extent to which people within a province settle in just one or in many nuclei. Then, through the ***crude population density*** (direct ratio population to ***land area***), we analyse the extent to which the territory of a province is thickly populated. Next, by comparing the accumulation of population against the accumulation of land uses, we examine whether population concentrates in a limited number of land uses. Finally, by comparing built land area to total land area we evaluate the degree to which a province's land area is uniformly developed for residential purposes.

We formulate dispersion indicators with a bottom up approach taking into account two leading vectors: the territorial vector and the population vector.

As for the territorial vector, our basic geographical units will be Spain's singular entities and municipalities. The indicators' primary components will be referred to them, though exceptionally to the provinces. By aggregation, we will calculate these indicators' primary components at the provincial level. On its side, the algorithms used to formulate the population dispersion indicators that we propose will yield provincial values. Once again, by aggregation, we will calculate the indicators at the regional and national levels. The network of singular entities provides greater granularity to address population dispersion, thus producing indicators that are more accurate. For this reason, we will analyse in a forthcoming Paper the association between SE-based and MUN-based indicators to gauge the possibility of focusing only on the latest ones in future updates of the indicators or when the calculation method is not feasible for SE-based indicators.

As for the population vector, we establish two leading elements for modelling population dispersion: the people and the locations where people reside. It gives rise to two approaches: "dispersion of people" versus "dispersion of locations." By creating this distinction, relevant issues surface concerning organizing the provision of FPS: On one hand, less dispersion of people will trigger economies of scale when providing FPS (including Reference Services, especially when centrality is high). On the other hand, even if the dispersion of people is less than the dispersion of locations, the need to guarantee universal access to FPS implies decreasing productivity in the supply of services to the population, depending on the province's zone, resulting in losses of economies of scale. Thus, and regarding decision-making, even if efficiency reasons would advise focussing on the dispersion of people, both people-based and location-based indicators should be combined for the set of indicators to be adequate from the perspective of equal access.

Proximity and centrality indicators are based on distances between SE and from SE to the capital of the province. The approach used to measure distances between municipalities is identical to the one used to measure distances between SE. Thus, indicators using distances as the primary component (as for proximity and centrality) will be formulated both SE-based and MUN-based. Normally, we will focus on the description of SE-based indicators, as it can be applied to MUN-based, *mutatis mutandis*. We measure distances using three criteria: straight-line distance, travel distance and travel duration. In addition, we distinguish both location-based (geographical) and population-based (population) proximity and centrality. Geographical proximity or centrality within a province derives from the simple average of distances between SE no matter their respective population ("**location distance**"). Population proximity or centrality within a province derives from the population-weighted average of distances between SE, thus giving more relevance to distances between more populated locations ("**population distance**").

For these two dimensions, we define three groups of indicators: absolute, relative and standardised.

Absolute indicators come from averages of distances between land uses; both simple averages (location distance) and weighted averages (population distance).

Relative indicators compare absolute population proximity to absolute geographical proximity (alternatively, centrality). They will be key, as both population and location-based indicators need to be jointly considered in the decision-making process to accommodate equal access requirements.

Standardised indicators allow for comparing provinces of different size and shape by normalising distances through the breadth of the province, thus providing dimensionless indicators. In order to build normalised indicators by rescaling distances via the breadth of the province, we define the adjusted diagonal of the province: the axes-aligned 2-dimensional bounding box's diagonal of each province, which we have then adjusted to improve inter-province comparability. In addition to rescaling distances via the province's diagonal, we have found the following algorithms to standardise proximity and centrality indicators that, in our view, will better capture the degree to which the population is spread across the territory:

- For proximity, we define the *Standardised Proximity Index (SPI)*. It uses the maximum attainable value of the spatial separation between land uses to normalise distances within a province. This maximum does not have a closed form solution, only approximations. In a region forming a perfect circle, the maximum value occurs when all the population is evenly distributed along the external edge. We have formulated this indicator based on distances between municipalities because, given the large number of SE in Spain, it would require such extensive calculation resources that, given the available ones, it is not feasible to work with SE.

- For centrality, we define:
    - The *Centralisation Ratio*, which compares the mean distance that population is located from the centre to the mean distance to the centre if population were uniformly distributed across the province with the same density in each municipality

    - The *Centralisation index (ACI),* which measures how rapidly land uses accumulate relative to land area as one moves progressively outward in concentric rings from the CBD.

Nuclearity indicators recourse to the number of nuclei in a province. A nucleus is a locus of intense residential development. For the purpose of this work, a nucleus will be an urban entity. We have defined it based on the widely used criterion in official statistical practice: any municipality with more than 10,000 inhabitants. In this work, both SE and municipalities with more than 10,000 inhabitants will be considered nuclei and we will formulate SE-based and MUN-based nuclearity indicators.

Density indicators are based on crude population density, which refers to municipalities. For provinces, we capture the average number of population per km^2 via the population-weighted average of the crude population densities of their municipalities. We use three ways for formulating the concept of land area referred to municipalities: total land area, urban land area, and built-up land area.

Concentration indicators use the variability of the population density across municipalities and concentration measures such as the Gini index and similar ones to gauge the extent to which the population is evenly distributed across land uses. To gauge this, once we have ordered the SE (or municipalities) in increasing order according to their population, if the accumulation of population goes in line with the accumulation of the number of land uses, then the population is evenly distributed. On the contrary, if it goes slower, the slower it goes the more concentrated would be that population. We will also use accumulation of land areas instead of accumulation of the number of land uses to juxtapose to the accumulation of population.

Finally, continuity indicators depend on crude density data. More specifically, on the degree to which municipalities' crude population density fits to an exponential pattern as a function of the distance from the centre of the province (CBD).

Primary components of the indicators for the dispersion's dimensions

Population in land uses

According to our objective, the observation units will be the people that use the land for residential purposes. We show the population included in this study in *Annex III. Table I.* In 2016, our base year, the population included was 46,206,955 inhabitants (99.25% of the total in Spain).

We will focus on two types of land uses: singular entities and municipalities. Our study includes 55,861 SE inhabited through 2003 to 2017,[18] which provides an outlook of the situation at given

[18] Initially, we built a database with base year 2016, meaning that the backbone of it was the SE inhabited in 2016. Please refer to the section on "Sources and databank." Afterwards, we have completed the pooled database for 2016 with the population data for the period 2017. As for this year, we have not excluded any SE; we just retain with zero population those SE having disappeared in 2017.

moments as well as the evolution over time. These 55,861 SE are in 8,102 municipalities (*Annex III. Table I*).

The province is our geographical unit of analysis: the algorithms used to formulate the population dispersion indicators that we propose yield provincial values. Our study includes 50 out of 52 provinces in Spain. We do not include Ceuta and Melilla. Ceuta has three SE (Benzú, Ceuta, and El Príncipe) at an average distance between them of 6.23 Km. Melilla has one SE. The population in both provinces represents a 0.37% of the total in 2016.

We emphasize that the most accurate indicators would be those based on singular entities. Our interest will be focused on these. Nonetheless, calculations based on distances between SE require more resources that those based on municipalities. For this reason, measuring indicators based on both types of land uses would provide an appraise on the extent to which, when needed, it would be possible to work with indicators based on average distance between municipalities within a province, instead of singular entities, without great loss of generality.

Calculation of distances between land uses

To measure the spatial separation between land uses within a province we use distances.[19] We use three ways for modelling the concept of distance: Straight-line distance, travel distance and travel duration.

Straight-line distance (km)

We calculate straight-line distances between any two land uses A and B (singular entities or municipalities) from their geographical coordinates (longitude and latitude).[20] Data on longitude and latitude are provided by the National Geographical Institute in European Terrestrial Reference System 1989 (ETRS89), according to the Royal Decree 1071/2007 of July 27, which regulates the geodetic system of official reference in Spain.

[19] No distance between two SE or municipalities of different provinces is involved in the calculations.

[20] For each singular entity, we define its centroid as its geographical coordinates provided by the National Geographical Institute (IGN). For each municipality, we define its centroid as the centroid of the singular entity that holds the capital as indicated by the IGN.

We compute straight-line distances with a basic formula for calculating spherical distances, the so-called Law of Cosines.[21] The equation is:

$$\boldsymbol{d} = R \times \mathrm{acos}\big(\cos(lat1) \times \cos(lat2) \times \cos(lon1 - lon2) + sin(lat1) \times sin(lat2)\big)$$

Where:

$\boldsymbol{R}$	$is\ the\ Earth's\ radius\ in\ Km = 6{,}378.137$
$\boldsymbol{lat1}$	$is\ the\ latitud\ of\ Population\ Unit\ A\ in\ radians$[22]
$\boldsymbol{lon1}$	$is\ the\ longitude\ of\ Population\ Unit\ A\ in\ radians$
$\boldsymbol{lat2}$	$is\ the\ latitud\ of\ Population\ Unit\ B\ in\ radians$
$\boldsymbol{lon2}$	$is\ the\ longitude\ of\ Population\ Unit\ B\ in\ radians$

With these calculations, we obtain the distance in *Km.* We notice that some authors consider additional metrics for the distance between two land uses, such as: 1/*d* (Torres, 2017); e^{-d} and *ln(d)* (Folch, 2012), which can be used and rescaled to fix the results to a particular range based on the research objectives (Campante and Do, 2009). In this work, we directly use straight-line distances.

For each province, we calculate two triangular matrices with straight-line distances: one for SE and another one for municipalities.[23] We call each distance in these matrices an "*observation point*". Please notice that in a given province i with $\#_i$ SE and μ_i municipalities the number of observation points is $\#_i(\#_i-1)/2$ for SE and $\mu_i(\mu_i-1)/2$ for municipalities (please refer to Annex II for the nomenclature). In total, we work with 128,783,590 observation points for SE and 841,372 for municipalities.

These triangular distance matrices are the inputs to calculate a number of proximity and centrality indicators. They provide basic descriptive statistics on the closeness between land uses within a province. We present some basic results in *Annex III. Table II.*

[21] This is the Law of Cosines equation, one of the versions of the Haversine formula for distances on spherical surfaces. We have selected it for its simplicity. It is documented that it is the simplest one. Nonetheless, it is also documented that due to an issue of computer crude, not mathematics, it is not "well-conditioned" for small distances (of the order of 1 meter or less). In our experience, we have found a negligible number of erroneous results when we applied the formula to calculate the distance between a location with itself. In addition, we have introduced internal coherence control and have identified a negligible number of locations for which distances calculated were inconsistent and therefore corrected manually or discarded.

[22] Thus, we have converted geographical coordinates from decimal degrees to radians multiplying by the constant $\pi/180$.

[23] We have developed the calculations by programming with Visual Basic. Including accesses from Excel to Bing Maps to obtein travel distances (Km) and travel durations (minutes).

According to our calculations, nationwide, the average straight-line distance between SE locations within the same province is 51.82 Km. It has a considerably high intra-province variability with coefficients of variation ranging from 48% in Cuenca to 101% in Las Palmas. If we focus on the provincial averages, the inter-province variability is lower though still high, with a coefficient of variation (CV) of 26% (Table 4 below).

Table 4. Simple average of straight-line distances between SE locations within the same province. National values (Km)	
NATIONAL AVERAGE	***51.82***
MAXIMUN PROVINCIAL AVERAGE	***86.19***
MINIMUN PROVINCIAL AVERAGE	***24.40***
STANDARD DEVIATION OF PROVINCIAL AVERAGES	***13.5***
CV OF PROVINCIAL AVERAGES	***26%***

Source: Annex III. Table II.

The minimum average straight-line distance between SE within a province is 24.40 Km, in Gipuzkoa, and the maximum one is 86.19 Km in Las Palmas de Gran Canaria. The maximum distance between SE within a province occurs in Illes Balears; it is 284.57 Km.

We notice that it is in the islands where we observe the maximum average distance between SE within a province as well as the maximum distance between SE within a province. This is caused by the spatial separation between islands. If we calculate only intra-island distances, discarding the distances between SE that are in different islands, even if within the same province, the average distance falls down considerably ("*island correction*"). The same applies to the maximum distance between SE.

Concerning the islands, two approaches are possible to develop population dispersion indicators: with or without the "*island correction*." We chose to formulate our indicators following the general rule of including all distances between SE within the same province, because distances between islands are a cost driver, especially regarding highly specialised centralised services.

Travel distance (km) and travel duration (minutes)

From the perspective of fundamental public services delivery, travel distance and travel duration from singular entities (or municipalities) to the points of service delivery are more relevant than straight-line distance. Therefore, we have modelled proximity and centrality indicators based on travel distance and travel duration, in addition to straight-line distance.

These data have been obtained from Bing Maps Geographical Information System (GIS), with travel distance in Km and travel duration in minutes. We have used Bing Maps Distance Matrix API through free access granted by Microsoft, which allowed for massive transactions from Excel to Bing Maps programmed with Visual Basic. We would have required the travel distance and duration for each one of the 128,783,590 observation points. This was not feasible for us. Thus, we have estimated average travel distances and travel durations at province level based on provincial random samples.

The sampling universe was the set of observation points. The sample size (n) was calculated for each province under simple random sampling for an infinite N(0,1) population to ensure a relative error (r) not greater than 1% at the level of confidence α = 95% (z = 2) with an anticipated estimation for the population CV of 30% within each province. Thus:[24]

$$n = \left(\frac{z * CV}{r}\right)^2$$

Where n is the theoretical sample size, r = 0.01 and CV = 0.30 are fixed a priori as said before. To better approach the objective that all SE in a province were included in at least one of the selected observation points, in some provinces, we enlarged the theoretical sample size by ensuring that the ratio of total SE to surveyed SE was at least one. The theoretical sample size (enlarged) was n = 196,000 and the effective sample size is actually ñ = 191,702.[25]

The CV of straight-line distances between SE within the same province resulted significantly higher than 30% in all provinces and so did sample estimates of the CV for travel distances and travel duration. Thus, producing estimation errors ranging from 1.05% to 3.07% for the sample average of travel distances, and 0.83% to 3.60% for the sample average of travel durations (*Annex III. Table III)*.

Estimation errors for the sample average were calculated as follows (in parts per unit):

[24] Cochran, W. G. (1984).

[25] We randomly selected rows and columns positions of the distances triangular matrices with the row number less than the column number. The results included diagonal and duplicate positions that were discarded.

$$\hat{r} = \frac{z \times \widehat{CV}}{\tilde{n}^{0.5}}$$

Where $\tilde{n}$ is the effective sample size and $\widehat{CV}$ the estimate of the coefficient of variation for each of the surveyed variables.

Taking advantage that we had both sample and population values for straight-line distance, we have used ratio estimators for improving the accuracy of travel distances and travel durations estimates at the provincial level. First, we have estimated two types of sample ratio in each province (i):

> ***Ratio 1 (i)***: Average travel distance / Average straight-line distance in province i ($\rho_1(i)$)
>
> ***Ratio 2 (i)***: Average travel duration / Average straight-line distance in province i ($\rho_2(i)$)

Subsequently, we have used them to estimate average travel distances and travel durations as follows:

$\boldsymbol{Travel\ distance\ (i)}$[26]

$$\psi(i) = Population\ average\ straight - line\ distance\ between\ SE\ in\ Province\ i\ \times Ratio\ 1\ (i)$$

$\boldsymbol{Travel\ duration\ (i)}$

$$\zeta(i) = Population\ average\ straight - line\ distance\ between\ SE\ in\ Province\ i\ \times Ratio\ 2\ (i)$$

We calculate relative errors for the sample ratios as well as for the ratio estimates of travel distances and travel durations as follows:

$$\hat{\hat{r}} = \frac{z\sigma_{\hat{\theta}_R}}{\hat{\theta}_R}$$

Where $\hat{\theta}_R$ is a generic name that represents both the estimator for the ratios ($\rho_1$and ρ_2) and the ratio estimator for travel distances (ψ) and travel durations (ζ); and $\sigma_{\hat{\theta}_R}$ represents its standard deviation. Indeed, following Cochran, W. G. (1984), for large samples, we can consider $\hat{\theta}_R$ an unbiased estimator with Normal distribution. Thus, the confidence interval for the corresponding population parameter would be $\theta_R \in (\hat{\theta}_R \pm z\sigma_{\hat{\theta}_R})$, where $r\theta_R = z\sigma_{\hat{\theta}_R}$. We summarise the nomenclature for the estimators we use and their estimation errors in Table 5 below.

[26] Please notice that in these formulas the term "population" is to indicate that the average has been calculated with all the observation points in the province instead of with the sample ones.

Table 5. Calculation of the relative errors of sample estimates for each province

Relative error for the sample average	$\hat{r} = \frac{z\,\widehat{CV}}{\tilde{n}^{0.5}}$			
$\widehat{CV}$ **Estimator of the coefficient of variation of each population surveyed**	**Variable**	**Sample estimator for the population average**	**Sample estimator for the population variance**	**Sample estimator for the population coefficient of variation ($\widehat{CV}$)**
	Straight-line distance (ξ)	Sample average: $\bar{\xi}$	Sample variance: $\hat{v}(\xi) = s_{\xi}^2$	$\widehat{cv}(\xi) = \frac{s_{\xi}}{\bar{\xi}}$
	Travel distance (ψ)	Sample average: $\bar{\psi}$	Sample variance: $\hat{v}(\psi) = s_{\psi}^2$	$\widehat{cv}(\psi) = \frac{s_{\psi}}{\bar{\psi}}$
	Travel duration (ζ)	Sample average: $\bar{\zeta}$	Sample variance: $\hat{v}(\zeta) = s_{\zeta}^2$	$\widehat{cv}(\zeta) = \frac{s_{\zeta}}{\bar{\zeta}}$

Relative error for the ratio estimators ($\hat{\theta}_R$)	$\hat{\hat{r}} = \frac{z\,\sigma_{\hat{\theta}_R}}{\hat{\theta}_R}$		
$\sigma_{\hat{\theta}_R}$ **Standard deviation of the ratio estimators**	**Variable**	**Ratio Estimator**	**Estimator's Variance**
	Ratio 1 ($\rho_1 = \frac{\psi}{\xi}$)	$\hat{\rho}_1 = \frac{\bar{\psi}}{\bar{\xi}}$	$\hat{v}(\hat{\rho}_1) = \hat{v}\left(\frac{\bar{\psi}}{\bar{\xi}}\right) = \frac{1}{\tilde{n}\bar{\Xi}^2}\left(s_{\psi}^2 + \hat{\rho}_1^2 s_{\xi}^2 - 2\hat{\rho}_1 s_{\psi\xi}\right)$
	Ratio 2 ($\rho_2 = \frac{\zeta}{\xi}$)	$\hat{\rho}_2 = \frac{\bar{\zeta}}{\bar{\xi}}$	$\hat{v}(\hat{\rho}_2) = \hat{v}\left(\frac{\bar{\zeta}}{\bar{\xi}}\right) = \frac{1}{\tilde{n}\bar{\Xi}^2}\left(s_{\zeta}^2 + \hat{\rho}_2^2 s_{\xi}^2 - 2\hat{\rho}_2 s_{\zeta\xi}\right)$
	Travel distance (ψ)	Ratio estimator: $\hat{\psi}_R = \hat{\rho}_1\,\bar{\Xi}$	$\hat{v}(\hat{\psi}_R) = \frac{1}{\tilde{n}}\left(s_{\psi}^2 + \hat{\rho}_1^2 s_{\xi}^2 - 2\hat{\rho}_1 s_{\psi\xi}\right)$
	Travel duration (ζ)	Ratio estimator: $\hat{\zeta}_R = \hat{\rho}_2\,\bar{\Xi}$	$\hat{v}(\hat{\zeta}_R) = \frac{1}{\tilde{n}}\left(s_{\zeta}^2 + \hat{\rho}_2^2 s_{\xi}^2 - 2\hat{\rho}_2 s_{\zeta\xi}\right)$

z =	2	Value of the distribution N (0,1) that leaves on its right a 0.025 probability (confidence level α = 0.95)

ñ =	Effective sample size
$\bar{\Xi}$ =	Average Straight-line distance calculated with the whole set of SE (whole population of observation points)

Source: Author's own work based on Cochran, 1984.

Estimation errors for the ratio estimators are significantly lower than that of the respective sample averages (*Annex III. Table IV)*: from 0.273% to 0.918% for travel distances and from 0.503% to 1.226% for travel durations.

Our results show that overall in Spain the average travel distance between any two SE within the same province is a 56% greater that the straight-line distance. Nationwide, we estimate an average travel distance of 80.72 Km. There is a high inter-province variability with a CV of 24% (Table 6).

In addition, we estimate a travel duration of 70.52 minutes (1.18 hours). It is remarkable that the inter-province variability for travel durations reaches a CV of 58% (Table 7). This is due to the high travel durations in the island territories far from the rest of provinces in the continental territory. In Canarias, this is compounded with the fact that the ratio travel distance to straight-line distance is the highest. Comparing with the results when considering the "*islands correction*," we see that the insular condition predominantly affects the duration of the journeys, which is significantly lower when we consider only intra-island displacements. Indeed, in Illes Balears, Las Palmas and Santa Cruz de Tenerife the average travel duration between SE when we consider only intra-island displacements is a 59% to 67% of the national average (*Annex III. Table IV*).

In our opinion, the two ratios that we have calculated ($\rho_1$and ρ_2) could be explored as proxy variables for measuring orography and one facet of insularity. Nonetheless, this is beyond the scope of this paper

Table 6. Simple average of travel distances between SE within the same province. National values (Km)	
NATIONAL AVERAGE	***80.72***
MAXIMUN PROVINCIAL AVERAGE	***134.34***
MINIMUN PROVINCIAL AVERAGE	***40.69***
STANDARD DEVIATION OF PROVINCIAL AVERAGES	***19.58***
CV OF PROVINCIAL AVERAGES	***24%***

Table 7. Simple average of travel duration between SE within the same province. National values (h)	
NATIONAL AVERAGE	***1.18***
MAXIMUN PROVINCIAL AVERAGE	***4.14***
MINIMUN PROVINCIAL AVERAGE	***0.61***
STANDARD DEVIATION OF PROVINCIAL AVERAGES	***0.68***
CV OF PROVINCIAL AVERAGES	***58%***

Source: Annex III. Table IV.

Breadth of the provinces

Absolute distances do not capture the extent to which singular entities and municipalities spread throughout the whole extension of the province. That extension is another cost driver that should be considered separately. Thus, it is meaningful to calculate normalised distances using a measure of the breadth of the province. This normalisation procedure makes comparisons of provinces of different shapes and sizes possible.

In the literature, we have found two possibilities to measure the breadth of the province for normalising purposes. One of them is using the square root of the surface area of the province. The other one is using the maximum spatial separation attainable between land uses within the province. The square root of the surface area of the province has the limitation that, for the islands, it does not represent properly the breadth of the provinces, and when used as a benchmark for normalising distances it produces overestimates. For that reason, we have substituted it by the **province's diagonal**, a proxy to the maximum attainable distance within the province. We detail below the definition and measurement of both the diagonal of the province and the **maximum spatial separation attainable between land uses** within the province.

Diagonal of the province

To build the province's diagonal, we start by calculating the geometric diagonal of the axes-aligned 2-dimensional bounding box of each province.[27] The axes-aligned 2-dimensional bounding box of each province (hereinafter bounding box or bounding rectangle) is the smallest axes-aligned rectangle that encloses the province' area, where the reference axes are the lines that run from north to south (vertical) and from east to west (horizontal). We use it as a proxy to the minimum 2-dimensional bounding box subject to no orientation constraints. Therefore, we have initially measured with Bing Maps the province maximum length from north to south and from east to west and calculated the mentioned geometric diagonal. We present the results in *Annex III. Table V*. Apart from the Islands, that hold the maximum values, the largest one is 295 Km in Zaragoza. The smallest one is 90 Km in Gipuzkoa.

[27] Following Ooi, B.C. (1993), the h-dimensional bounding boxes can be defined as a single dimensional array of k entries: (I_1, I_2 , ..., I_h) where I_s (s = 1 to h) is a closed bounded interval [a, b] describing the extent of the spatial object along dimension s.

To better approximate the concept of “maximum attainable distance within the province”, we adjust the mentioned diagonal controlling by the square root of the surface area of the province, the surface coverage in the bounding box and the actually attained maximum distance between SE. This improves inter-province comparability by addressing the effect of the gap between the surface area of each province and that of the axes-aligned bounding box. In addition, the adjustment method respects the inter-island’s distances reality.

To this end, we have calculated our province’s final diagonal as follows:

$$D_{adj} = \hat{d}_1 x_1 + \hat{d}_2 x_2 + \hat{d}_3 x_3 + \hat{b}$$

Where we have determine $\hat{d}_1, \hat{d}_2, \ \hat{d}_3 \ and \ \hat{b}$ based on the least squares criterion applied to the equation:

$$y = d_1 x_1 + d_2 x_2 + d_3 x_3 + b + \varepsilon$$

Where, for each province:

y = Geometric diagonal of the axes-aligned 2-dimensional bounding box (Km).

x_1 = Square root of the surface area (Km).

x_2 = Surface coverage within the axes-aligned 2-dimensional bounding box (in parts per unit).

x_3 = Maximum distance between SE within the province (Km).

As for peninsular provinces, we exclude the islands to determine $\hat{d}_1, \hat{d}_2, \ \hat{d}_3 \ and \ \hat{b}$. As for the islands, we include all provinces to determine these parameters so that we control by the gap between the surface area of each province and that of the axes-aligned bounding box, but we retain the effect of inter-island distances. We present the mentioned parameters in Figure 2.

Figure 2. Parameters estimates for the adjustment of D_{adj} 2016

Parameters	d_1	d_2	d_3	b	R^2
PENINSULAR STANDARD	1.57	-145.66	0.30	71.74	0.9863
ISLANDS STANDARD	1.24	-162.62	0.55	73.83	0.9780

Source: Authors’ own work based on the sources described in this paper

The adjusted $\boldsymbol{D_{adj}}$ (that we will name ***diagonal of the province*** hereinafter) is the final standard we propose to normalise distances for Proximity indicators. Our results are in *Annex III. Table*

V. Please notice that this benchmark is not suitable to standardise indicators based on travel duration because of lack of homogeneity with the magnitude of the indicators.

Overall, the average provincial diagonal in Spain is 203 Km. The diagonal of the provinces ranges from 91 Km in Gipuzkoa to 300 Km in Balears (295 in Badajoz, islands excluded). As for the surface area, the smallest is Gipuzkoa with 1,909 Km^2 (square root 44 Km) and the largest Badajoz with 21,766 Km^2 (square root 148 Km). The surface coverage in the bounding rectangle is 0.50 at national level; it ranges from 0.09 in Las Palmas to 0.70 in Ciudad Real. The maximum distance between SE within a province ranges from 77 Km in Gipuzkoa to 285 Km in Illes Balears (236 in Badajoz, islands excluded). (Figure 3).

Figure 3	Square root of the surface area of the province	Surface coverage in the bounding rectangle	Maximum distance between SE (Km)	Diagonal of the bounding rectangle*	Diagonal of the province D_{adj} ** (Km)*
TOTAL	**97.44**	**0.50**	**162.34**	**202.99**	**202.99**
Max SE	147.53	0.70	284.57	310.93	300.43
Min SE	43.69	0.09	76.64	89.81	91.05
Max SE	Badajoz	Ciudad Real	Balears	Balears	Balears
Min SE	Gipuzkoa	Palmas	Gipuzkoa	Gipuzkoa	Gipuzkoa
Excluding Islands					
Max SE	147.53	0.70	236.27	295.01	295.92
Min SE	43.69	0.40	76.64	89.81	91.05
Max SE peninsular	Badajoz	Ciudad Real	Badajoz	Zaragoza	Badajoz
Min SE peninsular	Gipuzkoa	Zaragoza	Gipuzkoa	Gipuzkoa	Gipuzkoa

Source: Authors' own work based on the sources described in this paper.
(*) TOTAL corresponds to the simple average of the diagonals of the provinces.
(**) Adjusted diagonal of the bounding rectangle.

Maximum distance attainable between land uses within a province

Following our findings in the literature review, we define the maximum distance attainable between land uses within a province as the maximum value attainable by the Spatial Separation Index or Venables (V_{max}).

The *Spatial Separation Index*[28] or *Venables Index* (*V*) is a population-based indicator that measures population separation instead of locations' separation. It is a sort of weighted average of the distances between land uses within the geographical unit, where each distance is given

[28] Originally proposed by Midelfart-Knarvik, K.H. et al. (2000; 2002). Please see Pereira et al., 2013 and 2015.

a different load depending on the population of the land uses, specifically the product of the two population weights. The system of products of the population weights of any two land uses is not a proper weights system as they don't sum 1. Nonetheless, we have retained the mentioned formulation in accordance to the literature because our final use of the Venables index is to build the *Standardised Proximity Index (SPI),* a normalised proximity indicator whose formulations using both the Venables index and that based on the proper system of population weights are equivalent.

Due to operational reasons, we formulate the *Venables Index* (*V*) only for municipalities. Below, we detail the formulation of the index for straight-line distance between municipalities within a province i (in Annex IV we provide further detail, including the formulations for travel distances and travel durations):

$$\boldsymbol{V(i)} = \tilde{s}\,'(i) \times T(i) \times \tilde{s}\,(i)$$

Where:

$\boldsymbol{T(i)}$ Is a triangular distance matrix whose entry d_{j1j2} is the straight-line distance between the municipalities j_1 and j_2 for under-diagonal entries and zero for the diagonal and upper-diagonal entries.

$\boldsymbol{\tilde{s}(i)}$ Is a column vector of population weights of the municipalities of a province:

$$\tilde{s}\,'(i) = \left(\frac{\pi_{i1}}{\pi_i}, \frac{\pi_{i2}}{\pi_i}, \ldots, \frac{\pi_{i\mu_i}}{\pi_i}\right)$$

The estimation of V_{max} is not trivial, because it has no closed-form solution. In a region forming a perfect circle, the maximum value (V_{max}) occurs when the population is evenly distributed along the external edge (Pereira, R.H.M. et al., 2013 and 2015): "*In a perfect circle shape, the maximum value of V would be obtained when all employment is evenly distributed along the external edge.*"… "*Thus, we have chosen to consider the 'opposite of maximum proximity' as a homogeneous distribution of values along the edge of the map. Although this solution is not the global maximum of the Venables index, it was considered a satisfactory solution for two reasons: a) intuitively, it is the opposite of a completely monocentric city with all employment in the center; and b) it is easy to calculate and does not require a specific algorithm. This normalization procedure makes it possible to compare urban areas of different shapes and sizes.*"

Therefore, an upper bound for the maximum spatial separation in a province would be the Spatial Separation Index when its municipalities are on the external edge of province's bounding circle. For the Spanish provinces, the percentage surface of that circle covered by the province's land area might differ considerably between provinces, depending on the shape of the province, and might produce lack of comparability.

To improve the comparability, we use the bounding box instead of the bounding circle. That is to say, we hold that an adequate proxy of the maximum spatial separation attainable is the Spatial Separation Index between the municipalities within a province when they are evenly distributed on the edge of the bounding box, where evenly means equidistant and with the same population (equal population weights). In Annex V, we provide some empirical evidence that would support our approach to approximate the maximum spatial separation attainable.

Our plan to estimate V_{max} is as follows.

— *First*, for each province i, we calculate an initial value of V_{max}:

$$VI_{max}(i) = \tilde{s}'(i) \times T_{max}(i) \times \tilde{s}(i)$$

Where $T_{max}(i)$ is a triangular distance matrix whose entry di_{j1j2}, $j_1 < j_2$, is the distance between municipalities j_1 and j_2 of province i after having distributed them evenly along the external edge of the axes-aligned 2-dimensional bounding box of province i, where evenly means equidistant and with the same population (equal population weights).

— *Second*, considering that:

i) Placing the population on the edge of the axes-aligned bounding box is different from placing it on the edge of the map of the province.

ii) The gap depends on the surface area of the province and the degree of fit of the shape of the province to the rectangle.

we have adjusted $VI_{max}(i)$ to control by the surface area of the province and its degree of fit to the rectangle, measured through the percentage of the axes-aligned bounding box 's surface covered by the land area of the province.

To this end, we have calculated our final V_{max} as follows:

$$V_{max} = \hat{m}_1 x_1 + \hat{m}_2 x_2 + \hat{m}_3 x_3 + \hat{b}$$

Where we have determined $\hat{m}_1, \hat{m}_2, \ \hat{m}_3 \ and \ \hat{b}$ based on the least squares criterion applied to the following equation:

$$y = m_1 x_1 + m_2 x_2 + m_3 x_3 + b + \varepsilon$$

Where, for each province:

y = $VI_{max}(i)$ (Km)

x_1 = Square root of the surface area of the province (Km)

x_2 = Surface coverage within the axes-aligned bounding box (in parts per unit)

x_3 = Ratio of the simple average of straight-lines distance to V (dimensionless)

The rationale behind is that $VI_{max}(i)$ is driven by the axes-aligned bounding box extension (captured by x_1), the shape of the province (captured by x_2), and the population loads given to each distance between any two municipalities within the province (captured by x_3).

As for peninsular provinces, we exclude the islands to determine $\hat{m}_1, \ \hat{m}_2, \ \hat{m}_3 \ and \ \hat{b}$. As for the islands, we include all provinces to determine these parameters so that we control by the gap between the surface area of each province and that of the axes-aligned bounding box, but we retain the effect of inter-island distances. We present the mentioned parameters in Figure 4.

The adjusted V_{max} (that we will name V_{max} hereinafter) is the final standard we propose to build the Standardised Proximity Index. Our results for V, VI_{max} and V_{max} are in Annex III. Table V.

We observe that when we give different loads to the distances between any two municipalities depending on their populations, thus moving from location distance to a proxy of population distance, the resulting separation decreases. Indeed, the average distance between the municipalities' locations within a province is 53.58 Km in 2016 while the Venable index, which is a proxy for population distance, is 17.70 Km. This indicates that typically, within a Spanish province, the population tends to concentrate in municipalities that are closer to each other than the entire set of municipalities.

Figure 4. Parameters for the adjustment of V_{max}

PENINSULAR STANDARD

Parameter / Year	m1	m2	m3	b	R^2
2003	0.4574	-9.3769	1.7243	4.1744	0.9292
2004	0.4570	-9.3502	1.7422	4.1418	0.9294
2005	0.4567	-9.3495	1.7453	4.1572	0.9292
2006	0.4563	-9.3297	1.7486	4.1657	0.9291
2007	0.4559	-9.3232	1.7542	4.1873	0.9290
2008	0.4556	-9.3023	1.7491	4.2075	0.9288
2009	0.4553	-9.2426	1.7408	4.2185	0.9287
2010	0.4551	-9.2129	1.7427	4.2091	0.9287
2011	0.4549	-9.1847	1.7447	4.2019	0.9287
2012	0.4547	-9.1495	1.7377	4.2142	0.9288
2013	0.4544	-9.1193	1.7281	4.2399	0.9288
2014	0.4544	-9.1210	1.7221	4.2362	0.9286
2015	0.4543	-9.0925	1.7128	4.2528	0.9287
2016	0.4541	-9.0908	1.6964	4.2964	0.9286
2017	0.4540	-9.0672	1.6848	4.3135	0.9286

ISLANDS STANDARD

Parameter / Year	m1	m2	m3	b	R^2
2003	0.5441	-78.1816	-0.4724	38.8150	0.9265
2004	0.5442	-78.1702	-0.4733	38.8037	0.9265
2005	0.5442	-78.1402	-0.4786	38.8025	0.9265
2006	0.5442	-78.1100	-0.4774	38.7805	0.9265
2007	0.5443	-78.0936	-0.4766	38.7622	0.9265
2008	0.5443	-78.0691	-0.4787	38.7564	0.9265
2009	0.5444	-78.0542	-0.4756	38.7389	0.9265
2010	0.5444	-78.0443	-0.4691	38.7138	0.9264
2011	0.5444	-78.0439	-0.4690	38.7102	0.9264
2012	0.5445	-78.0397	-0.4678	38.7031	0.9265
2013	0.5446	-78.0327	-0.4672	38.6953	0.9265
2014	0.5445	-78.0273	-0.4697	38.7081	0.9265
2015	0.5446	-78.0295	-0.4701	38.7092	0.9265
2016	0.5446	-78.0261	-0.4698	38.7086	0.9265
2017	0.5446	-78.0241	-0.4727	38.7179	0.9266

Source: Authors' own work based on the sources described in this paper

Nuclei of the Spain's provinces

A nucleus is a locus of intense residential development. For the purpose of this work, a nucleus will be an urban entity: singular entities (or, municipalities) with 10,001 or more inhabitants.[29] Typically, the widely used criterion in official statistical practice is to classify as urban any

[29] The traditional Spanish statistical classification determines as urban "*the set of population entities with 10,001 or more inhabitants.*" Please refer to Ministry of Development of Spain (2018). Áreas urbanas en España 2018. Constitución, Cuarenta años de las ciudades españolas. Retrieved from: https://apps.fomento.gob.es/CVP/handlers/pdfhandler.ashx?idpub=BAW058. Please refer also to Reig, E. et al. (2016). Retrieved from: https://www.fbbva.es/wp-content/uploads/2017/05/dat/DE_2016_IVIE_delimitacion_areas_rurales.pdf.

municipality with 10,001 or more inhabitants. However, for the purpose of this work, we have particularised it for SE and have defined nuclei both SE-based and MUN-based. In *Annex III. Table VII* we show the number of nuclei by province in Spain both SE-based and MUN-based. In 2016, there were 731 urban municipalities and 673 urban SE, ranging from 1 nucleus in Ávila, Palencia, Segovia and Soria to 72 (80) in Barcelona.

Land area of Spain's municipalities (Km^2)

Land area data refer to municipalities; they are not available for SE. We obtain the surface area of each province by aggregation. We use three ways for modelling the concept of ***land area***: ***total*** land area, ***urban*** land area, and ***built-up*** land area. As for the total land area, the source is the National Geographical Institute (IGN). For the rest, the official definitions and data come from the Cadastral Services (CS) of Spain. The urban land area corresponds to that for cadastral purposes, in accordance with the cadastral legislation: article 7.2 of the Consolidated Text of the Real Estate Cadastre Law.[30] The cadastral statistics subdivide the area of urban parcels into the area of built-up parcels and the area of unbuilt parcels. We work with the first two items: urban and built-up.

The National Cadastral Services do not provide data for Navarra and País Vasco because of their specific fiscal system. For these regions, we have retrieved the data from the Urban Information System (SIU) published by the Ministry of Development (MD).[31] In the case of Navarra, the SIU provides only information for 10% of the municipalities. We have completed them with information concerning urban land (υ) as reported by Navarra's regional Statistical Office (Nastat)[32] and using regression techniques to estimate the built-up land area (β) in each municipality of Navarra as a function of urban land as follows:[33]

[30] Please refer to: http://www.catastro.meh.es/documentos/estadisticas_Metodologia_Catastro_2015.pdf; http://www.catastro.minhap.gob.es/documentos/estadisticas/Nota%20metodol%C3%B3gica%20Estadistica%20ocupaci%C3%B3n.pdf; and http://www.catastro.minhap.gob.es/documentos/estadisticas_Metodologia_Catastro_2015.pdf.

[31] Retrieved from: https://apps.fomento.gob.es/CVP/handlers/pdfhandler.ashx?idpub=BAW055.

[32] Retrieved from: https://administracionelectronica.navarra.es/GN.InstitutoEstadistica.Web/DescargaFichero.aspx?Fichero=\web\agregados\2_espacio_fisico\21_espafis_territorio\espafis_territorio_mun_ext.xls.

[33] We notice that from Nastat we have only found data on urban land but not on built-up land. http://www.eustat.eus/bankupx/pxweb/es/spanish/-/PX_3951_super01.px#axzz6220fFzbX

1. First, we adjust urban land data from the SIU with those from Nastat (concerning the 10% of the municipalities with available data from the SIU). For these municipalities we obtain:

$$\widehat{N} = 1.0854N$$

$$1.0854 = \frac{Total\ urban\ land\ for\ the\ 10\%\ of\ municipalities\ as\ reported\ by\ Nastat}{Total\ urban\ land\ for\ the\ 10\%\ of\ municipalities\ as\ reported\ by\ the\ SIU}$$

$N = Total\ urban\ land\ for\ the\ 10\%\ of\ municipalities\ as\ reported\ by\ the\ SIU$

$\widehat{N} = Total\ urban\ land\ for\ the\ 10\%\ of\ municipalities\ as\ reported\ by\ Nastat$

2. Second, we adjust built-up land data from the SIU to keep invariant the percentage of built-up land over urban land as registered by the SIU (concerning the 10% of the municipalities with available data from the SIU). For these municipalities we get:

$$\hat{B} = \frac{B}{N}\widehat{N}$$

$B = Total\ built - up\ land\ for\ the\ 10\%\ of\ municipalities\ as\ reported\ by\ the\ SIU$

$\hat{B} = Total\ built - up\ land\ for\ the\ 10\%\ of\ municipalities\ as\ reported\ by\ Nastat$

3. Third, with the adjusted data for the mentioned 10% of Navarra's municipalities $(\widehat{N}, \hat{B})$, outliers excluded, we calculate the statistical association between built-up land and urban land conditioned to be a line through the origin. We obtain:

$$\widehat{\mathrm{B}} = 0.564\widehat{N}$$

4. Fourth, we calculate built-up land area for all of Navarra's municipalities through the equation:

$$\widehat{\beta} = 0.564\upsilon$$

In *Annex III. Table VI* we present the data about land area referred to 2016.

Crude population density in Spain's municipalities

The direct ratio population to land area: ***crude population density*** is a primary measure of population density that has been widely used as an indicator of population dispersion.

To capture the extent to which the territory of Spain's provinces is thickly populated, we use municipalities' crude population density as the primary measure, both for total, urban and built-up area, and then aggregate at the provincial, regional and national level using population-weighted averages.[34]

We avoid focusing our final analysis on crude population density for geographical levels higher than municipalities because crude population density, at that aggregation level, would hide relevant information. Indeed, relating population to total land area, we see that, overall, in Spain, total crude population density is 92 inhabitants per Km^2, below the EU average of 118 (*Annex III. Table VIII* and EUROSTAT (a)). However, the analysis by Rae, A. (2018) on "*Population Density in Europe*" shows that "*much of Spain appears to be empty; much more so than any other large European country… Yet characterising Spain as a sparsely populated country does not reflect the experience on the ground … So even though the settlement pattern appears sparse, people are actually quite tightly packed together*."[35]

In brief, with the available data, to capture in the best way the extent to which Spain's provinces are thickly populated throughout their territory, we use crude population density as primary component computed at the lowest level available of geographical breakdown. Then, we build more elaborated computations at the provincial level via population-weighted means, and by using three approaches for the concept of density: based on total land area (total density), on urban area (urban density) and on built-up area (residential density).

Municipalities' crude population density would also provide a first raw glance to concentration through its standard deviation, which according to our calculation is high. *Annex I. Table VIII*.

Sources and databank

To develop our indicators, we have built a database of singular entities for each of Spain's provinces with the following basic variables:

[34] Please notice that we use the term "crude" to distinguish, **at provincial, regional and national** level, the population-weighted from the surface-weighted average of the density of the municipalities (at municipal level density and crude density coincide). Therefore, we highlight that "crude" in this paper is not for distinguishing the ratio total population divided by the total area from that without subtracting areas devoted to open space, roadways, parks or similar public use and infrastructure areas. To address the latest features we work with the Spain's legal concepts on Cadastral definitions.

[35] Rae, A. (2018). Please refer also to Reig, E. et al. (2016).

INE code of the province

INE code of the municipality

INE code of the singular entity

Name of the singular entity

Name of the province

Latitude

Longitude

Population by age groups.

Concerning municipalities, we have built a database including information on land area: [36]

INE code of the province

INE code of the municipality

INE code of the ES holding the capital of the municipality

Name of the municipality

Name of the capital

Distance to CBD

Total land area

Urban land area

Built-up land area

Population by age groups.

To this end, we have matched the following sources:

INE: Continuous Municipal Register with information detail at the level of singular entities. It includes inhabited SE with their number of inhabitants for the period 2003-2017 by single ages. It has been provided by the INE under petition.

IGN: Geographic Register of singular entities, 2017.
Retrieved from https://www.ign.es/web/ign/portal in June 2018

IGN: Geographic Register of municipalities, 2017.
Retrieved from https://www.ign.es/web/ign/portal in August 2018

[36] Please notice that in each Municipality there is a singular entity that holds the capital. This SE is the centroid of the municipality, which is used for the calculation of distances between municipalities. Therefore, within the SE database we have all the information that we need on distances between municipalities.

CS Urban Real Estate Cadastre Statistics. Cadastral main variables: Urban surfaces plots. Retrieved from: http://www.catastro.minhap.gob.es/esp/estadistica_1.asp

MD Urban Information System (SIU). Retrieved from: https://apps.fomento.gob.es/CVP/handlers/pdfhandler.ashx?idpub=BAW055

EUSTAT www.eustat.eus http://www.eustat.eus/bankupx/pxweb/es/spanish/-/PX_3951_super01.px#axzz6220fFzbX

NASTAT https://www.navarra.es/home_es/Gobierno+de+Navarra/Organigrama/Los+departamentos/Economia+y+Hacienda/Organigrama/Estructura+Organica/Instituto+Estadistica/

We started with the INE's Continuous Municipal Register for 2016 (base year) which includes 58,358 inhabited SE with 46.557.008 inhabitants. We have added the information on latitude and longitude. Then, we continued working with provinces individually.

During this process, we found some incidences. First, for some SE in the INE's database we found no information in the IGN's database; so no location information was available (type I incidence). *In Annex III. Table IX*, we summarised the magnitude of this sort of incidences. The population excluded from the analysis for this reason is a 0.015 % of the total.

Second, we found SE in the IGN's database with no information on latitude and longitude: value zero (type II incidence). In *Annex III. Table X*, we show the details for this sort of incidence. The population excluded from the analysis for this reason is a 0.032% of the total.

Third, we have completed the pooled database for 2016 with the population data for the period 2003-2015 and 2017. Thus, we got an integrated database for the period 2003-2017. In doing so we have also found incidences because some SE were not population entities during the whole period (type III incidence). In *Annex III. Table XI*, we show the magnitude of this sort of incidence. The population excluded from the analysis for this reason is a 0.338 % of the total.

As for 2017, we have not excluded any SE; exceptionally, we only retain with zero population those having disappeared in 2017.

Finally, we added straight-line distances between SE and municipalities to the provincial databases building triangular matrices with straight-line distances. We have excluded 13 SE because the calculated straight-line distance from them to the rest of all SE in their province was disproportionately high in relation to the size of the territory (type IV incidence, affecting 13 SE and 0.0007% of the population)[37]. Please refer to *Annex III. Table XII*.

Therefore, we work with a databank of provincial databases with 55,861 inhabited SE covering a 99.25% of Spain's population in 2016 and their respective municipalities' databases. *Annex III. Table XIII* presents the details by province of the SE included and excluded in our analysis.

Formulation of indicators

For each of the six dimensions integrating dispersion, we have identified a set of indicators. We describe below the specific indicators that we have selected and provide some keys for their interpretation. Typically, we have adjusted the indicator's algorithm to ensure that low values point out high rates of dispersion. As a leading rule, we have used simple algorithm, as close as possible to intuition; considering that the ultimate objective is to facilitate their application in the decision-making process.

For some of the indicators, we have found in the literature specific analyses regarding desirable properties that such indicators ought to display. This sort of analysis is beyond the scope of this paper. We have only referred to some of these characteristics occasionally to further support the suitability of a number of indicators. The nomenclature that we use for the basic elements employed in formulating the indicators is found in Annex II. The technical details concerning the indicator's algorithm are located in Annex IV.

[37] As already said, we calculate distances with the Law of Cosines equation, one of the versions of the Haversine formula for distances on spherical surfaces. We have selected it for its simplicity. It is documented that it is the simplest one. Nonetheless, it is also documented that due to an issue of computer crude, not mathematics, it is not "well-conditioned" for small distances (of the order of 1 meter or less). In our experience, we have found a negligible number of erroneous results when we applied the formula to calculate the distance between a location with itself. In addition, we have introduced internal coherence control and have identified a negligible number of locations for which distances calculated were inconsistent and therefore corrected manually or discarded (affecting 13 SE and 0.0007% of the population).

Proximity indicators

The set of indicators that we use captures proximity within province i through the spatial separation between land uses: SE and municipalities within the province. To measure spatial separation we use ***distances***. The lower the distance the higher the proximity (and the lower the dispersion). We adjust the indicator's algorithm by calculating the ***inverse of the distance*** to ensure that low values of the proximity indicators point out high rates of dispersion. Please, notice that we always refer to distances between any two land uses within the same province: No distance between two SE and municipalities of different provinces is involved in the calculations. We list the proximity indicators that we will explore in Box 1.

Box 1. List of proximity indicators[38]

1. *Inverse of the simple average of the straight-line distances between SE ($PROXS_{SE1a}$)*
2. *Inverse of the simple average of travel distances between SE ($PROXS_{SE1b}$)*
3. *Inverse of the simple average of travel durations between SE ($PROXS_{SE1c}$)*
4. *Inverse of the weighted[39] average of straight-line distances between SE ($PROXW_{SE1d}$)*
5. *Inverse of the weighted average of travel distances between SE ($PROXW_{SE1e}$)*
6. *Inverse of the weighted average of travel durations between SE ($PROXW_{SE1f}$)*
7. *Ratio of population proximity to geographical proximity (SE & straight-line distance) ($PROXR_{SE1g}$)*
8. *Ratio of population proximity to geographical proximity (SE & travel distance) ($PROXR_{SE1h}$)*
9. *Ratio of population proximity to geographical proximity (SE & travel duration) ($PROXR_{SE1i}$)*
10. *Normalised geographical proximity (SE & straight-line distance) ($PROXN_{SE1j}$)*
11. *Normalised geographical proximity (SE & travel distance) ($PROXN_{SE1k}$)*
12. *Normalised population proximity (SE & straight-line distance) ($PROXN_{SE1l}$)*
13. *Normalised population proximity (SE & travel distance) ($PROXN_{SE1m}$)*
14. *Inverse of the simple average of straight-line distances between municipalities ($PROXS_{MUN2a}$)*
15. *Inverse of the simple average of travel distances between municipalities ($PROXS_{MUN2b}$)*
16. *Inverse of the simple average of travel durations between municipalities ($PROXS_{MUN2c}$)*
17. *Inverse of the weighted average of straight-line distances between municipalities ($PROXW_{MUN2d}$)*
18. *Inverse of the weighted average of travel distances between municipalities ($PROXW_{MUN2e}$)*
19. *Inverse of the weighted average of travel durations between municipalities ($PROXW_{MUN2f}$)*
20. *Ratio population proximity to geographical proximity (municipalities & straight-line distance) ($PROXR_{MUN2g}$)*
21. *Ratio of population proximity to geographical proximity (municipalities & travel distance) ($PROXR_{MUN2h}$)*
22. *Ratio of population proximity to geographical proximity (municipalities & travel duration) ($PROXR_{MUN2i}$)*
23. *Normalised geographical proximity (MUN & straight-line distance) ($PROXN_{MUN1j}$)*
24. *Normalised geographical proximity (MUN & travel distance) ($PROXN_{MUN1k}$)*
25. *Normalised population proximity (MUN & straight-line distance) ($PROXN_{MUN1l}$)*
26. *Normalised population proximity (MUN & travel distance) ($PROXN_{MUN1m}$)*
27. *Standardised Proximity Index (SPI) based on straight-line distance) ($PROXV_{MUN2n}$)*
28. *Standardised Proximity Index (SPI) based on travel distance ($PROXV_{MUN2o}$)*
29. *Standardised Proximity Index (SPI) based on travel duration ($PROXV_{MUN2p}$).*

In principle, most accurate indicators would be those based on distances between SE. Nonetheless, calculations based on distances between SE require more resources than those based on municipalities. For this reason, we will analyse in a forthcoming working paper the association between SE-based and municipalities-based indicators to gauge the possibility of

[38] Indicators based on distances between singular entities have SE in the subscript and those based on distances between municipalities have MUN in the subscript.
[39] We weight the distance between two different singular entities of province i by the product of their respective populations.

focusing only on municipalities' distances in future updates of the indicators or when the calculation method is not feasible for SE-based indicators.

For each province, we calculate the average distance between any two land uses[40] through simple and weighted averages (population weights). The simple average yields indicators associated with the geographical proximity (***$PROXS_{SE1a}$, $PROXS_{SE1b}$, $PROXS_{SE1c}$, $PROXS_{MUN2a}$, $PROXS_{MUN2b}$,*** and ***$PROXW_{MUN2c}$***). On the contrary, indicators based on weighted average reflect population proximity (***$PROXS_{SE1d}$, $PROXS_{SE1e}$, $PROXS_{SE1f}$, $PROXS_{MUN2d}$, $PROXS_{MUN2e}$,*** and ***$PROXW_{MUN2f}$***). Both simple and weighted averages are indicators of proximity in absolute terms.

The ratio between the two types of indicators is a relative magnitude that compares the proximity of the population with the proximity of the places where they live (***$PROXR_{SE1g}$, $PROXR_{SE1h}$, $PROXR_{SE1i}$, $PROXR_{MUN2g}$, $PROXR_{MUN2h}$,*** and ***$PROXR_{MUN2i}$***).

Absolute and relative indicators do not capture the extent to which locations or population spread throughout the whole territory of the province. A way to overcome this limitation is to calculate standardised indicators. In the literature, we have found two possibilities for normalising. The first is using the square root of the surface area of the province, and the second is to calculate the maximum average distance attainable between land uses when they are distributed in a way that maximises the distances between them. These normalisation procedures make possible the comparisons of provinces of different shapes and sizes.

As already explained, the square root of the surface area of the province has the limitation that, for the island territories, it does not represent properly the breadth of the provinces, and when used as a benchmark for normalising distances it produces proximity underestimations. For that reason, we have substituted it by the diagonal of the province as explained. Please notice that we do not standardise indicators based on travel duration because of lack of homogeneity between the magnitude of the indicator and the available benchmarks. Standardised indicators based on the diagonal of the province (***$PROXS_{SE1j}$, $PROXS_{SE1k}$, $PROXS_{SE1l}$, $PROXS_{SE1m}$, $PROXS_{MUN2j}$, $PROXS_{MUN2k}$,***

[40] We notice that some authors consider also the median. The advantage of the median is that it is more robust than the average so less influenced by outliers. Nonetheless, we use the average for simplicity and ease of interpretation.

$PROXS_{MUN2l}$, and $PROXW_{MUN2m}$) relate the average distance between land uses to a comparable magnitude approximating the maximum straight-line distance attainable within the province.

The indicator relating the distance between land uses and the maximum attainable one, the *Standardised Proximity Index*, is based on the *Spatial Separation Index* or *Venables index*. We calculate it based on straight–line distances (***$PROXV_{MUN2n}$***), travel distances (***$PROXV_{MUN2o}$***), and travel durations (***$PROXV_{MUN2p}$***). We notice that, by construction, this indicator standardises the population distance in each province in relation to its size, as the maximum attainable depends only on that size and it is unrelated to the number of land uses (see Annex V).

Centrality Indicators

The set of indicators we have chosen captures centrality within province i in two ways. First, through the ***distances*** between SE and municipalities to the centre of that province (CBD). We define the CBD as the capital of the province, which needs not to be its geographical centre. The same as for proximity, we compute the inverse of the distance to CBD so that the lower the centrality the higher the dispersion. Second, through the accumulation of land uses and their population around the CBD. In this case, we use ***land areas*** to gauge the extent to which the CBD accumulates population in relation to the accumulation of land area and the way in which accumulation of people around the CBD is faster than that of land area. The slower the accumulation, the higher the dispersion. We list the centrality indicators that we will explore in Box 2.

For each province, we calculate the average distance from land uses to the CBD through simple and weighted averages (population weights). The simple average yields indicators associated with geographical centrality (***$CBDdS_{SE3a}$, $CBDdS_{SE3b}$, $CBDdS_{SE3c}$, $CBDdS_{MUN4a}$, $CBDdS_{MUN4b}$, and $CBDdS_{MUN4c}$***). On the other hand, indicators based on weighted average reflect population centrality (***$CBDdW_{SE3d}$, $CBDdW_{SE3e}$, $CBDdW_{SE3f}$, $CBDdW_{MUN4d}$, $CBDdW_{MUN4e}$, and $CBDdW_{MUN4f}$***). As for the geographical centrality we don’t include the CBD in the calculations. On the contrary, we do include it for population centrality. Both simple and weighted averages are indicators of centrality in absolute terms.
The ratio between the two types of indicators is a relative magnitude that relates the centrality of the population to the centrality of the places where they live (***$CBDdR_{SE1g}$, $CBDdR_{SE1h}$, $CBDdR_{SE1i}$, $CBDdR_{MUN2g}$, $CBDdR_{MUN2h}$, and $CBDdR_{MUN2i}$***).

Box 2. List of centrality indicators[41]

1. *Inverse of the simple average of the straight-line distances from SE to CBD ($CBDdS_{SE3a}$)*
2. *Inverse of the simple average of the travel distances from SE to CBD ($CBDdS_{SE3b}$)*
3. *Inverse of the simple average of the travel durations from SE to CBD ($CBDdS_{SE3c}$)*
4. *Inverse of the weighted average of the straight-line distances from SE to CBD ($CBDdW_{SE3d}$)*
5. *Inverse of the weighted average of the travel distances from SE to CBD ($CBDdW_{SE3e}$)*
6. *Inverse of the weighted average of the travel durations from SE to CBD ($CBDdW_{SE3f}$)*
7. *Ratio population centrality to geographical centrality (SE & straight-line distance) ($CBDdR_{SE3g}$)*
8. *Ratio population centrality to geographical centrality (SE & travel distance) ($CBDdR_{SE3h}$)*
9. *Ratio population centrality to geographical centrality (SE & travel duration) ($CBDdR_{SE3i}$)*
10. *Normalised geographical centrality (SE & straight-line distance) ($CBDdN_{SE3j}$)*
11. *Normalised geographical centrality (SE & travel distance) ($CBDdN_{SE3k}$)*
12. *Normalised population centrality (SE & straight-line distance ($CBDdN_{SE3l}$)*
13. *Normalised population centrality (SE & travel distance) ($CBDdN_{SE3m}$)*
14. *Inverse of the simple average of the straight-line distances from municipalities to CBD ($CBDdS_{MUN4a}$)*
15. *Inverse of the simple average of the travel distances from SE municipalities to CBD ($CBDdS_{MUN4b}$)*
16. *Inverse of the simple average of the travel durations from municipalities to CBD ($CBDdS_{MUN4c}$)*
17. *Inverse of the weighted average of the straight-line distances from municipalities to CBD ($CBDdW_{MUN4d}$)*
18. *Inverse of the weighted average of the travel distances from municipalities to CBD ($CBDdW_{MUN4e}$)*
19. *Inverse of the weighted average of the travel durations from municipalities to CBD ($CBDdW_{MUN4f}$)*
20. *Ratio population centrality to geographical centrality (MUN & straight-line distance) ($CBDdR_{MUN3g}$)*
21. *Ratio population centrality to geographical centrality (MUN & travel distance) ($CBDdR_{MUN3h}$)*
22. *Ratio population centrality to geographical centrality (MUN & travel duration) ($CBDdR_{MUN3i}$)*
23. *Normalised geographical centrality (MUN & straight-line distance) ($CBDdN_{MUN3j}$)*
24. *Normalised geographical centrality (MUN & travel distance) ($CBDdN_{MUN3k}$)*
25. *Normalised population centrality (MUN & straight-line distance) ($CBDdN_{MUN3l}$)*
26. *Normalised population centrality (MUN & travel distance) ($CBDdN_{MUN3m}$)*
27. *Centralisation ratio ($CBDdCR_{MUN4n}$)*
28. *Centralisation index ($CBDdACI_{MUN4o}$).*

To build *normalised centrality indicators*, as already explained, we have chosen to use the diagonal of the province. These indicators (***$CBDdN_{SE1j}$, $CBDdN_{SE1k}$, $CBDdN_{SE1l}$, $CBDdN_{SE1m}$, $CBDdN_{MUN2j}$, $CBDdN_{MUN2k}$, $CBDdN_{MUN2l}$***, and ***$CBDdN_{MUN2m}$***) relate the average distance from land uses to the CBD with a comparable magnitude measuring the span of the province.

In addition to the mentioned indicators, we have found in the literature indicators that provide alternative ways of benchmarking. Though not standardised by the breadth of the province, they would provide some information on the extent to which population spreads around the centre and they are independent of the province's size.

The *Centralisation Ratio* (***$CBDdCR_{MUN4n}$***, only for municipalities) compares the mean distance population is located from the centre to the mean distance to the centre if population were uniformly distributed across the province with the same density in each municipality. When population is uniformly distributed across the province with the same density in each municipality the Centralisation Ratio is zero, pointing out high dispersion. If the population

[41] Exceptionally, we will propose some indicators whose reference is the province, our geographical unit of analysis, they have PROV in the subscript.

moves far away from a uniform distribution, the centralisation ratio diminishes pointing out high dispersion. Moreover, if population is actually more decentralised than a uniform distribution, the centralisation ratio can be negative. On the contrary, should the population reside in just one land use, the *Centralisation Ratio* would be one.

The *Centralisation Index* (***$CBDdACI_{MUN4o}$***, only for municipalities) computes the accumulation of population from the CBD compared to the corresponding accumulation of land area. The slower the population accumulation the higher the dispersion.

We would like to notice that some authors consider the density gradient indicator as an indicator for centrality (please refer to the point on the concentration indicators' formulation). Following Ottensmann, J.R. (2017 (a), (b)),[42] we have not included the density gradient as an indicator of centrality. We have included it as a concentration indicator.

Nuclearity Indicators

To formulate ***nuclearity*** indicators we have designated nuclei for each province: loci of intense residential development. For the purpose of this work, nuclei will be an urban entity (singular entities –or municipalities- with 10,001 or more inhabitants, according to the traditional Spanish statistical classification). The set of indicators that we use captures nuclearity within province i through the degree of mononuclearity: the lower the mononuclearity, the higher the dispersion. Our first indicator for nuclearity is the inverse of the number of nuclei, and thus we ensure that low values of the indicator point out high rates of dispersion. The percentage of population in the CBD over the whole set of nuclei is also a measure of mononuclearity. We use it as our second nuclearity indicator. The lower the share of the CBD in the total population of nuclei (thus the lower the nuclearity), the higher the dispersion. Nuclearity is maximised when the province has a mononuclear pattern of residential development: the CBD is the only nucleus. When the number of nuclei increases, nuclearity decreases; this would point out more dispersion.

We list the nuclearity indicators that we will explore in Box 3.

[42] *"The attraction of using the negative exponential model parameters to examine decentralization is clear. For any metropolitan area, the decentralization of population or housing units will result in a decrease in the negative exponential density gradient (all other things being equal). However, the truth of this statement does not necessarily imply the converse, that the density gradient can therefore be used as a reasonable measure of centralization. The relationship between the parameters for the negative exponential decline of density and the level of centralization is more complex." "However, the level of the density gradient is not significantly related to the level of centralization…"*

Box 3. List of nuclearity indicators

1. *Inverse of the number of nuclei SE-based (NUNoN$_{SE5a}$)*
2. *Share of the population in the CBD over the population in nuclei SE-based (NUSoP$_{SE5b}$)*
3. *Inverse of the number of nuclei MUN-based (NUNoN$_{MUN6a}$)*
4. *Share of the population in the CBD over the population in nuclei MUN-based (NUSoP$_{MUN6b}$).*

Density Indicators

The set of indicators that we use captures density within province i through the ***crude population densities*** of its municipalities.[43]

We use crude population density as the primary component, computed at the lowest level available of geographical breakdown and using three approaches to the concept of density: based on total land area, on urban area and on built-up area. Then, via population-weighted means, we have constructed more elaborate measures at the provincial level.

Thus, we propose three approaches to define population density at the provincial level. The first one captures the average number of residential units per km^2 for the total land area; we will refer to it as "***total density***." The second one is for the urban land area; we will refer to it as "***urban density***." The third one is for the built-up land area; we will refer to it as "***residential density***." The list of density indicators that we will explore at the provincial level is in Box 4.

Box 4. List of density indicators

1. *Population-weighted density based on total land* (DEPWD$_{MUN7a}$)
2. *Population-weighted density based on urban land* (DEPWD$_{MUN7b}$)
3. *Population-weighted density based on built-up land area* (DEPWD$_{MUN7c}$)
4. *Maximum density based on total land* (DENMAX$_{MUN7d}$)
5. *Maximum density based on urban land* (DENMAX$_{MUN7e}$)
6. *Maximum density based on built-up land area* (DENMAX$_{MUN7f}$)
7. *Minimum density based on total land* (DENMIN$_{MUN7g}$)
8. *Minimum density based on urban land* (DENMIN$_{MUN7h}$)
9. *Minimum density based on built-up land area* (DENMIN$_{MUN7i}$)
10. *Share of the population living in high density municipalities based on total land* (DENHIGH$_{MUN7j}$)
11. *Share of the population living in high density municipalities based on urban land* (DENHIGH$_{MUN7k}$)
12. *Share of the population living in high density municipalities based on built-up land area* (DENHIGH$_{MUN7l}$)
13. *Density of land use in the CBD based on total land* (DENCBD$_{MUN7m}$)
14. *Density of land use in the CBD based on urban land* (DENCBD$_{MUN7n}$)
15. *Density of land use in the CBD based on built-up land area* (DENCBD$_{MUN7o}$).

We would like to note that some authors (especially in the field of urban sprawl) consider the density gradient indicator as an indicator for population density. The rationale behind this indicator is to capture the rate of attenuation of the population density with growing distance from the CBD. We have chosen to associate the mentioned indicator to the degree of compactness of the population around the CBD and use it as a measure of concentration.

[43] We have not land area for SE

Concentration Indicators

To measure concentration we focus on the distribution of the population across the singular entities and municipalities. Especially, on its variability and on the extent to which a small number of locations concentrates a high share of the population. To this end, we rely on a set of indicators that measure any form of variability in population density and uneven or dissimilar distribution of the population. We list the concentration indicators that will explore in Box 5.

Box 5. List of concentration indicators

1. *Gini index for SE (CNGINI$_{SE8a}$)*
2. *Standardised Theil entropy index (SE) (CNSTHEI$_{SE8b}$)*
3. *Standardised Herfindahl index (SE) (CNSHHI$_{SE8c}$)*
4. *Coefficient of variation of densities (CNDCV$_{MUN9a}$)*
5. *Share of the population living in high density municipalities based on built land (CNHGDMUN9b)*
6. *Population density gradient (CNPDG$_{MUN9c}$)*
7. *Gini index for MUN based on population (CNGINI$_{MUN9d}$)*
8. *Gini index for MUN based on land areas (CNGINI$_{MUN9e}$)*
9. *Standardised Theil entropy index (MUN) (CNSTHEI$_{MUN9f}$)*
10. *Theil index (CNTHI$_{MUN9g}$)*
11. *Standardised Herfindahl index (MUN) (CNSHHI$_{MUN9h}$)*
12. *Raw geographic concentration index (CNRGCI$_{MUN9i}$)*
13. *Ellison and Glaesser (CNEG$_{MUN9j}$)*
14. *Delta index (also Hoover index) (CNDI$_{MUN9k}$)*
15. *Massey and Denton dissimilarity index for urban land (CNMDDI$_{MUN9}$l)*
16. *Massey and Denton dissimilarity index for built-up land (CNMDDI$_{MUN9m}$).*

We have borrowed indicators from several fields of analysis: "*variability of the population density*" (Coefficient of variation of densities, Share of the population living in high residential density municipalities, Population density gradient), income distribution (i.e. Gini or Theil indices), economic concentration (Herfindahl or Ellison and Glaeser indices) and social spatial segregation (Delta index or Massey and Deaton indices). Each of them reflects different facets of population concentration.

The "*Coefficient of variation of densities"* measures the variability of the population density across municipalities in relative terms of the mean value of that variable. The "*Share of population living in high residential density municipalities"* measures the population percentage of the most densely populated municipalities. The "*Population density gradient*" is the rate at which density falls from the centre. A high value means that density will decline sharply with increasing distance from the capital of the province, thus pointing out concentration in the CBD.

We calculate the "*Population density gradient*" (ϕ) through the exponential density function $\delta_{ij}^{0}(d_{ij}) = \delta_{CBD}e^{-\phi d_{ij}\varepsilon}$ applying OLS regression.

The rest of indicators have been designed to approach inequality or dissimilarity (or alternatively, evenness). Typically, this sort of indices are dimensionless with low values reflecting low concentration or equivalently high dispersion. *Gini indices* range between 0 and 1 the extent to which population concentrates in a few number of locations. *Theil indices* are based on the idea that "order" (the index equals 1) is associated with the concentration of the bulk of the population in only one location (maximum concentration) while "disorder" (the index equals 0) is associated with an even distribution of the population across the locations (high entropy; dispersion). Indices coming from the field of economic activity concentration (*Herfindahl index* -also Herfindahl-Hirschman index-, *Raw geographic concentration* index and *Ellison and Glaesser index*) measure the degree to which the SE population shares (alternatively those of municipalities) are equally distributed. Finally, fairly used in social spatial segregation, the family of dissimilarity indices (*Delta index* –or Hoover- index and *Massey and Denton index*) provide additional measures of the evenness with which two variables (population and surface) are distributed across municipalities.

We highlight that the calculation spatial unit appears as a key factor to study the dynamics of concentration. Indeed, the literature review has shown that generally the degree of concentration increases with the size of the chosen spatial units. The fact that concentration measures are sensitive to the size of the calculation geographical units is related to the Modifiable Area Problem (MAUP) and has been approached by some authors in the field of economic concentration (Bertinelli, L. et al. (2005)).

Typically, when data are grouped, the indices are sensitive to the definition and the number of categories used. The integration of two or more categories always implies a reduction of the calculated value of the index; unless the two of them have the same population share (aggregation implies erasing part of the differences) (Rainis R. et al. (2003)).

In Spain, some analyses in the field of economic concentration have indicated that the use of the province does not provide significant data due to the differences in size and spatial distribution of population (Santa María, M.J. et al. (2005)). There seems to be a certain consensus about the choice of local units as the most appropriate. That is the case in this work. We use local units: singular entity and municipality to calculate provincial indicators.

Continuity Indicators

To measure continuity we use two indicators based on crude density data. First, we use the ratio urban or built-up land area to total land area.[44] The lower the ratio the lower the continuity and the greater the dispersion. Second, we measure continuity through the degree to which municipalities' crude population density (δ_{ij}^{0}) fits to an exponential pattern as a function of the distance from the centre of the province, the CBD, (d_{ij}). The determination coefficient R^2 of the below exponential density function provides an indicator of continuity.[45]

$$\ln \delta_{ij}^{0}(d_{ij}) = \delta_0 - \phi d_{ij} + \varepsilon \qquad \text{[exponential density function]}$$

Where:

δ_{ij}^{0}	Crude population density of municipality j in province i (total land-based)
d_{ij}	Distance to the CBD of municipality j in province i
ϕ	Population density gradient
δ_0	CBD density

The lower the R^2 the lower the continuity and the greater the dispersion. We list the concentration indicators that we will explore in Box 6.

Box 6. List of continuity indicators

1.	*Ratio urban land area to total land area ($CNTRUT_{PROV10a}$)*
2.	*Ratio built-up land area to total land area ($CNTRUT_{PROV10b}$)*
3.	*R-square of the exponential density function ($CNTR2_{PROV10c}$).*

[44] Adapted from Ghandi, S.R. et al. (2016).
[45] Malpezzi, S. et al. (2001); Tsai, Y.H. (2005).

5. SUMMARY AND CONCLUSIONS

Population dispersion is one of the spending drivers in fundamental public services: education, health and essential social services (FPS), and thus influences the sustainability of public finances. Geographical areas where population is highly dispersed would need to offer services at higher rates of intensity of resources, to ensure equal access. As a driver of public expenditure, it has not yet been explored in Spain as much as other drivers have been, such as population ageing, with which it interacts.

We have developed this work in the context of the analysis of budgetary stability in Spain. Considering the de facto federal structure of Spain, which provides that FPS are mainly managed by Spain's Regions, the sustainability of fundamental public services at the national level is determined by the ability of Regional governments to comply with fiscal stability requirements.

We claim that the sustainability of public spending requires disruptive innovative solutions[46] to address the provision of essential public services in geographical areas with high population dispersion. The first step to integrate population dispersion into decision-making processes would be to ensure the availability of valid indicators, to provide evidence-based choices. An objective of this paper is to present a methodology through which valid indicators are used to quantify and measure population dispersion in Spain's Regions. A further objective is to provide a flexible tool for policy decision-making, integrating population dispersion in the analysis of the sustainability of FPS in Spain.

Our analyses show that population dispersion is a multidimensional concept and, therefore, its measurement requires a multidimensional approach. Following a literature review, we define population dispersion as a specific pattern of land use by the population for residential purposes that is represented by low values in one or more of six distinct dimensions: Proximity, Centrality, Nuclearity, Density, Concentration and Continuity.

[46] https://ec.europa.eu/health/sites/health/files/expert_panel/docs/012_disruptive_innovation_en.pdf

The literature review has provided us with definitions and algorithms for calculating indicators for these six dimensions. We have designed the algorithms to work with a bottom up approach. We start with a set of primary components to build the indicators. We use them to calculate indicators for provinces and then aggregate to the regional and national levels.

To measure population dispersion, we have introduced a set of ninety-four indicators grouped in six dimensions and two categories, depending on the basic cell that we set as calculation spatial unit: singular entity (SE) or municipality (MUN). We have selected the simplest possible indicators among those most frequently used in the literature, which we understand as a guarantee of compliance with basic suitability criteria, placing greater priority on practical concerns, such as ease of use. For some of them we have found in the literature analyses regarding desirable properties that such indicators ought to display. We have only referred to these characteristics occasionally to further support the suitability of a number of indicators. Analysing in depth these desirable properties goes beyond the scope of this paper.

In the literature review, we have found one indicator on population dispersion that is already used in Spain. Specifically, in the resource allocation model of the Spanish regions. In this context, the indicator selected for modelling population dispersion is the number of singular population entities. We have not included this indicator. In our view, the number of singular entities existing in each Region in a given year is not a suitable indicator to capture the additional costs of providing fundamental public services because of population dispersion. Main reasons are listed next. The indicator includes all SE, even if they are not inhabited. It does not take into account some relevant associated cost drivers, such as the distance between land uses, the distance to the capital of the province, and the extent to which most of the population is concentrated in locations closer to each other than the entire set of locations. The indicator is not normalised by the size of the province, though said size is one of the cost drivers specifically considered in the regional resource allocation model. In addition, we have verified through simulation techniques that the maximum spatial separation attainable by population entities within a province is unrelated to the number of entities, but only to the size of the province itself.

Quantifying and analysing the selected ninety-four indicators will be the task for a forthcoming paper, in application of the designed methodology. Here we focus on presenting the algorithms and the primary components to build the indicators, together with some quantitative data concerning these primary components, namely Population in land uses; distances between any two land uses within the same province; breadth of the province; nuclei; area of land uses (municipalities); and crude population density (municipalities).

We have organised the primary components to build indicators in a databank of provincial databases with both singular entities and municipalities as basic geographic cells. Our study includes 55,861 SE inhabited through 2003 to 2017 (base year = 2016) in 50 provinces, which provides an outlook of the situation at given moments as well as the evolution over time. These 55,861 SE are in 8,102 municipalities.

We summarise below the key basic features extracted from the quantification of the primary components that we have defined to build population dispersion indicators.

The average straight-line distance between SE within the same province in Spain is 51.82 Km (location distance). Nonetheless, moving between two locations is not always possible in a straight-line. The average travel distance within a province in Spain is 80.72 Km and the average travel duration is 70.52 minutes or 1.18 hours.

There is a significant variability among provinces, with a coefficient of variation of 26% for straight-line distance, 24% for travel distance and 57% for travel duration. The magnitude of the latter is due to the high travel durations registered in inter-island displacements. When this effect is corrected, the coefficient of variation for inter-province travel durations is 21%. In general, when the "*island correction*" is considered, the impact on national averages is negligible. Nonetheless, it is significant at province level in the island territories.

The observed high level of inter-province variability in the distances between locations is connected to the significant differences in the size of Spain's provinces. The diagonal of the provinces ranges from 91 Km in Gipuzkoa to 300 Km in Balears (295 in Badajoz, islands excluded). The maximum distance between SE within a province ranges from 77 Km in Gipuzkoa to 285 Km in Balears (236 in Badajoz, islands excluded). As for the surface area, the smallest is Gipuzkoa with 1,909 Km^2 and the largest Badajoz with 21,766 Km^2.

When quantifying the provinces' breath, we have verified that population distance in Spain's provinces is below location distance. Indeed, the average distance between the municipalities' locations within a province is 53.58 Km in 2016 while the Venable index, which is a proxy for population distance, is 17.70 Km. This indicates that typically, within a Spanish province, the population tends to concentrate in municipalities that are closer to each other than the entire set of municipalities. Thus, both population and location-based indicators need to be jointly considered in the decision-making process to accommodate equal access requirements.

Dispersion would be greater when the population recourses to several nodes of intense residential development (nuclei) to spread over the territory. As for nuclearity, we have defined singular entities (or, municipalities) with 10,001 or more inhabitants as nodes of intense residential development. In Spain, there were 673 in 2016 (alternatively, 731 for municipalities). Ranging from 1 nucleus in Ávila, Palencia, Segovia and Soria to 72 (80) in Barcelona.

Relating population to the extension of the territory, we see that, overall, in Spain, the ratio population to surface area (crude total density) is 92 inhabitants per Km^2 (below the EU average of 118). Typically, crude total density has been used as an indicator of population dispersion, associating low densities to high population dispersion. Nonetheless, as said, density is only one dimension of dispersion. Even more, as for density, we would need more fine-tuned indicators than crude population density, to capture properly the extent to which Spain is thickly populated throughout its territory.

Indeed, if we analyse population density in more detail, focusing on the built-up area, we can see that Spain registers a density of built-up area (the share of built-up area in total land area) of 1.36% in 2016, well below the EU average, according to Eurostat and OECD data. This would show that "*much of Spain appears to be empty; much more so than any other large European country… Yet characterising Spain as a sparsely populated country does not reflect the experience on the ground … So even though the settlement pattern appears sparse, people are actually quite tightly packed together*" Rae, A. (2018). This would further support approaching dispersion from several angles, in this case: "total density," "urban density," and "residential density;" as well as the relevance of measuring "population distance" together with "location distance." The variability of crude density provides a first raw glance to concentration. According to our calculation, it is high in Spain. Especially for crude total density (CV of 1.55).

We can point out that the basic components that measure different aspects of dispersion present a high variability among provinces. We lack international benchmarks except for population density and built-up area. However, the analysis of the inter-province variability together with the national averages as benchmarks can provide an insight of the population dispersion in Spain. We expect that that variability confronted to that of spending on FPS would provide some clues on the association between both magnitudes.

To carry out such an analysis, this work has provided us with valuables conclusions to guide next steps:

1. Population dispersion is a multidimensional concept, and using in isolation one individual dimension to capture dispersion (such as, for instance, density, which has been widely used to measure dispersion, not to say the number of singular population entities) may leads to rather different conclusions than when a more balanced definition is adopted
2. A good practice in approaching population dispersion measurement would be building a composite indicator capturing each of the six dimensions of proximity, centrality, nuclearity, density, concentration and continuity

3. Quite some indicators have been already used to measure population dispersion through its dimensions. Most relevant ones should be adapted to the Spanish regions and analysed, to construct a good composite indicator
4. In the process of selecting the most relevant indicators, at least the following criteria should be considered:
    a. To involve information both on population and locations dispersion, for the set of indicators to accommodate equal access requirements
    b. To involve both information on total, urban and residential density
    c. To be independent of the breadth of the provinces in order avoid confounding factors
    d. To use whenever possible indicators based on SE
    e. To prioritise indicators based on travel distances or travel durations where appropriate.

ANNEX I. LIST OF PAPERS SELECTED IN THE LITERATURE REVIEW

If the link to the paper does not work, please copy and paste the URL into the internet navigator

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
2018 (a)	OECD	Divided Cities Understanding Intra-urban Inequalities	OECD Publishing, Paris. https://www.oecd.org/publications/divided-cities-9789264300385-en.htm https://read.oecd-ilibrary.org/urban-rural-and-regional-development/divided-cities_9789264300385-en#page1
2018 (b)	OECD	Rethinking Urban Sprawl: Moving Towards Sustainable Cities	OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264189881-en https://read.oecd-ilibrary.org/environment/rethinking-urban-sprawl_9789264189881-en#page1
2018	Rae, A.	Think your country is crowded? These maps reveal the truth about population density across Europe.	The Conversation. https://theconversation.com/think-your-country-is-crowded-these-maps-reveal-the-truth-about-population-density-across-europe-90345
2017	EUROPEAN UNION (EU)	Background documents for the European Semester Measuring Competitiveness	DG Internal Market, Industry, Entrepreneurship and SMEs, FWC "Studies in the Area of European Competitiveness." Centre for European Economic Research (ZEW), Germany Austrian Institute of Economic Research (WIFO), Austria https://ec.europa.eu/docsroom/documents/28181/attachments/1/translations/en/renditions/pdf
2017	Gude, A. et al.	Heterogeneous spillovers among Spanish provinces: A generalized spatial stochastic frontier model	ECONOMIC DISCUSSION PAPERS. Departamento de Economía. Universidad de Oviedo https://www.unioviedo.es/oeg/ESP/esp_2017_03.pdf http://economia.uniovi.es/investigacion/papers
2017	Khoirunurrofik	Trends and determinants of the geographical distribution of economic activities: evidence from Indonesian manufacturing.	Journal of Indonesia Applied Economics, Vol.7 No.1, 2017: 18-47 jiae.ub.ac.id/index.php/jiae/article/download/190/175 http://www.google.es/url?sa=t&rct=j&q=&esrc=s&source=web&cd=312&cad=rja&uact=8&ved=0ahUKEwjryoLZh4jcAhXP2KQKHQhjCpg4tgIQFgqxMAE&url=http%3A%2F%2Fjiae.ub.ac.id%2Findex.php%2Fjiae%2Farticle%2Fdownload%2F190%2F175&usg=AOvVaw19wCvBCJnaJKeRjMVVS_42
2017	Maslikhina, V.	Spatial concentration of the manufacturing industry: evidence from Russia.	Volga State University of Technology, Russia. Original Scientific Paper. Paper number: 15(2017)4, 481, 509 – 517. http://dx.doi.org/10.5937/jaes15-14961 http://www.engineeringscience.rs/images/pdf/14961.pdf
2017	Ministry of Finances of Spain	Comisión de expertos para la revisión del modelo de financiación autonómica	https://www.hacienda.gob.es/CDI/sist%20financiacion%20y%20deuda/informaci%C3%B3nccaa/informe_final_comisi%C3%B3n_reforma_sfa.pdf
2017 (a)	Ottensmann, J.R.	The Degree of Centralization in Large Urban Areas in the U.S., 1950-2010.	Urban Patterns. Observations and research on urban structure https://urbanpatternsblog.files.wordpress.com/2017/01/urban-area-centralization.pdf
2017 (b)	Ottensmann, J.R.	Negative Exponential Model Parameters and Centralization in Large Urban Areas in the U.S., 1950-2010.	Urban Patterns. Observations and research on urban structure and Indiana University-Purdue University Indianapolis. https://urbanpatternsblog.files.wordpress.com/2017/02/negative-exponential-centralization.pdf https://www.researchgate.net/profile/John_Ottensmann/publication/313516071_Negative_Exponential_Model_Parameters_and_Centralization_in_Large_Urban_Areas_in_the_US_1950-2010/links/589d0d14aca272e6cd48fd73/Negative-Exponential-Model-Parameters-and-Centralization-in-Large-Urban-Areas-in-the-US-1950-2010.pdf
2017	Tian, S. et al.	Exploring Industrial Agglomeration of Manufacturing Industries in Shanghai using Duranton and Overman's K-density Function.	The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLII-2/W7, 2017. ISPRS Geospatial Week 2017, 18–22 September 2017, Wuhan, China https://www.int-arch-photogramm-remote-sens-spatial-inf-sci.net/XLII-2-W7/149/2017/isprs-archives-XLII-2-W7-149-2017.pdf
2017	Torres, T.	Ensayos sobre Economías de Aglomeración.	Universidad de Barcelona http://diposit.ub.edu/dspace/bitstream/2445/118777/1/TPTG_TESIS.pdf
2016	Anaya-Arenas, A.	Models for a Fair Humanitarian Relief Distribution.	Laval University of Quebec http://www.fsa.ulaval.ca/sirul/2016-003.pdf
2016	Centurion, I et al.	Desigualdad regional y concentración en el Uruguay.	Dirección de Descentralización e Inversión Pública Oficina de Planeamiento y Presupuesto http://www.otu.opp.gub.uy/sites/default/files/docsBiblioteca/Reporte%205_Desigualdad%20PIBR.pdf
2016	Gandhi, SR. et al.	Quantifying Urban Sprawl for Rajkot City using Geospatial Technology.	IJBES 3(2)/2016, 86-92 https://www.researchgate.net/publication/309443048_Quantifying_Urban_Sprawl_for_Rajkot_City_using_Geospatial_Technology/fulltext/58d27b2d92851cf4f8f5e6de/Quantifying-Urban-Sprawl-for-Rajkot-City-using-Geospatial-Technology.pdf https://www.researchgate.net/publication/309443048_Quantifying_Urban_Sprawl_for_Rajkot_City_using_Geospatial_Technology
2016	Kavanagh, L. et al.	Is Poverty Decentralizing.	Annals of the American Association of Geographers

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
			http://eprints.whiterose.ac.uk/102896/7/Is%20Poverty%20Decentralizing%20Quantifying%20Uncertainty%20in%20the%20Decentralization%20of%20Urban%20Poverty.pdf https://www.tandfonline.com/doi/full/10.1080/24694452.2016.1213156
2016	Martori, J.C. et al.	Real state bubble and urban population density: six Spanish metropolitan areas 2001-2011	The Annals of regional Science volume 56, pages369–392. https://link.springer.com/article/10.1007%2Fs00168-016-0743-z https://www.researchgate.net/profile/Joan-Carles-Martori/publication/291690116_Real_estate_bubble_and_urban_population_density_six_Spanish_metropolitan_areas_2001-2011/links/5c616bbd92851c48a9c994b3/Real-estate-bubble-and-urban-population-density-six-Spanish-metropolitan-areas-2001-2011.pdf
2016	OECD	regions at a Glance 2016	OECD https://www.oecd-ilibrary.org/docserver/reg_glance-2016-en.pdf?expires=1615804598&id=id&accname=guest&checksum=8E96D76B66E3D566414E31B3798BCE22
2016	Reig, E. et al.	Delimitación de áreas rurales y urbanas a nivel local. Demografía, coberturas de suelo y accesibilidad.	Fundación BBVA https://www.fbbva.es/wp-content/uploads/2017/05/dat/DE_2016_IVIE_delimitacion_areas_rurales.pdf
2016	Sobrino, J.	Localización industrial y concentración geográfica en México.	Estudios demográficos y urbanos, VOL. 31, NÚM. 1 (91), 2016, pp. 9-56 http://www.scielo.org.mx/pdf/educm/v31n1/0186-7210-educm-31-01-00009.pdf
2016	Stepler, R. et al.	U.S. Latino Population Growth and Dispersion.	Pew Research Centre http://assets.pewresearch.org/wp-content/uploads/sites/7/2016/09/PH_2016.09.08_Geography.pdf
2016	Van Egeraat, C. et al.	A Measure for identifying substantial geographic concentrations.	Papers in regional Science https://core.ac.uk/download/pdf/297027339.pdf
2015	Allen, R. et al.	More reliable inference for the dissimilarity index of segregation	Econometrics Journal (2015), volume 18, pp. 40–66. DOI: 10.1111/ectj.12039 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5054828/pdf/ECTJ-18-40.pdf
2015	Alvarez, C. et al.	Índice de dispersión poblacional distrital (IDP) para la estimación de necesidades de recursos humanos en salud del primer nivel de atención	*An Fac Med. 2015;76:41-8* http://revistasinvestigacion.unmsm.edu.pe/index.php/anales/article/view/10969/9902
2015	Dauth, W. et al.	Long-run processes of geographical concentration and dispersion Evidence from Germany	IAB Discussion Paper 27/2015. Articles on labour market issues http://doku.iab.de/discussionpapers/2015/dp2715.pdf
2015	De la Fuente, A.	El impacto de los factores geográficos sobre las necesidades de gasto autonómicas: una aproximación empírica	FEDEA http://pareto.uab.es/wp/2015/95715.pdf
2015	FEDEA	Presentation for the workshop on regional financing	FEDEA http://www.fedea.net/docs/hacienda-autonomica/WRFR_propuesta_medida_dispersion.pdf
2015	Goerlich, F.J. et al.	Estimaciones de la población rural y urbana a nivel municipal	Estadística Española · Volumen 57, número 186 / 2015, pp. 5-28 http://observatoriopoblacion.aragon.es/wp-content/uploads/2018/11/Estimaciones-de-la-poblaci%C3%B3n-rural-y-urbana-a-nivel-municipal.pdf
2015	Junta de Castilla y León	Propuesta de medida de la dispersión.	FEDEA. Workshop sobre la reforma del sistema de financiación de las comunidades autónomas de régimen común, 20 de noviembre de 2015. https://www.fedea.net/docs/hacienda-autonomica/WRFR_propuesta_medida_dispersion.pdf https://www.fedea.net/workshop-regional/
2015	Lee, S.	The role of urban spatial structure in reducing VMT and GHG emissions	University of Illinois at Urbana-Champaign, 2015 https://www.ideals.illinois.edu/bitstream/handle/2142/88966/LEE-DISSERTATION-2015.pdf?sequence=1
2015	Pereira, R.H.M. et al.	Quantifying urban centrality: a simple index proposal and international comparison	Ipea 189 Discussion Paper http://www.ipea.gov.br/portal/images/stories/PDFs/TDs/ingles/dp_189.pdf
2015	Rastvortseva, S. et al.	Regional specialization and geographical concentration of industry in Russia	regional Science Inquiry, Vol. VII, (2), 2015, pp. 97-106 http://www.rsijournal.eu/ARTICLES/December_2015/9.pdf https://www.researchgate.net/profile/Svetlana-Rastvortseva/publication/301799374_Regional_specialization_and_geographical_concentration_of_industry_in_Russia/links/5728bd1c08aef7c7e2c0c3c2/Regional-specialization-and-geographical-concentration-of-industry-in-Russia.pdf

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
			https://www.researchgate.net/profile/Svetlana-Rastvortseva/publication/273145600_Analyses_of_regional_specialization_and_geographical_concentration_of_industry_in_Russia/links/54fabd170cf23e66f032b4ef/Analyses-of-regional-specialization-and-geographical-concentration-of-industry-in-Russia.pdf
2015	Reyes, R.J. et al.	Concentración Geográfica de las Patentes como Indicador de Innovación Tecnológica en México en los años 2009 – 2014	XXXIX Conference ANPAD, 2015 https://www.researchgate.net/profile/Porto_Geciane/publication/282672542_Concentracion_Geografica_de_las_Patentes_como_Indicador_de_Innovacion_Tecnologica_en_Mexico_en_los_anos_2009_-_2014/links/56180df908ae78721f9a8d4e/Concentracion-Geografica-de-las-Patentes-como-Indicador-de-Innovacion-Tecnologica-en-Mexico-en-los-anos-2009-2014.pdf?origin=publication_detail
2015	Van Egeraat, C. et al.	A Measure for identifying substantial geographic concentrations	MPRA Paper No. 65954, posted 14. August 2015 04:56 UTC. https://mpra.ub.uni-muenchen.de/65954/1/MPRA_paper_65954.pdf
2014	Bertaud, A. et al.	The Spatial Distribution of Population in 57 World Cities: The Role of Markets, Planning, and Topography	Coller School of Management. Tel Aviv University https://coller.tau.ac.il/sites/nihul.tau.ac.il/files/media_server/Recanati/management/elrov/june2014_symposium/Malpezzi.pdf
2014	Boontore, A.	Sprawl and distance travelled	Centre for Advanced Spatial Analysis. University college London, April 2014 http://discovery.ucl.ac.uk/1426738/1/Boontore.pdf
2014	Cerina, F. et al.	Network communities within and across borders	Scientific Reports volume 4, Article number: 4546 (2014) https://www.nature.com/articles/srep04546.pdf
2014	Dijkstra, L. et al.	A harmonised definition of cities and rural areas: the new degree of urbanisation.	Regional Working Paper 2014. WP 01/2014. European Commission Directorate-General for regional and Urban Policy (DG REGIO). https://ec.europa.eu/regional_policy/sources/docgener/work/2014_01_new_urban.pdf
2014	Eisenberg, J.	Team member distance and innovative team performance - the influence of leadership styles and team dynamics	State University of New Jersey. https://rucore.libraries.rutgers.edu/rutgers-lib/45621/PDF/1/play/
2014	Fernández, M. et al.	Factores explicativos de la evolución reciente de la distribución de la población en Galicia	Universidad de Zaragoza. International Conference on regional Science https://old.reunionesdeestudiosregionales.org/Zaragoza2014/htdocs/pdf/p1191.pdf
2014	Lemelin, A. et al.	Measuring urban agglomeration. A refoundation of the mean city-population size index	Documentos de Trabajo N.º 1430. Banco de España. https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/14/Fich/dt1430e.pdf
2014	Nuñez, G.	Medición multidimensional de la concentración de población	Universidad Autónoma de Chiapas. México https://espacioimasd.unach.mx/index.php/Inicio/article/download/39/107/ https://www.researchgate.net/publication/330746206_Medicion_Multidimensional_de_la_Concentracion_de_Poblacion/fulltext/5c52567192851c22a39d2e9d/Medicion-Multidimensional-de-la-Concentracion-de-Poblacion.pdf
2014	Prat, J.M. et al.	Análisis de la distribución geográfica, a nivel comarcal, de los alojamientos turísticos del pirineo catalán mediante la aplicación de algunos índices de concentración geográfica	Revista de análisis turístico, nº 17, 1º semestre 2014, pp. 27-38 https://revistas.ucm.es/index.php/AGUC/article/download/45196/42549/ https://analisis-turistico.aecit.org/index.php/AECIT/article/view/174 https://dialnet.uniRioja.es/servlet/articulo?codigo=4784129
2014	Prat, J.M. et al.	Análisis de la evolución de la concentración geográfica de los establecimientos de turismo rural en Cataluña	Anales de Geografía 2014, vol. 34, núm. 1 155-177. ISSN: 0211-9803 http://dx.doi.org/10.5209/rev_AGUC.2014.v34.n1.45196
2014	Ranasinghe, J.	Research into Drivers of Service Costs in Rural Areas Rapid Evidence Assessment - Literature Review	GOV.UK developer docs https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/388598/Rural_literature_review.pdf
2014	Rodríguez, J.M. et al.	Dispersion indicator in relation to the services	Jornadas de Estadística de las Comunidades Autónomas http://www.jecas.es/2014_Asturias/ponencias/H1.pdf http://www.jecas.es/2014_Asturias/descargap.html
2014	Santic, D.	Spatial concentration of population in Serbia 1981–2011 measured with the Hoover index	University of Belgrade - Faculty of Geography. ORIGINAL SCIENTIFIC PAPER DOI: 10.2298/ZMSDN1448461S https://www.researchgate.net/profile/Danica_Santic/publication/273294505_Spatial_concentration_of_population_in_Serbia_1981-2011_measured_with_the_Hoover_index/links/56dea1fb08aedf2bf0c9bdcb/Spatial-concentration-of-population-in-Serbia-1981-2011-measured-with-the-Hoover-index.pdf https://www.researchgate.net/profile/Danica_Santic/publication/273294505_Spatial_concentration_of_population_in_Serbia_1981-2011_measured_with_the_Hoover_index/links/56dea1fb08aedf2bf0c9bdcb/Spatial-concentration-of-population-in-Serbia-1981-2011-measured-with-the-Hoover-index.pdf

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
2014	Zhong, C. et al.	High-speed rail accessibility: a comparative analysis of urban access in Los Angeles, San Francisco, Madrid, and Barcelona.	EJTIR 14 (4), 2014, pp. 468-488. http://diposit.ub.edu/dspace/bitstream/2445/68516/1/634210.pdf
2014	Zurita, J.	Análisis de la concentración y competencia en el sector bancario	BBVA https://www.bbvaresearch.com/wp-content/uploads/2014/09/WP-concentraci%C3%B3n-y-competencia-sector-bancario.pdf
2013	Angulo, A.M. et al.	Análisis de concentración geográfica de la productividad: el caso de las empresas de manufacturas del Valle del Ebro	Estadística Española Volumen 55, número 180 / 2013, pp. 95-118 https://www.ine.es/ss/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3B+filename%3D180_6.pdf&blobkey=urldata&blobtable=MungoBlobs&blobwhere=612%2F135%2F180_6.pdf&ssbinary=true https://www.ine.es/ss/Satellite?L=0&c=INERevEstad_C&p=1254735226759&pagename=ProductosYServicios%2FPYSLayout&_charset_=utf-8&cid=1259942510685&submit=Ir
2013	Franch, X. et al.	Un análisis espacial de las pautas de crecimiento y concentración de la población a partir de series homogéneas: España (1877-2001)	Investigaciones regionales, 25 – Páginas 43 a 65 https://old.aecr.org/images//ImatgesArticles/2013/5/3_Franch.pdf
2013	Jurado, I. et al.	Políticas sociales, concentración geográfica y sistemas electorales	Cuadernos económicos de ICE Nº 85 http://www.revistasice.com/index.php/CICE/article/view/6055/6055
2013	Lis-Gutiérrez, J.P.	Medidas de concentración y estabilidad de mercado. Una aplicación para Excel	Grupo de Estudios Económicos - Superintendencia de Industria y Comercio de Colombia https://mpra.ub.uni-muenchen.de/47615/1/MPRA_paper_47615.pdf
2013	Liu, Z.	Geographical Concentration of Manufacturing Industries in China —Measurements and Determinants	University of Connecticut. DigitalCommons@UConn https://opencommons.uconn.edu/cgi/viewcontent.cgi?referer=http://www.google.es/url?sa=t&rct=j&q=&esrc=s&source=web&cd=38&ved=0ahUKEwjgz6Lx9IncAhUJshQKHTZ_Dsc4HhAWCGkwBw&url=http%3A%2F%2Fopencommons.uconn.edu%2Fcgi%2Fviewcontent.cgi%3Farticle%3D6339%26context%3Ddissertations&usg=AOvVaw20hmYdwhp8q79KfA_Ro997&httpsredir=1&article=6339&context=dissertations
2013	Malpezzi, S.	Population Density: Some Facts and Some Predictions	Cityscape: A Journal of Policy Development and Research. Volume 15. Number 3. https://www.huduser.gov/portal/periodicals/cityscpe/vol15num3/ch13.pdf
2013	Muñiz, I. et al.	Anatomía de la dispersión urbana en Barcelona.	EURE, vol. 39, no 116, enero 2013, pp. 189-219 https://www.researchgate.net/profile/Miquel-angel-Garcia-lopez/publication/270072050_Anatomia_de_la_dispersion_urbana_en_Barcelona/links/582ad37508ae102f071fea32/Anatomia-de-la-dispersion-urbana-en-Barcelona.pdf
2013	Pereira, R.H.M. et al.	Quantifying Urban Centrality: A Simple Index Proposal and International Comparison	Geographical Analysis (2013) 45, 77–89 22/publication/298419278_Quantifying_Urban_Centrality_A_Simple_Index_Proposal_and_International_Comparison_vol_45_pg_77_2013/links/5aa7bb0aaca272f7a16380e9/Quantifying-Urban-Centrality-A-Simple-Index-Proposal-and-International-Comparison-vol-45-pg-77-2013.pdf https://www.researchgate.net/profile/Rafael-Pereira-22/publication/258727927_Urban_Centrality_A_Simple_Index/links/59de4f1f458515376b29d6d5/Urban-Centrality-A-Simple-Index.pdf https://www.researchgate.net/profile/Rafael-Pereira-
2012	Atienza, M. et al.	Concentración y crecimiento en Chile: una relación negativa ignorada	EURE (Santiago) vol.38 no.114 Santiago mayo 2012 https://scielo.conicyt.cl/pdf/eure/v38n114/art10.pdf
2012	Azcárate, M.V. et al.	Análisis desarrollo urbano áreas metropolitanas Madrid y Granada	Universidad de Alcalá de Henares http://www.geogra.uah.es/simurban1/PDF/pdf_2012/segundo_capitulo.pdf
2012	Campos, C.	The Geographical Concentration of Industries	Office for National Statistics. UK. Publications https://webarchive.nationalarchives.gov.uk/20160106205639/https://ons.gov.uk/ons/dcp171766_272232.pdf http://webarchive.nationalarchives.gov.uk/20160106205639/http://www.ons.gov.uk/ons/rel/regional-trends/regional-economic-analysis/the-geographical-concentration-of-industries/art-geographical-concentration.html
2012	EUROSTAT	Eurostat regional yearbook 2012	EUROSTAT http://ec.europa.eu/regional_policy/en/newsroom/news/2012/10/eurostat-regional-yearbook-2012-and-online-interactive-tools-launched https://www.ab.gov.tr/files/ardb/evt/1_avrupa_birligi/1_6_raporlar/1_5_eurostat/eurostat_regioanal_yearbook_2012.PDF
2012	Folch, D.C.	The Centralization Index as a Measure of Local Spatial Segregation	Arizona State University https://repository.asu.edu/attachments/94051/content//tmp/package-GobAVc/Folch_asu_0010E_12085.pdf

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
2012	Santos-Preciado, J.M. et al.	Methodologies for the measurement of urban dispersion, in a GIS environment. Application to the study of the Community of Madrid	XV Congreso Nacional de Tecnologías de la Información Geográfica, Madrid, AGE-CSIC, 19-21 de Septiembre de 2012 http://www.age-geografia.es/tig/2012_Madrid/ponencia2/SantosPreciados_final_imp.pdf
2011	Berti, S. et al.	Lagrangian Drifter Dispersion in the Southwestern Atlantic Ocean	American Meteorological Society https://journals.ametsoc.org/doi/pdf/10.1175/2011JPO4541.1
2011	De Vries, A. et al.	System of regional Indicators of Navarre	Observatorio regional de Navarra https://nasuvinsa.es/sites/default/files/images/Sistema_Indicadores_regionales_Navarra.pdf
2011	EUROSTAT	Eurostat regional yearbook 2010	EUROSTAT http://ec.europa.eu/eurostat/documents/3217494/5728777/KS-HA-11-001-EN.PDF
2011	Salom, A.	Las limitaciones al crecimiento poblacional y espacial establecidas por la normativa regional y urbanística	INAP. Monografías http://www.izenpe.eus/s15-4812/es/contenidos/informacion/bibl_digital/es_documen/adjuntos/Limitaciones.pdf
2011	Waltman, L. et al.	Globalisation of science in kilometres	Cornell University Library https://arxiv.org/ftp/arxiv/papers/1103/1103.3648.pdf
2010	Ayuda, M.I. et al.	Long-run regional population disparities in Europe during modern economic growth: a case study of Spain	Ann Reg Sci (2010) 44:273–295 DOI 10.1007/s00168-008-0260-9 https://economia_aplicada.unizar.es/sites/economia_aplicada.unizar.es/files/archivos/55/Collantes/ayuda_et_al._2010.pdf
2010	Campante, F.R. et al.	A Centered Index of Spatial Concentration: Expected Influence Approach and Application to Population and Capital Cities	Semantic Scholar. Allen Institute for AI (AI2) file://erc-per-ccdd.central.sepg.minhac.age/FolderRedirection01$/KD000169/Desktop/SSRN-id1362349.pdf https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID1362349_code518410.pdf?abstractid=1362349&mirid=1 https://pdfs.semanticscholar.org/2327/a33ec6b39c96a3d0a70fd746df061af666cd.pdf
2010	EUROSTAT	Eurostat regional yearbook 2010	EUROSTAT http://ec.europa.eu/eurostat/documents/3217494/5727301/KS-HA-10-001-EN.PDF/1ba3cf6a-5e25-44c1-99f9-fada17625212
2010	Hindle, G.A. et al.	Developing geographical indicators of mileage-related costs: a case study exploring travelling public services in English local areas	The Journal of the Operational Research Society https://www.researchgate.net/profile/Giles_Hindle/publication/32009092_Developing_geographical_indicators_of_mileage-related_costs_A_case_study_exploring_travelling_public_services_in_English_local_areas/links/56373eba08ae88cf81bd51b3/Developing-geographical-indicators-of-mileage-related-costs-A-case-study-exploring-travelling-public-services-in-English-local-areas.pdf?origin=publication_detail https://www.researchgate.net/publication/32009092_Developing_geographical_indicators_of_mileage-related_costs_A_case_study_exploring_travelling_public_services_in_English_local_areas
2010	Martínez-López, M.	A sociological and urban perspective about the social densities in the urban centre	Miguel Martinez López. Universidad Complutense de Madrid. Depto. Sociología II. http://www.miguelangelmartinez.net/?Una-perspectiva-sociologica-y http://www.miguelangelmartinez.net/IMG/pdf/2010_Densidades_centro_urbano_Marco_teorico_Miguel.pdf
2010	Monkkonen, P.	Measuring Residential Segregation in Urban Mexico: Levels and Patterns	Working Paper 2010-05. Berkeley Institute of Urban and regional Development. University of California https://iurd.berkeley.edu/wp/2010-05.pdf
2010	Otto, A. et al.	Long-term processes of regional concentration and dispersion - fuzzy evidence for Western Germany	50th Congress of the European regional Science Association: "Sustainable regional Growth and Development in the Creative Knowledge Economy", 19-23 August 2010, Jönköping, Sweden. Provided in Cooperation with ERSA https://www.econstor.eu/bitstream/10419/119036/1/ERSA2010_0537.pdf
2010	Rueda, S. et al.	Plan of Urban Sustainability Indicators of Vitoria-Gasteiz	Departamento de Medioambiente de la ciudad de Vitoria-Gasteiz. http://www.vitoria-gasteiz.org/wb021/http/contenidosEstaticos/adjuntos/es/89/14/38914.pdf
2010	Salazar, X. et al.	Las empresas en Santa Cruz, ¿continúan en una estructura monocéntrica?	Revista Latinoamericana de Desarrollo Económico. Versión On-line ISSN 2074-4706 http://www.scielo.org.bo/pdf/rlde/n13/v8n13a02.pdf
2010	Vázquez, R. et al.	Principales características demográficas con fuerte impacto en la ordenación sanitaria y financiación de la Comunidad Autónoma de Galicia	FEGAS https://extranet.sergas.es/catpb/Docs/cas/Publicaciones/Docs/PlOrSanitaria/PDF-1865-es.pdf
2009	Campante, FR. et al.	A Centered Index of Spatial Concentration: Expected Influence Approach and Application to Population and Capital Cities	Singapore Management University http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.208.5639&rep=rep1&type=pdf

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
2009	Jiang, C. et al.	The Measurement Method of Population Dynamics and New Measurement Index	International Journal of Biology Vol.1 No. 1 http://www.ccsenet.org/journal/index.php/ijb/article/download/664/641
2009	PUCC	Índices de Medición de la Concentración	Pontificia Universidad Católica de Chile. Departamento de Ingeniería Eléctrica http://hrudnick.sitios.ing.uc.cl/alumno09/concentra/indices/indices.htm
2009	Subero Munilla, J.M.	Métodos de análisis de la eficacia espacial de las redes de transporte colectivo de infraestructura fija, ensayo de indicadores de oferta	Universidad Politécnica de Cataluña. Doctoral Thesis *Retrieved from:* https://upcommons.upc.edu/bitstream/handle/2117/93941/TJSM1de4.pdf
2009	Zhemoyda, O. et al.	Concentration of agriculture and competitive advantages of East-European countries.	113th EAAE Seminar "A resilient European food industry and food chain in a challenging world" https://ageconsearch.umn.edu/bitstream/58147/2/Zhemoyda.pdf
2008	Aso, Y.	Social Overhead Capital Development and Geographical Concentration	Policy Research Institute, Ministry of Finance, Japan, Public Policy Review, Vol.4, No.1, December 2008. https://eaber.org/wp-content/uploads/2011/05/PRI_Aso_2008.pdf
2008	Campante, F.R. et al.	A Centered Index of Spatial Concentration: Axiomatic Approach with an Application to Population and Capital Cities	Social Science Research Network – SSRN. https://ssrn.com/abstract=1362349 http://dx.doi.org/10.2139/ssrn.1362349
2008	Combes, P.P. et al.	The rise and fall of spatial inaqualities in France : a long-run perspective	Document de Travail n°2008-60 GREQAM Groupement de Recherche en Economie Quantitative d'Aix-Marseille - UMR-CNRS 6579 Ecole des Hautes Etudes en Sciences Sociales Universités d'Aix-Marseille II et III https://halshs.archives-ouvertes.fr/halshs-00349293/document
2008	Cutrini, E.	Using entropy measures to disentangle regional from national localization patterns	regional Science and Urban Economics https://u-pad.unimc.it/retrieve/handle/11393/38492/579/article-RSUE-2008-inpress.pdf
2008	Goerlich, F.J. et al.	Empirical Evidence of Population Concentration in Spain, 1900-2001	Population 2008/4 (Vol. 63) https://www.cairn-int.info/load_pdf.php?ID_ARTICLE=E_POPU_804_0731 https://www.cairn-int.info/article-E_POPU_804_0731--empirical-evidence-of-population-concent.htm
2008	He, C. et al.	Globalization, Institutional Change, and Industrial Location: Economic Transition and Industrial Concentration in China	regional Studies http://www.tandfonline.com/loi/cres20 https://www.tandfonline.com/doi/pdf/10.1080/00343400701543272?needAccess=true
2008	Marmolejo, C.	La localización intrametropolitana de las actividades de la información: un análisis para la región metropolitana de Barcelona 1991-2001	Revista electrónica de geografía y ciencias sociales. Universidad de Barcelona http://www.ub.edu/geocrit/sn/sn-268.htm
2008	Moyano Pesquera, P.B.	The incidence of the dispersion of the population in the cost of the sanitary services in the Spanish autonomous communities	Paper presented in the "XI Congreso de Economía de Castilla y León" http://www.jcyl.es/web/jcyl/binarios/874/471/G16_69_%20Incidencia%20de%20la%20dispersi%C3%B3n%20de%20la%20poblaci%C3%B3n%20en%20los%20costes%20sanitarios.pdf?blobheader=application%2Fpdf%3Bcharset%3DUTF-8&blobheadername1=Cache-Control&blobheadername2=Expires&blobheadername3=Site&blobheadervalue1=no-store%2Cno-cache%2Cmust-revalidate&blobheadervalue2=0&blobheadervalue3=JCYL_EconomiaEmpleo&blobnocache=true
2008	Roca Cladera, J.	La localización intrametropolitana de las actividades de la información: un análisis para la región metropolitana de Barcelona 1991-2001	Centro de Política de Suelo y Valoraciones de la Universidad Politécnica de Cataluña http://www.ub.edu/geocrit/sn/sn-268.htm
2007	Ayuntamiento de Oviedo	Diagnóstico del sistema ambiental. Agenda 21 5.8. Medio ambiente urbano	Ayuntamiento de Oviedo http://www.oviedo.es/documents/12103/daa9755e-0c84-415f-a059-99af6d309748
2007	Castañeda Véliz, C.C.	Indicadores de Concentración: Una revisión del marco conceptual y la experiencia internacional	OSIPTEL https://studylib.es/doc/7467222/marco-conceptual-sobre-indicadores-de-concentraci%C3%B3n-de https://docplayer.es/43932446-Indicadores-de-concentracion-una-revision-del-marco-conceptual-y-la-experiencia-internacional-1.html
2007	De Dominicis, L. et al.	The Spatial Distribution of Economic Activities in Italy	TI 2007-094/3. Tinbergen Institute Discussion Paper https://papers.tinbergen.nl/07094.pdf

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
2007	Do, Q-A et al.	A Centered Measure of Spatial Concentration: A Gravity-Based Approach with an Application to Population and Capital Cities	Research Gate https://www.researchgate.net/profile/Quoc-Anh_Do/publication/228692254_A_centered_measure_of_spatial_concentration_A_gravity-based_approach_with_an_application_to_population_and_capital_cities/links/02e7e527cb01a580a7000000/A-centered-measure-of-spatial-concentration-A-gravity-based-approach-with-an-application-to-population-and-capital-cities.pdf
2007	Gorbea, S.	Principales revistas latinoamericanas en ciencias bibliotecológica y de la información: su difusión y su concentración temática y geográfica	Investigación Bibliotecológica: archivonomía, bibliotecología e información, Vol. 21, Núm. 42 (2007) http://rev-ib.unam.mx/ib/index.php/ib/rt/printerFriendly/4119/51452
2007	Nieto, V.M.	La aglomeración como una causa de la innovación en Colombia	Planeación & Desarrollo. Volumen XXXVIII, Número 1 (2007) 97-139. https://colaboracion.dnp.gov.co/CDT/Desarrollo%20Empresarial/La%20aglomeraci%C3%B3n%20como%20una%20causa%2097-139%20-%20JUNIO%2023.pdf
2007	OECD	Transport, Urban Form and Economic Growth	Document prepared for the Round Table of the OECD (Organization for Economic Cooperation and Development), Berkeley, CA, March 2006 https://one.oecd.org/document/74200707/en/pdf
2006	Alonso, O.	La concentración geográfica de la industria en Galicia	Fundación Pedro Barrié de la Maza. Instituto de Estudios Económicos de Galicia http://alonso-villar.webs.uvigo.es/Libro-Barrie.pdf
2006	Franz, G. et al.	Urban Sprawl. How Useful Is This Concept?	European regional Science Association http://www-sre.wu.ac.at/ersa/ersaconfs/ersa06/papers/105.pdf
2006	Giovanni, L. et al.	Describing Income Inequality Theil Index and Entropy Class Indexes.	FAO. http://www.fao.org/3/am343e/am343e.pdf
2006	Goerlich, F.J. et al.	La localización de la población española sobre el territorio: un siglo de cambios: un estudio basado en series homogéneas (1900-2001)	Fundación BBVA https://www.fbbva.es/wp-content/uploads/2017/05/dat/DE_2006_IVIE_localizacion_poblacion_espanola.pdf
2006	Lee, B.	Urban spatial structure, commuting, and growth in us metropolitan areas	A Dissertation Presented to the Faculty of the Graduate School University of Southern California http://digitallibrary.usc.edu/cdm/compoundobject/collection/p15799coll127/id/32544/rec/7709
2006	Marcus, M.J.	Density: How Concentrated is Our Population	Indiana Business Review, Summer 2006 http://www.ibrc.indiana.edu/ibr/2006/summer/pdfs/1_density.pdf
2006	Picchizzolu, P.	Entrepreneurial Risk and the Geographical Concentration of Industries: UK 1996	Semantic Scholar. Allen Institute for AI (AI2) *This document provides methods for the calculation of industrial concentration. Retrieved from:* http://193.205.83.2/~confeco/papers/picchizzolu.pdf https://pdfs.semanticscholar.org/31c7/3100fcf41eb7bad28b9f26ec339aaabc981b.pdf
2005	Auyda, M.I. et al.	Explicando la concentración a largo plazo de la población española, 1860-2000	DOCUMENTOS DE TRABAJO 2005-3 Facultad de Ciencias Económicas y Empresariales. Universidad de Zaragoza. https://www.researchgate.net/profile/M-Ayuda/publication/28143728_Explicando_la_concentracion_a_largo_plazo_de_la_poblacion_espanola_1860-2000/links/0fcfd50efcd14a32c8000000/Explicando-la-concentracion-a-largo-plazo-de-la-poblacion-espanola-1860-2000.pdf
2005	Bertinelli, L.	Geographical Agglomeration: Ellison and Glaeser's Index Applied to the Case of Belgian Manufacturing Industry	regional Studies, Vol. 39.5, pp. 567–583, July 2005 https://www.tandfonline.com/doi/pdf/10.1080/00343400500151806?needAccess=true
2005	Coelho Avila, P.	Interpretation of Population Density Gradients: A Brazilian Perspective	Third Urban Research Symposium on "Land Development, Urban Policy and Poverty Reduction" sponsored by the World Bank and IPEA (Instituto de Pesquisa Econômica Aplicada). https://www.researchgate.net/profile/Paulo-Avila-4/publication/228156025_Interpretation_of_Population_Density_Gradients_A_Brazilian_Perspective/links/0f317531bad0010aab000000/Interpretation-of-Population-Density-Gradients-A-Brazilian-Perspective.pdf

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
2005	Couch, C. et al.	Decline and sprawl: An evolving type of urban development - Observed in Liverpool and Leipzig	European Planning Studies, 13(1), 117–136 https://www.researchgate.net/profile/Dieter_Rink/publication/263453477_Decline_and_sprawl_An_evolving_type_of_urban_development_-_Observed_in_Liverpool_and_Leipzig/links/54aabefa0cf2bce6aa1d5f4d/Decline-and-sprawl-An-evolving-type-of-urban-development-Observed-in-Liverpool-and-Leipzig.pdf?origin=publication_detail
2005	Hierro, M. et al.	Concentración poblacional y dispersión regional y migratoria en España, 1986-2003	Documentos de trabajo 2005-4. Departamento de Economía. Universidad de Cantabria https://ceddar.org/wp-content/uploads/DT-2005-4.pdf
2005	Martí-Henneberg, J.	Empirical Evidence of regional Population Concentration in Europe, 1870–2000	POPULATION, SPACE AND PLACE *Popul. Space Place 11, 269–281 (2005)* Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psp.373 http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.510.1138&rep=rep1&type=pdf https://scholar.google.co.id/citations?user=r5jGFAsAAAAJ&hl=fr
2005	Santa María, M.J. et al.	La concentración espacial de la industria en España	Universidad de Alicante. XXXI Congreso de la Asociación Española de Ciencia regional https://rua.ua.es/dspace/bitstream/10045/2766/1/AECR_2006.pdf
2005	Tsai, Y-H.	Quantifying urban form: Compactness versus 'sprawl'.	Urban Studies, 42(1), 141– https://www.researchgate.net/profile/Yu-Hsin_Tsai2/publication/228985804_Quantifying_Urban_Form_Compactness_versus_%27Sprawl%27/links/0912f512442382fe4f000000/Quantifying-Urban-Form-Compactness-versus-Sprawl.pdf?origin=publication_detail
2005	Wassmer, R.W. et al.	Causes of Urban Sprawl (Decentralization) in the United States: Natural Evolution, Flight from Blight, and the Fiscalization of Land Use	California State University https://www.researchgate.net/profile/Robert_Wassmer/publication/229048043_Causes_of_urban_sprawl_decentralization_in_the_United_States_natural_Evolution_flight_from_blight_and_the_fiscalization_of_land_use/links/0c96052142b1d79401000000/Causes-of-urban-sprawl-decentralization-in-the-United-States-natural-Evolution-flight-from-blight-and-the-fiscalization-of-land-use.pdf?origin=publication_detail
2004	Aiginger, K.	Industrial specialisation and geographic concentration: two sides of the same coin? Not for the European Union	Journal of Applied Economics, Vol. VII, No. 2 (Nov 2004), 231-248 https://ageconsearch.umn.edu/bitstream/37627/2/aisinger.pdf
2004	Bárcena Martín, E. et al.	Medición de la privación y del bienestar entre poblaciones mediante el índice de Gini.	ESTADÍSTICA ESPAÑOLA. Vol. 46, Núm. 155, 2004, págs. 49 a 76 http://www.ine.es/ss/Satellite?L=0&c=INERevEstad_C&p=1254735226759&pagename=ProductosYServicios%2FPYSLayout&_charset_=UTF-8&cid=1259925024729 https://www.ine.es/ss/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3B+filename%3D155_3.pdf&blobkey=urldata&blobtable=MungoBlobs&blobwhere=393%2F222%2F155_3%2C0.pdf&ssbinary=true
2004	Cosby, K.L.	From Mission to Megacity: The Changing Concentration of the Los Angeles City-system	Brigham Young University. BYU ScholarsArchive https://scholarsarchive.byu.edu/cgi/viewcontent.cgi?article=1026&context=etd
2004	Glaeser, E.L.. et al.	Sprawl and urban growth	En J. V. Henderson, P. Nijkamp, E. S. Mills, P. C. Cheshire & J. F. Thisse (Eds.), Handbook of regional and Urban Economics (vol. 4, pp. 2481–2527). Amsterdam: Elsevier http://commons.trincoll.edu/wp-content/blogs.dir/411/files/2013/09/sprawl-and-urban-growth-glaeser.pdf
2004	Hildebrandt, A. et al.	Determinants of Geographical Concentration Patterns in Central and Eastern European Countries	Focus on European Economic Integration Q1-04. Central bank of the Republic of Austria (OeNB) https://www.oenb.at/dam/jcr:271cab85-918b-4c69-9a25-5a484bf5241f/determinants_tcm16-20265.pdf
2004	Lemelin, A.	Medición de la desigualdad y de la concentración en "Métodos cuantitativos"	L'Institut national de la recherche scientifique (INRS). Canada http://www.lemelin-metho.ucs.inrs.ca/wp-content/uploads/es1_4.pdf http://www.lemelin-metho.ucs.inrs.ca/
2004	Paluzi, E. et al.	The geographical concentration of industry across Spanish regions, 1856-1995	DOCUMENTS DE TREBALL DE LA DIVISIÓ DE CIÈNCIES JURÍDIQUES ECONÒMIQUES I SOCIALS. Col·lecció d'Economia. Universitat de Barcelona. https://www.researchgate.net/profile/Daniel-Tirado/publication/36729485_The_geographical_concentration_of_industry_across_Spanish_regions_1856-1995/links/00b4951ececcdd00db000000/The-geographical-concentration-of-industry-across-Spanish-regions-1856-1995.pdf
2004	Park, S-C	An Empirical Assessment of Geographical Concentration of Industries in East Asia	Korea Institute for International Economic Policy https://www.kiep.go.kr/boardDownload.es?bid=0034&list_no=5726&seq=1

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
2004	Quinn, L.M. et al.	Assumptions and Limitations of the Census Bureau. Methodology Ranking Racial and Ethnic Residential Segregation in Cities and Metro Areas	Census Bureau USA https://www.census.gov/hhes/www/housing/resseg/pdf/quinn.pdf
2004	Son, Y. et al.	Measuring Urban Form: Is Portland Winning the War on Sprawl?	Journal of the American Planning Association, 70(2), 210–225 https://www.researchgate.net/profile/Gerrit_Jan_Knaap/publication/232973827_Measuring_Urban_Form_Is_Portland_Winning_the_War_on_Sprawl/links/0deec5308b774c32ec000000/Measuring-Urban-Form-Is-Portland-Winning-the-War-on-Sprawl.pdf?origin=publication_detail
2003	Alonso, O. et al.	Spillovers geográficos y sectoriales de la industria.	Revista de Economía Aplicada Número 32 (vol. XI), 2003, págs. 77 a 95 http://www.redalyc.org/pdf/969/96917632004.pdf
2003	Barrios, S. et al.	Industry mobility and geographic concentration in the European Union.	Göttingen State and University Library. Core Discussion Paper 2003/43 http://webdoc.sub.gwdg.de/ebook/serien/e/CORE/dp2003-43.pdf
2003	Otterstrom, S.M. et al.	Deserts and oases: the continuing concentration of population in the American Mountain West	Journal of Rural Studies 19 (2003) 445–462 https://www.researchgate.net/profile/J-Matthew-Shumway/publication/238069873_Deserts_and_Oases_The_Continuing_Concentration_of_Population_in_the_American_Mountain_West/links/5b082fddaca2725783e552b3/Deserts-and-Oases-The-Continuing-Concentration-of-Population-in-the-American-Mountain-West.pdf
2003	Rainis R. et al.	The effects of data aggregation on the spatial analysis of population concentration	Jurnal Teknologi, 39(B) Dis. 2003: 47–62. Universiti Teknologi Malaysia https://journals.utm.my/jurnalteknologi/article/view/448/438 file://erc-per-ccdd.central.sepg.minhac.age/FolderRedirection01$/KD000169/Desktop/448-445-1-PBTheEffectsofDataAggregationontheSpatialAnalysisofPopulationConcentration.pdf
2003	Tirado, D.A, et al.	Industrial agglomerations and wage gradients: the Spanish economy in the interwar period	ERSA conference papers ersa03p113, European Regional Science Association http://www-sre.wu.ac.at/ersa/ersaconfs/ersa03/cdrom/papers/113.pdf http://diposit.ub.edu/dspace/bitstream/2445/11973/1/103.pdf
2002	Aiginger, K. et al.	regional concentration in the USA and Europe	Weltwirtschaftliches Archiv, Vol. 138(4), pp. 1-28, 2002 (jointly with Leitner, W.) https://karl.aiginger.wifo.ac.at/fileadmin/publications/2002/regconcfin_01.pdf
2002	Ewing et al.	Measuring sprawl and its impact	Rutgers University and Smart Growth America http://smartgrowthamerica.us/documents/MeasuringSprawlTechnical.pdf
2002	Iceland, J. et al.	Racial and Ethnic Residential Segregation in the United States: 1980-2000	United States Census Bureau https://www.census.gov/prod/2002pubs/censr-3.pdf
2002	Iceland, J. et al.	Racial and Ethnic Residential Segregation in the United States: 1980-2000	United States Census Bureau https://www.census.gov/content/dam/Census/library/publications/2002/dec/censr-3.pdf
2002	Lage de Sousa, F.	A Localização da Indústria de Transformação Brasileira nas Últimas Três Décadas	Fundação Getúlio Vargas – FGV. Escola de Pós-Graduação em Economia – EPGE Dissertação submetida à Congregação da Escola de Pós-Graduação em Economia (EPGE) para obtenção do grau de Mestre em Economia http://www.race.nuca.ie.ufrj.br/teses/fgv/Sousa,F.pdf
2002	Midelfart-Knarvik K.H. et al.	Integration and Industrial Specialisation in the European Union	Revue économique, volume 53, n°3, 2002. pp. 469-481 file://erc-per-ccdd.central.sepg.minhac.age/FolderRedirection01$/KD000169/Desktop/MidelfartOvermanReddingVenables-RE2002.pdf https://www.persee.fr/doc/reco_0035-2764_2002_num_53_3_410418
2002	Rodriguez, J.	Distribución regional de la población de América Latina y el Caribe: tendencias, interpretaciones y desafíos para las políticas públicas	Centro Latinoamericano y Caribeño de Demografía (CELADE) - División de Población. CEPAL https://digitallibrary.un.org/record/489690/files/%5BE_%5DLC_L.1831-P-ES.pdf
2002	Sanchis, A. et al.	Estructura de Mercado y Concentración	Universidad de Valencia https://www.uv.es/~jamc/eia/econind1
2002	Sousa, F.	localização da indústria de transformação brasileira nas últimas três décadas.	Rio de Janeiro FGV. http://bibliotecadigital.fgv.br/dspace;/bitstream/handle/10438/41/1448.pdf?sequence=1&isAllowed=y
2001	Alonso, O. et al.	An Analysis of the Geographic Concentration of Industry in Spain	Working Papers 0105, Universidade de Vigo. Departamento de Economía Aplicada https://dialnet.uniRioja.es/descarga/articulo/2252256.pdf
2001	Galster, G. et al.	Wrestling sprawl to the ground: Defining and measuring an elusive concept.	Housing Policy Debate, 12, 681–717 http://www.gulfofmaine.org/2/wp-content/uploads/2014/03/Wrestling-Sprawl-to-the-Ground-Defining-and-Measuring-an-Elusive-Concept.pdf

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
2001	Malpezzi, S. et al.	Measuring "sprawl". Alternative measures of urban form in U.S. metropolitan areas.	Madison, WI: The Center for Urban Land Economics Research, University of Wisconsin http://smalpezzi.marginalq.com/documents/Alternative%20Measures%20of%20Urban%20Form.doc http://smalpezzi.marginalq.com/materials/
2001	Martori, J.C. et al.	Classical models of urban population density. The case of Barcelona Metropolitan Area.	Nederlandse Geografische Studies. https://www.researchgate.net/publication/23730354_Classical_models_of_urban_population_density_The_case_of_Barcelona_Metropolitan_Area 41st Congress of the European regional Science Association. 29 August-1 September 2001 (Zagreb). http://www-sre.wu.ac.at/ersa/ersaconfs/ersa01/papers/full/46.pdf
2001	Nuñez, S. et al.	El grado de concentración en las ramas productivas de la economía española	Banco de España-BDE https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/01/Fic/dt0113.pdf
2000	Mayor, M. et al.	Una aproximación al gradiente de densidad de población	Anales de Economía Aplicada. XIV Reunión ASEPELT-España. Oviedo, 22 y 23 de junio de 2000. https://www.asepelt.org/ficheros/File/Anales/2000%20-%20Oviedo/Trabajos/PDF/161.pdf
2000	Midelfart-Knarvik K.H. et al.	The Location of European Industry._European Economy	European Union Economic Papers (ECFIN/318/00-EN (Revised 25.2.2000)) https://ec.europa.eu/economy_finance/publications/pages/publication11136_en.pdf
2000	Torrens, P.M. et al.	Measuring Sprawl	CASA Working Paper 27, Center for Advanced Spatial Analysis (CASA), University College London https://www.ucl.ac.uk/bartlett/casa/sites/bartlett/files/migrated-files/paper27.pdf
2000	Wassmer, R.W.	Urban Sprawl in a U.S. Metropolitan Area: Ways to Measure and a Comparison of the Sacramento Area to Similar Metropolitan Areas in California and the U.S.	Sacramento, CA: Department of Public Policy and Administration, California State University https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID241975_code000928510.pdf?abstractid=241975&mirid=1
1999	Bertaud, A. et al.	The spatial distribution of population in 35 world cities: The role of markets, planning, and topography.	Wisconsin-Madison CULER working papers 01–03, University of Wisconsin, Centre for Urban Land Economics Research. http://mcrit.com/scenarios/files/documents/Bertaud%20and%20Malpezzi%20Part%20One.pdf
1998	Anas, A. et al.	Urban spatial structure	UCI School of Social Sciences https://www.socsci.uci.edu/~ksmall/JEL%20Paper.pdf
1998	Callejón, M.	Concentración geográfica de la industria y economías de aglomeración	Universidad de Barcelona. Facultad de Ciencias Económicas y Empresariales http://www.ub.edu/graap/pdfcallejon/Callejon-geo-concentr.pdf
1997	Ellison, G. et al.	Geographic Concentration in U.S. Manufacturing Industries: A Dartboard Approach	Journal of Political Economy 105: 889-927 http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.548.1673&rep=rep1&type=pdf
1997	Ferreira, E. et al.	Una nota sobre el cálculo de índice Gini.	ESTADÍSTICA ESPAÑOLA. Vol. 39, Núm. 142, 1997, págs. 207 a 218 http://www.ine.es/ss/Satellite?L=0&c=INERevEstad_C&p=1254735226759&pagename=ProductosYServicios%2FPYSLayout&_charset_=utf-8&cid=1259925027717&submit=Ir https://www.ine.es/ss/Satellite?blobcol=urldata&blobheader=application%2Fpdf&blobheadername1=Content-Disposition&blobheadervalue1=attachment%3B+filename%3D142_9.pdf&blobkey=urldata&blobtable=MungoBlobs&blobwhere=1005%2F265%2F142_9%2C0.pdf&ssbinary=true
1997	Otterstrom, S.M.	An Analysis of Population Dispersal and Concentration in the United States, 1790-1990: The Frontier, Long Waves, and the Manufacturing Connection.	Louisiana State University. LSU Digital Commons https://digitalcommons.lsu.edu/cgi/viewcontent.cgi?article=7585&context=gradschool_disstheses https://digitalcommons.lsu.edu/gradschool_disstheses/6586/
1990	Fossett, M.	New Methods for Measuring and Analysing Segregation	Springer https://link.springer.com/content/pdf/10.1007%2F978-3-319-41304-4.pdf
1990	Fossett, M.	New Methods for Measuring and Analysing Segregation	Springer https://link.springer.com/content/pdf/10.1007%2F978-3-319-41304-4.pdf
1988	Massey, D. and Denton, N.	The Dimensions of Residential Segregation	Social Forces, Volume 67, Issue 2, December 1988, Pages 281–315. https://academic.oup.com/sf/article-pdf/67/2/281/6514769/67-2-281.pdf
1977	Stephan, E.	An Empirically Standardized Measure of Population Concentration	Geographical Analysis https://onlinelibrary.wiley.com/doi/epdf/10.1111/j.1538-4632.1977.tb00582.x

YEAR	AUTHOR/S	TITLE	INSTITUTION/RETRIEVED FROM
1973	Ramos, L.M.	La concentración de la población y el proceso de urbanización: Veracruz 1940-1970	Centro de Estudios Económicos y Sociales de la Facultad de Economía de la Universidad Veracruzana https://cdigital.uv.mx/bitstream/handle/123456789/9310/1973_vol2_no2_p315-334.pdf;jsessionid=0EA68E49B9B4D0740BB2BCDF2F533535?sequence=1 https://cdigital.uv.mx/handle/123456789/9310
1951	Clark, C.	Urban Population Densities	Journal of the Royal Statistical Society 114(4), 490–96 http://www.brown.edu/Departments/Economics/Faculty/Matthew_Turner/ec2410/readings/Clarke_RSA_1951.pdf https://www.jstor.org/stable/2981088?seq=1
-	-	Spatial Distribution and Density of Population	Statistics Bureau of Japan http://www.stat.go.jp/info/meetings/cambodia/pdf/a02_chap.pdf
-	CEPAL	Distribución de la población urbana	CEPAL https://celade.cepal.org/redatam/pryesp/cairo/WebHelp/Metalatina/distribucion_de_la_poblacion_urbana.htm
-	INECC	Índice de dispersión de la población por municipio	Instituto Nacional de Ecología y Cambio Climático. Gobierno de Méjico https://www.inecc.gob.mx/emapas/
-	SINCHI	Population Density.	Indicators of the database INIRIDA. Instituto Amazónico de Investigaciones Científicas SINCHI https://www.sinchi.org.co/inirida/indicadores https://www.sinchi.org.co/files/Base%20de%20Datos%20Inirida/PDF/01_Densidad%20de%20poblacion.pdf
-	SINCHI	Population Gini Index.	Indicators of the database INIRIDA. Instituto Amazónico de Investigaciones Científicas SINCHI https://www.sinchi.org.co/inirida/indicadores https://www.sinchi.org.co/files/Base%20de%20Datos%20Inirida/PDF/07_Indice%20Gini%20de%20la%20poblacion.pdf
-	SINCHI	Population regional Concentration Index.	Indicators of the database INIRIDA. Instituto Amazónico de Investigaciones Científicas SINCHI https://www.sinchi.org.co/inirida/indicadores https://www.sinchi.org.co/files/Base%20de%20Datos%20Inirida/PDF/06_Concentracion%20geografica%20de%20la%20poblacion.pdf
-	SINCHI	Index of Urban Population Concentratio.	Indicators of the database INIRIDA. Instituto Amazónico de Investigaciones Científicas SINCHI https://www.sinchi.org.co/files/Base%20de%20Datos%20Inirida/PDF/14_indice%20de%20concentracion%20de%20la%20poblacion%20urbana%20o%20de%20Pinchemel.pdf

ANNEX II. NOMENCLATURE

Geographical unit of analysis: Spain's provinces except Ceuta and Melilla

$\mathbf{P_i}$ $\quad i \in I = \{1 \ldots 50\}$

We notice that this work uses the Spanish administrative nomenclature of territorial units: singular entities, municipalities and provinces. It is aligned with the European Nomenclature of Territorial Units for Statistics, the NUTS classification that subdivides the territory of the EU Member States into territorial units that favours administrative units already existing in the Member States. The NUTS are complemented at the lower level by local administrative units (LAU). This is a different approach from that based on the new tool of the population grids, which is an alternative to population statistics for administrative areas. Please refer to EUROSTAT for further detail on this new tool: EUROSTAT (c) and (d).

Land uses: Parcels of land for population residence:

○ *Municipality (MUN)*

$\mathbf{M_j}$	$j \in J = \{1 \ldots \mu\}$	$\mathbf{J}$	set of Spain's municipalities
$\mathbf{M_{ij}}$	$j \in J_i = \{1 \ldots \mu_i\}$	$\mathbf{J_i}$	set of municipalities of province P_i
		$\boldsymbol{\mu}$	number of municipalities in Spain
		$\boldsymbol{\mu_i}$	number of municipalities in province i

Municipalities in Spain are the basic local entity of the territorial organization of the State. They have legal personality and full capacity to exercise their powers.[47] The territory and the population are elements of the municipality. The Continuous Municipal Register (INE) recognises that the national territory is administratively divided into regions,[48] provinces, municipalities and minor local entities; and, traditionally, within the municipalities it subdivides into population entities, which in turn can be singular or collective (intermediate unit between the singular population entity and the municipality consisting of a group of several singular entities). For the purpose of this work, municipalities are basic territorial cells for the analysis and represent land uses whose location is determined by the centroid of the municipality that we will define below.

○ *Singular population entity (in this paper singular entity —SE)*

$\mathbf{se_k}$	$k \in K = \{1 \ldots \#\}$	$\mathbf{K}$	set of Spain's singular entities
$\mathbf{se_{ik}}$	$k \in K_i = \{1 \ldots \#_i\}$	$\mathbf{K_i}$	set of singular entities of province P_i
$\mathbf{se_{ijk}}$	$k \in K_{ij} = \{1 \ldots \#_{ij}\}$	$\mathbf{K_{ij}}$	set of singular entities of municipality j of province i
		$\mathbf{\#}$	number of singular entities in Spain
		$\mathbf{\#_i}$	number of singular entities in province i
		$\mathbf{\#_{ij}}$	number of singular entities in municipality j of province i

A singular entity[49] is understood to be any habitable area of the municipal terminality, inhabited or exceptionally inhabited, clearly differentiated within it and which is known by a specific denomination that identifies it without possibility of confusion. An area is considered habitable when there are habited dwellings or in a condition to be inhabited. An area is considered clearly differentiated when the buildings and dwellings belonging to the same area may be perfectly identified on the ground and the whole area is known by a denomination. Consequently, the seasonal residential urbanisations and areas may be singular population entities even when they are inhabited for certain periods of the year. No dwelling may belong simultaneously to two or more singular entities. A municipality may consist of one or more singular population entities. If in a municipality there do not exist clearly differentiated areas the municipality will be considered a unique entity. A municipality may contain only one single entity or group its population into different singular population entities, when there are different habitable areas, clearly differentiated and with their own name. The total population of the municipality will be the sum of those who reside in each of these entities.[50] For the purpose of this work, singular entities are basic territorial cells for the analysis and represent land uses whose location is determine by its centroid as defined below.

Observation units ○ *Population.*

π_{ijk}	population living in singular entity $\mathbf{se}_{ijk}$.
π_{ik}	total population living in singular entity k of province i.
π_{ij}	total population living in municipality j of province i.
π_i	total population living in province i.
π	total population living in Spain.

[47] Article 11 of Law 7/1985 of 2 April, Regulator of the Local Regimen Bases: https://www.boe.es/buscar/pdf/1985/BOE-A-1985-5392-consolidado.pdf.

[48] Please refer to http://www.ine.es/nomen2/Metodologia.do?L=1. In the EU context, Spain is organised into 17 self-governing regions (Autonomous Communities), 2 autonomous cities (Ceuta and Melilla), 50 provinces and 8, 131 municipalities: https://portal.cor.europa.eu/divisionpowers/Pages/Spain-intro.aspx. Please be aware that for statistical purposes all EU regions are NUTS (Nomenclature of Territorial Units for Statistics). The NUTS classification subdivides the territory of the EU Member States into territorial units that favours administrative units already existing in the Member States. The NUTS are complemented at the lower level by local administrative units (LAU). In this paper, Spain's regions refer to Autonomous Communities.
https://ec.europa.eu/eurostat/documents/3859598/10967554/KS-GQ-20-092-EN-N.pdf/9d57ae79-3ee7-3c14-da3e-34726da385cf?t=1591285035000.

[49] INE's definition as in http://www.ine.es/nomen2/Metodologia.do?L=1 and https://www.boe.es/boe/dias/2015/03/24/pdfs/BOE-A-2015-3109.pdf

[50] https://www.ine.es/nomen2/AyudaResul.do

Our observation units will be **residential units.** Specifically, **people** that use the land of singular entities or alternatively municipalities for residential purposes.

Centroid of singular entity Geographical coordinates of the SE provided by the National Geographical Institute.

Centroid of municipality Centroid of the singular entity that holds the capital of the municipality as indicated by the National Geographical Institute.

Observation point Regarding the calculations on distances, we define an "*observation point*" as the distance between any two singular entities (alternatively two municipalities) within the same province. To calculate average distances at regional and national level we will consider the number of observation points within the same province, excluding additional couples of SE from different provinces and regions. Please notice that in a given province the number of observation points for SE is $\#_i(\#_i-1)/2$ and for municipalities $\mu_i(\mu_i-1)/2$.

Central Business District —CBD Singular entity holding the capital of the province.

α_{ij} Total land area of municipality j of province i[51]

α_i Total land area of province i

β_i Total land area of province i CBD

Δ^{ν}_{ij} Land area of type ν of municipality j of province i, with $\nu = \begin{cases} 0 \; for \; total \; land \; \left(\Delta^{\mathbf{0}}_{ij} = \alpha_{ij}\right) \\ 1 \; for \; urban \; land \\ 2 \; for \; built-up \; land \end{cases}$

Δ^{ν}_{i} Land area of province i of type ν ($\Delta^{\mathbf{0}}_{i} = \alpha_i$)

λ_I Diagonal of the axes-aligned 2-dimensional bounding box of province i (bounding box or bounding rectangle)[52]

$D_{adj}(i)$ Diagonal of province i: Adjusted diagonal of the bounding box

δ^{ν}_{ij} Crude(*) population density in municipality j of province i with $\nu = \begin{cases} 0 \; for \; total \; land \; - global \; density \\ 1 \; for \; urban \; land \; - urban \; density \\ 2 \; for \; built-up \; land \; - residential \; density \end{cases}$

$$\delta^{\nu}_{ij} = \frac{\pi_{ij}}{\Delta^{\nu}_{ij}}$$

δ^{ν}_{i} (*) Crude population density in province i with ν as before

$$\delta^{\nu}_{i} = \frac{\pi_{i}}{\Delta^{\nu}_{i}}$$

δ^{ν}_{iCBM} Crude population density in the municipality hosting the CBM of province I with ν as before

δ^{ν}_{R} Crude population density in Region R with ν as before

$d_i[j_1,j_2]$ (**) Straight-line distance between the centroids of municipalities j_1 and j_2 within province i

$d_i[k_1,k_2]$ (**) Straight-line distance between the centroids of singular entities k_1 and k_2 within province i

$d_{ij}[k_1,k_2]$ (**) Straight-line distance between the centroids of singular entities k_1 and k_2 of municipality j within province i

$\boldsymbol{R_n}$ Region n, $n \in \mathbf{N} = \{1 \dots 17\}$

$\{i \mid i \in \boldsymbol{R_n}\}$ Set of provinces in Region n

$\#\boldsymbol{R_n}$ Number of provinces in the Region $\boldsymbol{R_n}$

ES Spain

(*) We use the term "*crude"* to distinguish the ratio a province's population to its surface area from the population density of that province calculated as the weighted average of the population density of its municipalities. At municipal level, which is the basic level to measure density in this work, both concepts are the same.

(**) We notice that some authors consider additional metrics for the distance between two land uses, such as: $1/d_i[k1,k2]$ (Torres, 2017); $e^{-d_i[k1,k2]}$ and $ln(d_i[k1,k2])$ (Folch, 2012) which can be used and rescaled to fix the results to a particular range based on the research objectives (Campante and Do, 2009). In this work, we use original distances.

---

[51] Regarding the territorial extension, the Spanish statistical system provides only data at the municipal level. Therefore, no data are available regarding the surface area of singular entities.

[52] Following Ooi, B.C. (1993), the h-dimensional bounding boxes can be defined as a single dimensional array of h entries: $(I_1, I_2, \dots, I_h)$ where I_s (s = 1 to h) is a closed bounded interval [a, b] describing the extent of the spatial object along dimension s.

ANNEX III. TABLES

Annex III. Table I
Land uses and observation units included in the analyses in 2016

province	Number of singular entities	Number of municipalities	Population included
TOTAL	***55,861***	***8,102***	***46,206,955***
Almería	600	102	700,316
Cádiz	207	44	1,239,580
Córdoba	326	75	791,551
Granada	468	168	909,386
Huelva	204	79	519,172
Jaén	346	97	647,691
Málaga	238	100	1,622,733
Sevilla	232	105	1,939,405
Huesca	712	202	220,623
Teruel	333	236	136,945
Zaragoza	386	292	947,548
Asturias	5,809	78	1,039,875
Illes Balears	308	67	1,107,119
Palmas	509	34	1,097,227
SC Tenerife	559	53	998,950
Cantabria	921	102	582,176
Ávila	441	248	162,456
Burgos	1,137	371	352,228
León	1,385	211	473,157
Palencia	456	191	164,602
Salamanca	733	362	334,136
Segovia	379	209	154,806
Soria	483	183	89,999
Valladolid	263	225	523,634
Zamora	512	248	180,405
Albacete	305	87	392,114
Ciudad Real	163	102	506,743
Cuenca	330	238	201,030
Guadalajara	494	288	247,209
Toledo	339	204	686,277
Barcelona	1,218	311	5,521,013
Girona	972	221	750,404
Lleida	934	231	433,646
Tarragona	444	183	782,833
Alicante	282	141	1,835,689
Castellón	320	135	577,831
Valencia	452	265	2,534,502
Badajoz	275	164	676,652
Cáceres	319	219	400,103
Coruña	9,228	92	1,108,892
Lugo	8,817	67	335,173
Ourense	3,462	92	312,633
Pontevedra	5,819	60	937,103
Madrid	550	179	6,433,562
Murcia	836	45	1,461,376
Navarra	850	272	640,339
Álava	422	51	324,113
Bizkaia	550	111	1,146,731
Gipuzkoa	290	88	709,491
La Rioja	243	174	315,776

Source: INE and authors' own work based on the sources described in this paper.

Annex III. Table II (Start)

Basic data on straight-line distances between SE within the same province in 2016 It continues

Province	Simple average of straight-line distances between SE within the same province (Km)	Sum of straight-line distances between singular entities within the same province (Km)	Observation points per province (Number)	Maximum straight-line distance between SE (Km)	Inter-province coefficient of variation of the average straight-line distance CV*
TOTAL	***51.82***	***6,673,339,765***	***128,783,590***	***284.57***	***0.26***
Almería	50.71	9,111,999	179,700	149.60	0.55
Cádiz	51.80	1,104,445	21,321	116.80	0.50
Córdoba	63.19	3,347,301	52,975	175.50	0.57
Granada	60.71	6,633,867	109,278	198.21	0.56
Huelva	50.38	1,043,189	20,706	130.24	0.54
Jaén	65.87	3,931,524	59,685	169.61	0.56
Málaga	51.31	1,447,140	28,203	157.84	0.58
Sevilla	63.49	1,701,270	26,796	166.78	0.54
Huesca	58.09	14,704,367	253,116	172.03	0.50
Teruel	66.12	3,654,885	55,278	180.79	0.50
Zaragoza	71.85	5,338,676	74,305	208.63	0.54
Asturias	54.90	926,046,186	16,869,336	214.53	0.63
Illes Balears	83.80	3,961,723	47,278	284.57	0.75
Palmas	86.19	11,143,235	129,286	274.91	1.01
SC Tenerife	76.74	11,967,875	155,961	211.13	0.65
Cantabria	45.42	19,243,037	423,660	134.98	0.54
Ávila	47.79	4,636,167	97,020	124.09	0.52
Burgos	64.84	41,876,695	645,816	188.39	0.54
León	66.36	63,599,448	958,420	191.23	0.54
Palencia	49.88	5,175,056	103,740	142.93	0.57
Salamanca	51.34	13,774,368	268,278	166.1	0.53
Segovia	43.86	3,141,555	71,631	123.39	0.53
Soria	51.16	5,955,383	116,403	141.25	0.50
Valladolid	52.95	1,824,296	34,453	131.31	0.51
Zamora	57.28	7,492,812	130,816	174.75	0.52
Albacete	57.40	2,661,249	46,360	174.14	0.56
Ciudad Real	73.99	976,829	13,203	209.21	0.52
Cuenca	66.53	3,611,658	54,285	163.66	0.48
Guadalajara	59.61	7,258,946	121,771	165.67	0.53
Toledo	60.63	3,473,337	57,291	204.26	0.57
Barcelona	43.81	32,468,658	741,153	122.95	0.52
Girona	42.54	20,074,227	471,906	132.15	0.57
Lleida	58.78	25,610,056	435,711	172.55	0.53
Tarragona	42.64	4,193,452	98,346	142.27	0.67
Alicante	51.43	2,037,716	39,621	134.16	0.59
Castellón	43.71	2,230,804	51,040	126.01	0.55
Valencia	57.63	5,873,614	101,926	174.87	0.57
Badajoz	80.15	3,019,674	37,675	236.27	0.53
Cáceres	73.20	3,712,754	50,721	216.02	0.52
Coruña	53.78	2,289,682,943	42,573,378	177.51	0.55
Lugo	53.88	2,094,080,791	38,865,336	158.48	0.54
Ourense	39.14	234,511,055	5,990,991	129.88	0.58
Pontevedra	42.33	716,589,269	16,927,471	136.67	0.54
Madrid	48.54	7,328,529	150,975	137.68	0.50
Murcia	50.32	17,563,994	349,030	145.95	0.51
Navarra	43.43	15,670,785	360,825	154.58	0.56
Álava	30.93	2,747,880	88,831	92.76	0.53
Bizkaia	25.99	3,923,438	150,975	83.32	0.55
Gipuzkoa	24.40	1,022,408	41,905	76.64	0.55
La Rioja	39.42	1,159,201	29,403	119.88	0.58

Source: Authors' own work based on the sources described in this paper.

* National values correspond to the CV of provincial averages.

Annex III. Table II (Conclusion)

Basic data on straight-line distances between SE within the same province in 2016. Island correction Conclusion

ISLAND CORRECTION	Average of straight-line distances between SE within the same Island (Km)	Sum of straight-line distances between SE within the same Island (Km)	Number of observation points within the same Island (Number)	Maximum straight-line distance between SE within the same Island (Km)
Illes Ballears	***35.80***	***888,451***	***24,820***	***97.37***
Formentera	6.34	228	36	14.23
Ibiza	14.49	3,666	253	35.19
Mallorca	37.65	850,065	22,578	97.37
Menorca	17.66	34,491	1,953	45.54
Las Palmas	***19.96***	***1,544,649***	***77,386***	***101.28***
Fuerteventura	29.78	52,719	1,770	101.28
Gran Canaria	19.80	1,455,665	73,536	46.72
Lanzarote	17.43	36,264	2,080	54.18
SC Tenerife	***24.66***	***1,484,479***	***60,190***	***80.70***
Gomera	9.28	43,216	4,656	22.80
Hierro	7.83	822	105	21.75
La Palma	15.01	163,269	10,878	39.56
Santa Cruz de Tenerife	28.67	1,277,172	44,551	80.70

Source: Authors' own work based on the sources described in this paper.

* National values correspond to the CV of provincial averages.

Annex III. Table III (Start)

Sample size of observation points and sample estimates of travel distance and travel duration by provinces in 2016 — It continues

PROVINCE	Theoretical sample size	Effective sample size	Travel distance ψ (Km) Simple Average	Standard deviation *	CV*	Relative Error** for the estimate of the sample average (%)	Travel duration ζ (minutes) Simple Average	Standard deviation*	CV*	Relative Error** for the estimate of the sample average (%)
	n	$\tilde{n}$	$\bar{\psi}$	s_ψ	$\frac{s_\psi}{\bar{\psi}}$	$\hat{r} = \frac{z\,\widehat{CV}}{\tilde{n}^{0.5}}$	$\bar{\zeta}$	s_ζ	$\frac{s_\zeta}{\bar{\zeta}}$	$\hat{r} = \frac{z\,\widehat{CV}}{\tilde{n}^{0.5}}$
Total/Average	***196,600***	***191,702***	***82.84***	***19.41***	***0.23***	***1.77***	***77.82***	***40.34***	***0.52***	***1.57***
Almería	3,600	3,547	87.54	47.70	0.54	1.83	76.39	32.53	0.43	1.43
Cádiz	3,600	3,282	82.57	44.64	0.54	1.89	68.73	30.90	0.45	1.57
Córdoba	3,600	3,461	91.12	50.28	0.55	1.88	78.43	36.59	0.47	1.59
Granada	3,600	3,539	99.56	52.37	0.53	1.77	80.57	36.28	0.45	1.51
Huelva	3,600	3,326	77.26	41.92	0.54	1.88	72.39	33.76	0.47	1.62
Jaén	3,600	3,486	109.55	58.97	0.54	1.82	99.74	49.13	0.49	1.67
Málaga	3,600	3,372	84.21	46.54	0.55	1.90	76.46	35.53	0.46	1.60
Sevilla	3,600	3,355	88.81	44.80	0.50	1.74	77.58	33.85	0.44	1.51
Huesca	3,600	3,566	91.00	41.49	0.46	1.53	77.30	32.36	0.42	1.40
Teruel	3,600	3,475	100.21	47.46	0.47	1.61	83.89	37.79	0.45	1.53
Zaragoza	3,600	3,489	104.25	57.07	0.55	1.85	80.73	38.75	0.48	1.63
Asturias	6,000	5,999	87.27	52.08	0.60	1.54	75.02	39.14	0.52	1.35
Illes Balears	3,600	3,469	111.59	77.97	0.70	2.37	205.67	190.49	0.93	3.15
Palmas	3,600	3,556	131.52	120.21	0.91	3.07	243.06	260.61	1.07	3.60
SC Tenerife	3,600	3,552	128.14	77.01	0.60	2.02	232.71	171.29	0.74	2.47
Cantabria	3,600	3,586	75.96	40.32	0.53	1.77	63.82	30.12	0.47	1.58
Ávila	3,600	3,521	71.43	36.88	0.52	1.74	67.36	31.39	0.47	1.57
Burgos	3,600	3,589	94.06	46.83	0.50	1.66	80.49	34.23	0.43	1.42
León	3,600	3,592	104.08	55.81	0.54	1.79	85.56	39.53	0.46	1.54
Palencia	3,600	3,535	68.31	36.81	0.54	1.81	55.51	24.97	0.45	1.51
Salamanca	3,600	3,568	73.64	38.19	0.52	1.74	62.48	26.29	0.42	1.41
Segovia	3,600	3,500	60.56	32.72	0.54	1.83	57.86	27.83	0.48	1.63
Soria	3,600	3,524	70.71	33.34	0.47	1.59	58.67	25.07	0.43	1.44
Valladolid	3,600	3,426	70.82	35.63	0.50	1.72	56.86	25.25	0.44	1.52
Zamora	3,600	3,548	80.46	41.01	0.51	1.71	66.85	29.49	0.44	1.48
Albacete	3,600	3,445	89.09	46.08	0.52	1.76	78.95	35.63	0.45	1.54
Ciudad Real	3,600	3,151	99.97	48.99	0.49	1.75	84.65	39.22	0.46	1.65
Cuenca	3,600	3,471	98.56	45.67	0.46	1.57	83.61	35.90	0.43	1.46
Guadalajara	3,600	3,547	89.02	44.95	0.50	1.70	75.26	33.59	0.45	1.50
Toledo	3,600	3,474	85.19	45.47	0.53	1.81	65.85	30.26	0.46	1.56
Barcelona	3,600	3,576	66.36	30.96	0.47	1.56	54.24	19.77	0.36	1.22
Girona	3,600	3,567	67.88	40.88	0.60	2.02	64.48	32.49	0.50	1.69
Lleida	3,600	3,579	97.54	48.41	0.50	1.66	93.33	42.05	0.45	1.51
Tarragona	3,600	3,523	61.87	36.86	0.60	2.01	52.82	25.38	0.48	1.62
Alicante	3,600	3,451	74.11	40.47	0.55	1.86	58.58	25.43	0.43	1.48
Castellón	3,600	3,445	71.62	40.02	0.56	1.90	67.79	31.63	0.47	1.59
Valencia	3,600	3,545	83.48	46.06	0.55	1.85	64.72	31.20	0.48	1.62
Badajoz	3,600	3,407	107.14	56.73	0.53	1.81	87.67	41.45	0.47	1.62
Cáceres	3,600	3,478	112.09	55.83	0.50	1.69	91.27	39.92	0.44	1.48
Coruña	10,000	9,999	80.19	41.93	0.52	1.05	69.27	28.62	0.41	0.83
Lugo	9,000	8,998	82.90	43.14	0.52	1.10	73.09	30.16	0.41	0.87
Ourense	3,600	3,599	65.62	38.40	0.59	1.95	62.41	29.06	0.47	1.55
Pontevedra	6,000	5,993	71.48	39.58	0.55	1.43	60.49	25.68	0.42	1.10
Madrid	3,600	3,555	70.88	33.77	0.48	1.60	58.14	21.47	0.37	1.24
Murcia	3,600	3,582	72.06	35.40	0.49	1.64	60.28	25.37	0.42	1.41
Navarra	3,600	3,580	67.80	34.86	0.51	1.72	57.16	26.04	0.46	1.52
Álava	3,600	3,525	46.24	23.45	0.51	1.71	42.47	19.09	0.45	1.51
Bizkaia	3,600	3,544	42.13	22.32	0.53	1.78	43.40	19.73	0.45	1.53
Gipuzkoa	3,600	3,442	40.73	20.21	0.50	1.69	36.67	13.83	0.38	1.29
La Rioja	3,600	3,363	66.41	36.13	0.54	1.88	57.39	26.09	0.45	1.57

Source: Authors' own work based on Cochran, W.G. (1984) and the sources described in this paper.
Microsoft Bing Maps accessed with Distance Matrix API between Mach and June 2019.

* National values correspond to inter-province standard deviation and CV of provincial averages.

** National values correspond to the average of provinces' relative errors.

z = 2 Value of the N (0,1) distribution that leaves on its right a 0.025 probability (confidence level a α = 0.95).

Annex III. Table III (Conclusion)

Sample size of observation points and sample estimates of travel distance and duration by provinces in 2016. Island correction Conclusion

ISLAND CORRECTION	Travel distances ψ (Km)				Travel durations ζ (minutes)			
	Simple Average	Standard deviation*	CV*	Relative Error** for the estimate of the sample average (%)	Simple Average	Standard deviation*	CV*	Relative Error** for the estimate of the sample average (%)
ILLES BALEARS	**40.47**				**38.78**			
Formentera	8.08	4.56	0.56	0.00	12.93	6.64	0.51	0.00
Ibiza	19.40	9.16	0.47	0.00	24.61	10.56	0.43	0.00
Mallorca	52.25	25.50	0.49	2.38	48.51	21.23	0.44	2.13
Menorca	25.43	16.43	0.65	10.76	30.75	16.35	0.53	8.86
LAS PALMAS	**27.32**				**41.15**			
Fuerteventura	41.39	25.12	0.61	0.00	49.10	31.12	0.63	0.00
Gran Canaria	43.32	23.41	0.54	2.39	50.10	20.88	0.42	1.84
Lanzarote	24.42	13.88	0.57	0.00	32.24	17.83	0.55	0.00
SC TENERIFE	**41.94**				**40.73**			
Gomera, La	25.92	10.90	0.42	8.13	38.21	14.42	0.38	7.30
Hierro, El	15.18	8.69	0.57	11.17	18.77	10.06	0.54	10.46
Palma, La	31.49	19.09	0.61	7.53	42.52	22.19	0.52	6.49
Tenerife	49.35	27.72	0.56	3.60	42.89	18.92	0.44	2.82

Source: Authors' own work based on Cochran W.G. (1984) and the sources described in this paper.

* National values correspond to inter-province standard deviation and CV of provincial averages.

** National values correspond to the average of provinces' relative errors.

Annex III. Table IV (Start)

Ratio estimates of travel distance and duration by provinces in 2016

It continues

PROVINCE	Ratio estimate of average travel distance (Km)			Ratio estimate of average travel duration (minutes)		
	Ratio 1 * sample estimate	Ratio estimate of average travel distance (Km)	Relative error* of the ratio estimate of travel distance (%)	Ratio 2* sample estimate (min/Km)	Ratio estimate of average travel duration (minutes)	Relative error* of the ratio estimate of travel duration (%)
TOTAL	***1.56***	***80.72***	***0.3640***	***1.36***	***70.52***	***0.6100***
Almería	1.74	88.37	0.7512	1.52	77.11	1.2127
Cádiz	1.61	83.47	0.7648	1.34	69.48	0.9255
Córdoba	1.45	91.54	0.4637	1.25	78.79	0.7574
Granada	1.65	100.43	0.8129	1.34	81.28	1.1212
Huelva	1.54	77.49	0.5414	1.44	72.60	0.6710
Jaén	1.67	110.15	0.7558	1.52	100.29	1.0979
Málaga	1.61	82.60	0.6447	1.46	75.00	1.2256
Sevilla	1.38	87.54	0.5371	1.20	76.47	1.0510
Huesca	1.59	92.21	0.5756	1.35	78.33	0.7423
Teruel	1.50	99.15	0.4425	1.26	83.00	0.6530
Zaragoza	1.47	105.95	0.5098	1.14	82.05	0.6666
Asturias	1.60	87.80	0.4949	1.37	75.47	0.8648
Illes Balears	1.35	113.52	0.6091	2.50	209.22	1.0464
Palmas	1.56	134.34	0.7354	2.88	248.26	1.0130
SC Tenerife	1.64	126.13	0.7316	2.98	229.06	0.9740
Cantabria	1.68	76.17	0.7258	1.41	64.00	1.0511
Ávila	1.49	71.13	0.6549	1.40	67.07	0.7581
Burgos	1.46	94.65	0.4647	1.25	80.99	0.8286
León	1.54	102.24	0.5597	1.27	84.04	0.9105
Palencia	1.37	68.16	0.4529	1.11	55.39	0.7705
Salamanca	1.42	72.83	0.5822	1.20	61.79	0.8345
Segovia	1.38	60.54	0.4251	1.32	57.85	0.5178
Soria	1.38	70.71	0.4421	1.15	58.67	0.6066
Valladolid	1.35	71.32	0.4037	1.08	57.26	0.5302
Zamora	1.42	81.25	0.5417	1.18	67.50	0.8129
Albacete	1.55	88.94	0.7438	1.37	78.82	1.0458
Ciudad Real	1.35	99.60	0.4542	1.14	84.34	0.6799
Cuenca	1.48	98.38	0.4490	1.25	83.46	0.6715
Guadalajara	1.50	89.28	0.4666	1.27	75.48	0.6822
Toledo	1.40	85.17	0.4346	1.09	65.84	0.8192
Barcelona	1.52	66.51	0.5503	1.24	54.36	0.9136
Girona	1.60	68.20	0.5991	1.52	64.79	0.9125
Lleida	1.68	98.92	0.6932	1.61	94.65	0.9582
Tarragona	1.41	60.02	0.5494	1.20	51.24	1.0166
Alicante	1.47	75.48	0.5555	1.16	59.66	1.0508
Castellón	1.65	72.31	0.6349	1.57	68.44	1.0465
Valencia	1.46	83.90	0.5112	1.13	65.04	0.7829
Badajoz	1.35	107.82	0.4009	1.10	88.22	0.5993
Cáceres	1.53	112.17	0.5214	1.25	91.34	0.7869
Coruña	1.50	80.71	0.2725	1.30	69.72	0.5025
Lugo	1.54	83.12	0.3139	1.36	73.28	0.5187
Ourense	1.68	65.76	0.5448	1.60	62.54	0.9020
Pontevedra	1.70	72.17	0.4492	1.44	61.08	0.6321
Madrid	1.46	70.75	0.5718	1.20	58.03	0.8651
Murcia	1.43	71.82	0.4849	1.19	60.08	0.8657
Navarra	1.59	68.89	0.6579	1.34	58.07	1.0046
Álava	1.49	46.23	0.6021	1.37	42.46	0.8241
Bizkaia	1.62	42.22	0.5632	1.67	43.49	0.8921
Gipuzkoa	1.67	40.69	0.7367	1.50	36.64	1.1503
La Rioja	1.69	66.77	0.9178	1.46	57.70	1.2635

Source: Authors' own work based on the sources described in this paper.
* National values correspond to the provinces' weighted average (weights = share of total straight-line distance between SE within a province).
Ratio 1 = Travel distance/straight-line distance.
Ratio 2 = Travel duration/straight-line distance.

Annex III. Table IV (Conclusion)

Ratio estimates of travel distance and duration by provinces. Island correction in 2016 Conclusion

ISLAND CORRECTION	Ratio estimate of average travel distance (Km)			Ratio estimate of average travel duration (minutes)		
	Ratio 1* estimate	Ratio estimate of average travel distance (Km)	Relative error* of the ratio estimate of travel distance (%)	Ratio 2* estimate (min/Km)	Ratio estimate of average travel duration (minutes)	Relative error* of the ratio estimate of travel duration (%)
ILLES BALEARS	**1.40**	**50.20**	**0.9399**	**1.32**	**47.27**	**1.3208**
Formentera	1.28	8.08	5.7824	2.04	12.93	7.0866
Ibiza	1.34	19.40	1.0306	1.70	24.61	1.4584
Mallorca	1.40	52.66	0.8522	1.30	48.89	1.2135
Menorca	1.50	26.54	3.0600	1.82	32.09	3.9119
LAS PALMAS	**1.02**	**20.38**	**3.3797**	**2.21**	**44.20**	**1.9602**
Fuerteventura	1.34	40.01	0.5665	1.65	49.10	1.7105
Gran Canaria	1.00	19.80	3.5457	2.24	44.42	1.9781
Lanzarote	1.40	24.42	0.8047	1.85	32.24	1.6041
SC TENERIFE	**1.77**	**43.59**	**2.3924**	**1.70**	**41.93**	**2.3948**
Gomera, La	3.03	28.17	5.9228	4.47	41.52	6.8442
Hierro, El	1.94	15.18	8.0637	2.40	18.77	7.9916
Palma, La	2.09	31.31	3.8533	2.82	42.27	3.1887
Tenerife	1.68	48.26	2.0826	1.46	41.95	2.1391

Source: Authors' own work based on the sources described in this paper.

* Values at the provincial level correspond to the islands' weighted average (weights = total straight-line distance between SE within an island).

Annex III. Table V

Basic data on provinces dimensions in 2016

Province	Square root of the surface area of the province (Km)*	Province maximum distance from north to south (Km)*	Province maximum distance from east to west (Km)*	Maximum straight-line distance between SE (Km)*	Diagonal of the axes-aligned 2-dimensional bounding box (Km)*	Diagonal of the province D_{adj} (Km)*	V (Km)	VI_{max} (Km)	V_{max} (Km)
TOTAL	***97.44***	***133.60***	***150.38***	***162.34***	***202.99***	***202.99***	***17.68***	***50.66***	***49.94***
Almería	93.66	136	133	149.60	190.22	192.78	20.77	48.09	46.02
Cádiz	86.23	116	121	116.80	167.62	164.75	23.92	41.14	42.56
Córdoba	117.35	171	140	175.50	221.00	224.52	21.75	55.63	57.58
Granada	112.43	155	182	198.21	239.06	242.17	18.39	61.11	55.99
Huelva	100.64	156	122	130.24	198.04	191.01	18.32	51.16	50.00
Jaén	116.14	128	160	169.61	204.90	208.71	23.06	51.87	55.45
Málaga	85.49	107	165	157.84	196.66	192.80	17.32	49.13	44.71
Sevilla	118.47	151	167	166.78	225.14	226.37	15.90	57.90	59.54
Huesca	125.01	173	139	172.03	221.92	224.60	27.33	56.35	58.90
Teruel	121.64	167	178	180.79	244.08	244.11	32.32	62.78	58.51
Zaragoza	131.43	202	215	208.63	295.01	282.39	12.43	75.95	70.12
Asturias	102.97	86	216	214.53	232.49	214.39	16.13	58.23	53.00
Illes Balears	70.65	161	266	284.57	310.93	300.43	33.56	75.19	67.15
Palmas	63.79	189	235	274.91	301.57	290.62	38.69	73.65	65.00
SC Tenerife	58.10	135	200	211.13	241.30	242.73	24.32	59.04	59.11
Cantabria	72.53	84	138	134.98	161.55	159.83	13.59	40.09	38.25
Ávila	89.72	120	132	124.09	178.39	175.55	20.30	45.79	44.30
Burgos	118.42	193	149	188.39	243.82	242.83	22.93	62.35	58.10
León	124.77	135	189	191.23	232.26	235.79	24.48	58.36	59.79
Palencia	89.74	145	94	142.93	172.80	169.21	18.48	43.22	44.08
Salamanca	111.13	116	155	166.10	193.60	195.69	18.57	49.35	53.92
Segovia	82.77	106	127	123.39	165.42	164.35	15.88	42.24	41.88
Soria	101.51	122	147	141.25	191.03	189.51	18.20	48.67	49.84
Valladolid	90.06	136	129	131.31	187.45	184.94	9.12	47.92	51.08
Zamora	102.77	124	150	174.75	194.62	202.54	20.86	49.87	50.19
Albacete	122.14	155	171	174.14	230.79	233.41	24.07	57.73	59.23
Ciudad Real	140.76	136	208	209.21	248.52	253.09	29.10	62.19	66.24
Cuenca	130.92	159	172	163.66	234.23	234.70	30.86	60.23	61.76
Guadalajara	110.30	130	168	165.67	212.42	213.17	13.72	54.28	56.85
Toledo	123.98	118	212	204.26	242.63	237.84	28.81	59.47	58.81
Barcelona	87.91	125	118	122.95	171.90	170.08	12.76	44.43	45.68
Girona	76.87	96	131	132.15	162.41	163.44	18.47	41.04	38.77
Lleida	110.30	176	128	172.55	217.62	217.72	21.74	55.17	54.28
Tarragona	79.39	116	125	142.27	170.53	175.55	19.66	43.91	40.34
Alicante	76.27	115	114	134.16	161.93	166.91	21.20	41.29	38.83
Castellón	81.44	119	115	126.01	165.49	166.59	15.33	42.20	42.07
Valencia	103.97	170	129	174.87	213.40	215.37	14.84	54.18	53.17
Badajoz	147.53	168	229	236.27	284.02	291.45	32.45	71.93	70.21
Cáceres	140.95	161	221	216.02	273.43	276.13	31.02	69.07	67.09
Coruña	89.16	146	131	177.51	196.16	204.23	22.74	50.05	45.16
Lugo	99.29	159	96	158.48	185.73	180.83	24.34	45.26	47.48
Ourense	85.28	86	134	129.88	159.22	152.44	17.27	39.01	41.66
Pontevedra	67.04	109	89	136.67	140.72	150.36	13.64	35.24	35.48
Madrid	89.56	142	130	137.68	192.52	190.12	9.26	49.54	50.53
Murcia	106.37	154	144	145.95	210.84	207.91	20.74	53.81	51.64
Navarra	99.00	156	145	154.58	212.98	210.16	19.62	49.97	49.87
Álava	54.47	82	86	92.76	118.83	123.66	6.78	29.32	33.98
Bizkaia	47.07	53	84	83.32	99.32	98.03	7.46	24.73	26.81
Gipuzkoa	43.69	55	71	76.64	89.81	92.01	11.95	22.37	22.77
La Rioja	70.91	80	119	119.88	143.39	141.92	13.91	35.87	36.26

Source: Authors' own work based on the sources described in this paper.

* TOTAL equals the simple average of the diagonals of the provinces.

Annex III. Table VI

Basic data on land area by province in 2016

PROVINCES	Surface (Km²) GIS	Surface area of the SE holding the CBD (Km²) GIS	Urban area (Km²) CS & SIU	Built–up area (Km²) CS & SIU	Not built-up area (Km²) CS & SIU	Urban area /Total %	Built-up area /Total %	Not built-up area /Total %
TOTAL	**504,688**	**15,690**	**11,325**	**6,870**	**4,456**	**2.24**	**1.36**	**0.88**
Almería	8,773	296	184	97	87	2.10	1.11	0.99
Cádiz	7,436	12	248	172	76	3.34	2.32	1.02
Córdoba	13,771	1,255	174	124	50	1.26	0.90	0.36
Granada	12,640	81	205	118	87	1.62	0.93	0.68
Huelva	10,128	151	137	78	59	1.35	0.77	0.58
Jaén	13,489	424	133	87	46	0.99	0.65	0.34
Málaga	7,308	395	349	210	140	4.78	2.87	1.91
Sevilla	14,036	141	333	241	92	2.38	1.72	0.66
Huesca	15,627	161	81	53	29	0.52	0.34	0.18
Teruel	14,796	440	56	35	20	0.38	0.24	0.14
Zaragoza	17,275	974	224	132	92	1.30	0.77	0.53
Asturias	10,604	187	285	166	119	2.69	1.57	1.12
Balears	4,992	209	236	172	63	4.72	3.46	1.27
Palmas	4,070	103	234	115	119	5.75	2.82	2.94
SC Tenerife	3,375	150	189	114	75	5.61	3.38	2.23
Cantabria	5,261	36	179	125	54	3.41	2.38	1.03
Ávila	8,050	231	88	52	36	1.10	0.65	0.45
Burgos	14,023	107	170	93	77	1.22	0.66	0.55
León	15,568	39	223	122	100	1.43	0.79	0.64
Palencia	8,053	95	69	48	21	0.85	0.60	0.26
Salamanca	12,349	39	123	74	49	0.99	0.60	0.39
Segovia	6,851	164	110	63	47	1.61	0.92	0.69
Soria	10,303	272	49	27	22	0.47	0.26	0.21
Valladolid	8,111	197	162	98	65	2.00	1.20	0.80
Zamora	10,561	149	94	58	36	0.89	0.55	0.34
Albacete	14,918	1,127	109	73	36	0.73	0.49	0.24
Ciudad Real	19,813	285	168	101	67	0.85	0.51	0.34
Cuenca	17,140	911	89	60	30	0.52	0.35	0.17
Guadalajara	12,167	235	175	79	96	1.44	0.65	0.79
Toledo	15,370	232	359	199	160	2.34	1.30	1.04
Barcelona	7,729	98	730	490	240	9.45	6.34	3.11
Girona	5,909	39	265	173	92	4.49	2.92	1.57
Lleida	12,166	212	135	80	56	1.11	0.65	0.46
Tarragona	6,303	55	286	170	116	4.54	2.70	1.83
Alicante	5,816	201	480	306	174	8.25	5.27	2.98
Castellón	6,632	109	200	108	92	3.01	1.62	1.39
Valencia	10,811	139	438	293	145	4.05	2.71	1.34
Badajoz	21,766	1,440	200	135	65	0.92	0.62	0.30
Cáceres	19,868	1,750	121	77	44	0.61	0.39	0.22
Coruña	7,950	38	480	272	209	6.04	3.42	2.62
Lugo	9,858	330	137	77	60	1.39	0.78	0.61
Ourense	7,273	85	105	62	43	1.44	0.85	0.60
Pontevedra	4,495	118	327	185	142	7.29	4.13	3.16
Madrid	8,022	606	964	570	394	12.02	7.11	4.92
Murcia	11,314	886	551	270	281	4.87	2.39	2.48
Navarra**	9,801	25	260	133	127	2.65	1.36	1.29
Alava**	2,967	277	104	73	31	3.50	2.46	1.04
Bizkaia**	2,216	41	131	94	37	5.92	4.24	1.68
Gipuzkoa**	1,909	61	90	67	23	4.71	3.53	1.18
La Rioja	5,028	79	82	48	35	1.64	0.95	0.69

Source: Authors' own work based on the sources described in this paper.

****** Estimates for urban surface area related data. The National Cadastral Services does not provide data for País Vasco and Navarra because of their specific fiscal system. For these regions, we have estimated the data based on the Urban Information System (SIU) published by the Ministry of Development and the regional Offices of Statistics.

Urban and built-up surface: The urban land and built-up area correspond to the respective concepts for cadastral purposes. That is, the one considered in accordance with the cadastral legislation, article 7.2 of the Consolidated Text of the Real Estate Cadastre Law (http://www.catastro.meh.es/documentos/estadisticas_Metodologia_Catastro_2015.pdf and http://www.catastro.minhap.gob.es/documentos/estadisticas/Nota%20metodol%C3%B3gica%20Estadistica%20ocupaci%C3%B3n.pdf). The estimations for País Vasco and Navarra's provinces are based on the concepts used by the SIU (https://apps.fomento.gob.es/CVP/handlers/pdfhandler.ashx?idpub=BAW055).

Note: Please note that in this table population and surface area correspond to total in official registries, including that of SE dropped to build the database use in this work.

Annex III. Table VII

Nuclei (*) by provinces in 2016

Province	Nuclei SE Based	Nuclei MUN Based
TOTAL	***673***	***731***
Almería	11	12
Cádiz	23	21
Córdoba	11	14
Granada	17	22
Huelva	13	14
Jaén	13	15
Málaga	26	19
Sevilla	37	26
Huesca	5	5
Teruel	2	2
Zaragoza	6	6
Asturias	8	20
Illes Balears	18	24
Palmas	21	22
SC Tenerife	15	20
Cantabria	10	10
Ávila	1	1
Burgos	3	3
León	5	6
Palencia	1	1
Salamanca	4	4
Segovia	1	1
Soria	1	1
Valladolid	4	4
Zamora	2	2
Albacete	6	6
Ciudad Real	13	13
Cuenca	2	2
Guadalajara	3	4
Toledo	11	13
Barcelona	72	80
Girona	12	20
Lleida	5	5
Tarragona	16	15
Alicante	35	36
Castellón	10	11
Valencia	56	51
Badajoz	9	9
Cáceres	4	2
Coruña	9	21
Lugo	2	5
Ourense	4	6
Pontevedra	6	24
Madrid	53	49
Murcia	37	29
Navarra	10	10
Álava	2	3
Bizkaia	20	20
Gipuzkoa	14	18
La Rioja	4	4

Source: authors own work based the sources described in this paper.

(*) The traditional Spanish statistical classification determines as urban "*the set of population entities with 10,001 or more inhabitants*." Please refer to Ministry of Development of Spain (2018). Áreas urbanas en España 2018. Constitución, Cuarenta años de las ciudades españolas. Retrieved from: https://apps.fomento.gob.es/CVP/handlers/pdfhandler.ashx?idpub=BAW058. We have particularised it for SE and MUN.

Annex III. Table VIII

Crude population density in Spain's provinces by type of surface area in 2016

2016	Population	Surface area (Km²)	Urban area (Km²)	Built-Up area (Km²)	Total density Inhabitants per Km²	Urban density Inhabitants per Km²	Residential density Inhabitants per Km²
Total Nacional*	**46,386,463**	**504,688**	**11,325**	**6,870**	**92**	**4,096**	**6,752**
Almería	704,297	8,773	184	97	80	3,823	7,240
Cádiz	1,239,889	7,436	248	172	167	4,991	7,195
Córdoba	791,610	13,771	174	124	57	4,551	6,395
Granada	915,392	12,640	205	118	72	4,473	7,749
Huelva	519,596	10,128	137	78	51	3,789	6,620
Jaén	648,250	13,489	133	87	48	4,874	7,439
Málaga	1,629,298	7,308	349	210	223	4,662	7,765
Sevilla	1,939,775	14,036	333	241	138	5,817	8,042
Huesca	221,079	15,627	81	53	14	2,715	4,199
Teruel	136,977	14,796	56	35	9	2,465	3,904
Zaragoza	950,507	17,275	224	132	55	4,237	7,182
Asturias	1,042,608	10,604	285	166	98	3,655	6,266
Illes Balears	1,107,220	4,992	236	172	222	4,698	6,419
Palmas	1,097,800	4,070	234	115	270	4,690	9,579
SC Tenerife	1,004,124	3,375	189	114	297	5,307	8,806
Cantabria	582,206	5,261	179	125	111	3,246	4,642
Ávila	162,514	8,050	88	52	20	1,841	3,105
Burgos	360,995	14,023	170	93	26	2,118	3,872
León	473,604	15,568	223	122	30	2,128	3,875
Palencia	164,644	8,053	69	48	20	2,397	3,427
Salamanca	335,985	12,349	123	74	27	2,740	4,539
Segovia	155,652	6,851	110	63	23	1,414	2,482
Soria	90,040	10,303	49	27	9	1,841	3,361
Valladolid	523,679	8,111	162	98	65	3,226	5,359
Zamora	180,406	10,561	94	58	17	1,924	3,131
Albacete	392,118	14,918	109	73	26	3,611	5,386
Ciudad Real	506,888	19,813	168	101	26	3,024	5,021
Cuenca	201,071	17,140	89	60	12	2,250	3,379
Guadalajara	252,882	12,167	175	79	21	1,447	3,215
Toledo	688,672	15,370	359	199	45	1,917	3,454
Barcelona	5,542,680	7,729	730	490	717	7,592	11,320
Girona	753,576	5,909	265	173	128	2,843	4,365
Lleida	434,041	12,166	135	80	36	3,208	5,458
Tarragona	792,299	6,303	286	170	126	2,772	4,652
Alicante	1,836,459	5,816	480	306	316	3,826	5,994
Castellón	579,245	6,632	200	108	87	2,899	5,387
Valencia	2,544,264	10,811	438	293	235	5,806	8,675
Badajoz	684,113	21,766	200	135	31	3,416	5,067
Cáceres	403,665	19,868	121	77	20	3,334	5,262
Coruña	1,122,799	7,950	480	272	141	2,338	4,133
Lugo	336,527	9,858	137	77	34	2,450	4,375
Ourense	314,853	7,273	105	62	43	2,999	5,119
Pontevedra	944,346	4,495	327	185	210	2,884	5,092
Madrid	6,466,996	8,022	964	570	806	6,705	11,343
Murcia	1,464,847	11,314	551	270	129	2,656	5,418
Navarra	640,647	9,801	260	133	65	2,467	4,813
Álava	324,126	2,967	104	73	109	3,119	4,438
Bizkaia	1,147,576	2,216	131	94	518	8,745	12,204
Gipuzkoa	717,832	1,909	90	67	376	7,977	10,647
La Rioja	315,794	5,028	82	48	63	3,830	6,633
Maximum					806	8,745	12,204
Minimum					9	1,414	2,482
Standard deviation					142	1801	2639
CV					1.55	0.44	0.39

Source: Authors' own work based on the sources described in this paper. Base year = 2016.

* Please notice that, due to the definition of crude density, the mean and standard deviation at national level of the provincial distribution should be weighted by the surfaces of the provinces.

Annex III. Table IX

SE and population excluded from the analysis due to incidence type I

SE in the INE's database with no information in the IGN's database

PROVINCE	# SE EXCLUDED	INHABITANTS IN SE EXCLUDED	POPULATION EXCLUDED (%)
TOTAL	**189**	**6,828**	**0.015**
Pontevedra	183	6,071	
Murcia	6	757	

Source: Authors' own work based on the sources described in this paper.

Annex III. Table X

SE and population excluded due to incidence type II

SE in the IGN's database with no information on latitude and longitude (value zero)

PROVINCE	# SE EXCLUDED	INHABITANTS IN SE EXCLUDED	POPULATION EXCLUDED (%)
TOTAL	**663**	**14,965**	**0.032**
Almería	15	358	
Cádiz	1	94	
Granada	2	124	
Huesca	4	21	
Asturias	259	1,633	
SC Tenerife	1	1	
Cantabria	1	1	
Salamanca	3	9	
Segovia	2	46	
Ciudad Real	7	137	
Toledo	3	1,530	
Barcelona	25	792	
Girona	46	1,632	
Lleida	18	203	
Tarragona	10	352	
Castellón	20	725	
Valencia	1	25	
Badajoz	1	2,544	
Coruña	101	1,691	
Lugo	66	825	
Ourense	25	473	
Pontevedra	34	878	
Murcia	3	391	
Navarra	9	255	
Álava	1	13	
Bizkaia	5	212	

Source: Authors' own work based on the sources described in this paper.

Annex III. Table XI

SE and population excluded due to incidence type III

SE not Population Units during the whole period

PROVINCE	# SE EXCLUDED	INHABITANTS	POPULATION EXCLUDED (%)
TOTAL	**1,628**	**157,379**	**0.338**
Almería	18	3,623	
Cádiz	5	215	
Córdoba	11	59	
Granada	14	5,882	
Huelva	12	424	
Jaén	14	559	
Málaga	5	6,530	
Sevilla	5	370	
Huesca	40	389	
Teruel	10	32	
Zaragoza	8	2,959	
Asturias	161	1,100	
Illes Balears	6	101	
Palmas	6	573	
SC Tenerife	12	5,173	
Cantabria	4	29	
Ávila	6	58	
Burgos	20	8,767	
León	9	447	
Palencia	7	42	
Salamanca	70	1,840	
Segovia	8	800	
Soria	10	41	
Valladolid	2	45	
Zamora	1	1	
Albacete	1	4	
Ciudad Real	3	8	
Cuenca	4	41	
Guadalajara	21	5,673	
Toledo	9	865	
Barcelona	37	20,814	
Girona	24	1,515	
Lleida	25	192	
Tarragona	20	9,114	
Alicante	5	770	
Castellón	20	686	
Valencia	14	9,737	
Badajoz	6	4,917	
Cáceres	7	3,562	
Coruña	532	12,216	
Lugo	177	501	
Ourense	50	1,747	
Pontevedra	52	258	
Madrid	78	33,434	
Murcia	37	2,323	
Navarra	19	53	
Álava	0	0	
Bizkaia	11	531	
Gipuzkoa	8	8,341	
La Rioja	4	18	

Source: Authors' own work based on the sources described in this paper.

Annex III. Table XII

SE and population excluded due to incidence type IV

The distance from the SE to the rest of all SE is disproportionately high in relation to size of the territory

PROVINCE	# SE EXCLUDED	INHABITANTS	POPULATION EXCLUDED (%)
TOTAL	**13**	**336**	**0.0007**
Málaga	1	35	
Huesca	2	46	
Barcelona	3	61	
Girona	2	25	
Castellón	1	3	
Lugo	2	28	
Pontevedra	1	36	
Bizkaia	1	102	

Source: Authors' own work based on the sources described in this paper.

Annex III. Table XIII

SE and population included in the calculations of dispersion indicators (base year = 2016)

Province	# SE INCLUDED	# SE NOT INCLUDED		
		TOTAL NOT INCLUDED	OF WHICH: # SE INCIDENCES	OF WHICH: # SE with POP = 0
TOTAL	**55,861**	**5,834**	**2,497**	**3,337**
Almería	600	64	33	31
Cádiz	207	10	6	4
Córdoba	326	22	11	11
Granada	468	16	16	0
Huelva	204	23	12	11
Jaén	346	33	14	19
Málaga	238	6	6	0
Sevilla	232	8	5	3
Huesca	712	77	46	31
Teruel	333	24	10	14
Zaragoza	386	29	8	21
Asturias	5,809	1,135	420	715
Illes Balears	308	9	6	3
Palmas	509	21	6	15
SC Tenerife	559	19	13	6
Cantabria	921	10	5	5
Ávila	441	18	6	12
Burgos	1,137	42	20	22
León	1,385	17	9	8
Palencia	456	13	7	6
Salamanca	733	255	73	182
Segovia	379	10	10	0
Soria	483	20	10	10
Valladolid	263	9	2	7
Zamora	512	2	1	1
Albacete	305	11	1	10
Ciudad Real	163	15	10	5
Cuenca	330	6	4	2
Guadalajara	494	32	21	11
Toledo	339	17	12	5
Barcelona	1,218	98	65	33
Girona	972	104	72	32
Lleida	934	89	43	46
Tarragona	444	39	30	9
Alicante	282	7	5	2
Castellón	320	70	41	29
Valencia	452	24	15	9
Badajoz	275	12	7	5
Cáceres	319	16	7	9
Coruña	9,228	1,262	633	629
Lugo	8,817	997	245	752
Ourense	3,462	230	75	155
Pontevedra	5,819	429	270	159
Madrid	550	234	78	156
Murcia	836	93	46	47
Navarra	850	100	28	72
Álava	422	6	1	5
Gipuzkoa	290	11	8	3
Bizkaia	550	21	17	4
La Rioja	243	15	4	11
Ceuta		3	3	
Melilla		1	1	

Pro-memoria: Population excluded because of incidences

Total Population excluded because of incidences	Total Population excluded because of incidences (%)	Population in Ceuta and Melilla (%)	TOTAL Population excluded (%)
179,508	**0.3856**	**0.3660**	**0.7516**

Source: Authors' own work based on the sources described in this paper.

ANNEX IV. INDICATORS FORMULATION: THE STATISTICAL TOOLBOX AND RATIONALE

Proximity indicators

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$PROXS_{SE1a}$	***Inverse of the simple average of straight-line distances*** between singular entities within province i.	$PROXS_{SE1a}(i) = \frac{\#_i(\#_i - 1)}{2}[\mathbf{1}' \times \tau(i) \times \mathbf{1}]^{-1} \quad [1a]$ Where: $\mathbf{1}$ is a column vector $\#_i \times 1$ with 1 in all entries. $\boldsymbol{\tau(i)}$ is a $\#_i \times \#_i$ triangular distance matrix whose entry d_{k1k2} is the straight-line distance between the singular entities k_1 and k_2 within province i ($d_i[k_1,k_2]$) for under-diagonal entries and zero for the diagonal and upper-diagonal entries. In detail: $PROXS_{SE1a}(i) = \left[\frac{\sum_{k_1,k_2=1, k_1<k_2}^{\#_i} d_i[k_1,k_2]}{\#_i(\#_i-1)/2}\right]^{-1}$	Regional $PROXS_{SE1a}(R_n) = \left[\sum_{i \in R_n} \frac{\#_i(\#_i-1)}{2}\right]\left[\sum_{i \in R_n} \mathbf{1}' \times \tau(i) \times \mathbf{1}\right]^{-1}$ National $PROXS_{SE1a}(ES) = \left[\sum_{i=1}^{50} \frac{\#_i(\#_i-1)}{2}\right]\left[\sum_{i=1}^{50} \mathbf{1}' \times \tau(i) \times \mathbf{1}\right]^{-1}$	Low values of $PROXS_{SE1a}$ mean high dispersion. Distance calculations are independent of the distribution of the population throughout SE. This indicator reflects geographical proximity instead of population proximity.
$PROXS_{SE1b}$	***Inverse of the simple average of travel distances*** between singular entities within province i.	$PROXS_{SE1b}(i) = \frac{\#_i(\#_i - 1)}{2}[\mathbf{1}' \times \overline{\tau}(i) \times \mathbf{1}]^{-1} \quad [1b]$ Where $\mathbf{1}$ is a column vector $\#_i \times 1$ with 1 in all entries. $\boldsymbol{\overline{\tau}(i)}$ is a $\#_i \times \#_i$ triangular distance matrix whose entry td_{k1k2} is the travel distance between the singular entities k_1 and k_2 within province i for under-diagonal entries and zero for the diagonal and upper-diagonal entries. $\boldsymbol{td_{k_1,k_2}}$ is our estimated travel distance between the singular entities k_1 and k_2 within province i. We have calculated it using the ratio ρ_{1i} "travel distance to straight-line distance" for province i as follows: $td_{k_1,k_2} = \rho_{1i} d_i[k_1,k_2]$. We have estimated ρ_{1i} with a random sample of observation points[53] as described in point 5 of this paper. In detail: $PROXS_{SE1b}(i) = \left[\frac{\sum_{k_1,k_2=1, k_1<k_2}^{\#_i} \rho_{1i} d_i[k_1,k_2]}{\#_i(\#_i-1)/2}\right]^{-1}$	Regional $PROXS_{SE1b}(R_n) = \left[\sum_{i \in R_n} \frac{\#_i(\#_i-1)}{2}\right]\left[\sum_{i \in R_n} \mathbf{1}' \times \overline{\tau}(i) \times \mathbf{1}\right]^{-1}$ National $PROXS_{SE1b}(ES) = \left[\sum_{i=1}^{50} \frac{\#_i(\#_i-1)}{2}\right]\left[\sum_{i=1}^{50} \mathbf{1}' \times \overline{\tau}(i) \times \mathbf{1}\right]^{-1}$	Low values of $PROXS_{SE1b}$ mean high dispersion. Travel distance calculations are independent of the distribution of the population throughout SE. This indicator reflects geographical proximity instead of population proximity.

[53] Please remember that for modelling proximity indicators, we have considered "*observation points*". An "*observation point*" is a distance between two singular entities (alternatively two municipalities) **within the same province.** To calculate averages at the regional and national level we will consider the number of observation points within the same province, excluding additional couples of SE from different provinces and regions.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$PROXS_{SE1c}$	***Inverse of the simple average of travel durations*** between singular entities within province i.	$$PROXS_{SE1c}(i) = \frac{\#_i(\#_i-1)}{2}\left[\mathbf{1}' \times \bar{\bar{\tau}}(i) \times \mathbf{1}\right]^{-1} \quad [1c]$$ Where: $\mathbf{1}$ is a column vector $\#_i \times 1$ with 1 in all entries. $\bar{\bar{\boldsymbol{\tau}}}(\boldsymbol{i})$ is a $\#_i \times \#_i$ triangular distance matrix whose entry tdr_{k1k2} is the travel duration between the singular entities k_1 and k_2 within province i for under-diagonal entries and zero for the diagonal and upper-diagonal entries. $\boldsymbol{tdr_{k_1,k_2}}$ is our estimated travel duration between the singular entities k_1 and k_2 within province i. We have calculated it using the ratio ρ_{2i} "travel duration to straight-line distance" for province i as follows: $tdr_{k_1,k_2} = \rho_{2i} d_i[k_1,k_2]$. We have estimated ρ_{2i} with a random sample of observation points, as described in point 5 of this paper. In detail: $$PROXS_{SE1c}(i) = \left[\frac{\sum_{k_1,k_2=1, k_1<k_2}^{\#_i} \rho_{2i} d_i[k_1,k_2]}{\#_i(\#_i-1)/2}\right]^{-1}$$	Regional $$PROXS_{SE1c}(R_n) = \left[\sum_{i\in R_n} \frac{\#_i(\#_i-1)}{2}\right]\left[\sum_{i\in R_n} \mathbf{1}' \times \bar{\bar{\tau}}(i) \times \mathbf{1}\right]^{-1}$$ National $$PROXS_{SE1c}(ES) = \left[\sum_{i=1}^{50} \frac{\#_i(\#_i-1)}{2}\right]\left[\sum_{i=1}^{50} \mathbf{1}' \times \bar{\bar{\tau}}(i) \times \mathbf{1}\right]^{-1}$$	Low values of $PROXS_{SE1c}$ mean high dispersion. Travel duration calculations are independent of the distribution of the population throughout SE. This indicator reflects geographical proximity instead of population proximity.
$PROXW_{SE1d}$	***Inverse of the weighted average of straight-line distances*** between singular entities of province i.	$$PROXW_{SE1d}(i) = \left[\frac{\sigma'(i) \times \tau(i) \times \sigma(i)}{\sigma'(i) \times \tau(1) \times \sigma(i)}\right]^{-1} \quad [1d]$$ Where: ***σ (i)*** is a column vector $\#_i \times 1$ of populations of province i singular entities: $$\sigma(i)' = (\pi_{i1}, \pi_{i2}, \dots, \pi_{i\#_i})$$ **τ(1)** is a $\#_i \times \#_i$ triangular distance matrix whose entry d_{k1k2} is 1 for under-diagonal entries and zero for the diagonal and upper-diagonal entries. We weight the distance between two different singular entities of province i by the product of their respective populations. Therefore, the set of weights for the distances between SE k_1 and k_2 of province i is defined as follows (please, notice that this is a proper system of weights as their sum is 1): $$\frac{\pi_{ik_1}\pi_{ik_2}}{\sum_{k_1,k_2=1, k_1<k_2}^{\#_i} \pi_{ik_1}\pi_{ik_2}}$$ In detail: $$PROXW_{SE1d}(i) = \left[\frac{\sum_{k_1,k_2=1, k_1<k_2}^{\#_i} \pi_{ik_1}\pi_{ik_2} d_i[k_1,k_2]}{\sum_{k_1,k_2=1, k_1<k_2}^{\#_i} \pi_{ik_1}\pi_{ik_2}}\right]^{-1}$$	Regional $$PROXW_{SE1d}(R_n) = \left[\frac{\sum_{i\in R_n} \sigma'(i) \times \tau(i) \times \sigma(i)}{\sum_{i\in R_n} \sigma'(i) \times \tau(1) \times \sigma(i)}\right]^{-1}$$ National $$PROXW_{SE1d}(ES) = \left[\frac{\sum_{n=1}^{17}\sum_{i\in R_n} \sigma'(i) \times \tau(i) \times \sigma(i)}{\sum_{n=1}^{17}\sum_{i\in R_n} \sigma'(i) \times \tau(1) \times \sigma(i)}\right]^{-1}$$	More populated SE have higher influence on the average distance. $PROXW_{E1d}$ is higher than $PROXS_{SE1a}$ when more populated SE within the province tend to be close to each other. $PROXW_{SE1d}(i) = PROXS_{SE1a}(i)$ when all the SE are equidistant; also, when the population is equally distributed across SE. This indicator reflects population proximity.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
PROXW_{SE1e}	***Inverse of the weighted average of travel distances*** between singular entities of province i.	$\mathrm{PROXW}_{SE1e}(i)=\left[\frac{\sigma'(i)\times\bar{\tau}(i)\times\sigma(i)}{\sigma'(i)\times\tau(1)\times\sigma(i)}\right]^{-1}$ $[1e]$ We weight the travel distance between two different singular entities of province i by the product of their respective populations. In detail: $\mathrm{PROXW}_{SE1e}(i)=\left[\frac{\sum_{k_1,k_2=1,\ k_1<k_2}^{\#_i}\pi_{ik_1}\pi_{ik_2}\rho_{1i}d_i[k_1,k_2]}{\sum_{k_1,k_2=1,\ k_1<k_2}^{\#_i}\pi_{ik_1}\pi_{ik_2}}\right]^{-1}$	Regional $\mathrm{PROXW}_{SE1e}(R_n)=\left[\frac{\sum_{i\in R_n}\sigma'(i)\times\bar{\tau}(i)\times\sigma(i)}{\sum_{i\in R_n}\sigma'(i)\times\tau(1)\times\sigma(i)}\right]^{-1}$ National $\mathrm{PROXW}_{SE1e}(ES)=\left[\frac{\sum_{n=1}^{17}\sum_{i\in R_n}\sigma'(i)\times\bar{\tau}(i)\times\sigma(i)}{\sum_{n=1}^{17}\sum_{i\in R_n}\sigma'(i)\times\tau(1)\times\sigma(i)}\right]^{-1}$	More populated SE have higher influence on the average travel distance. PROXW_{SE1e} is higher than PROXS_{E1b} when more populated SE within the province tend to be close to each other regarding travel distance, $\mathrm{PROXW}_{SE1e}(i)=\mathrm{PROXS}_{SE1b}(i)$ when all the SE are equidistant; also, when the population is equally distributed across SE. This indicator reflects population proximity.
PROXW_{SE1f}	***Inverse of the weighted average of travel durations*** between singular entities of province i.	$\mathrm{PROXW}_{SE1f}(i)=\left[\frac{\sigma'(i)\times\bar{\bar{\tau}}(i)\times\sigma(i)}{\sigma'(i)\times\tau(1)\times\sigma(i)}\right]^{-1}$ $[1f]$ We weight the travel duration between two different singular entities of province i by the product of their respective populations. In detail: $\mathrm{PROXW}_{SE1f}(i)=\left[\frac{\sum_{k_1,k_2=1,\ k_1<k_2}^{\#_i}\pi_{ik_1}\pi_{ik_2}\rho_{2i}d_i[k_1,k_2]}{\sum_{k_1,k_2=1,\ k_1<k_2}^{\#_i}\pi_{ik_1}\pi_{ik_2}}\right]^{-1}$	Regional $\mathrm{PROXW}_{SE1f}(R_n)=\left[\frac{\sum_{i\in R_n}\sigma'(i)\times\bar{\bar{\tau}}(i)\times\sigma(i)}{\sum_{i\in R_n}\sigma'(i)\times\tau(1)\times\sigma(i)}\right]^{-1}$ National $\mathrm{PROXW}_{SE1f}(ES)=\left[\frac{\sum_{n=1}^{17}\sum_{i\in R_n}\sigma'(i)\times\bar{\bar{\tau}}(i)\times\sigma(i)}{\sum_{n=1}^{17}\sum_{i\in R_n}\sigma'(i)\times\tau(1)\times\sigma(i)}\right]^{-1}$	More populated SE have higher influence on the average travel distance. PROXW_{SE1f} is higher than PROXS_{E1c} when more populated SE within the province tend to be close to each other regarding travel duration. $\mathrm{PROXW}_{SE1f}(i)=\mathrm{PROXS}_{SE1c}(i)$ when all the SE are equidistant regarding travel duration; also, when the population is equally distributed across SE. This indicator reflects population proximity.
PROXR_{SE1g} PROXR_{SE1h} PROXR_{SE1i}	***Ratio of population proximity to geographical proximity***.	$\mathrm{PROXR}_{ES1g}(i)=\mathrm{PROXW}_{ES1_d}(i)/\mathrm{PROXS}_{ES1_a}(i)$ $[1g]$ $\mathrm{PROXR}_{ES1h}(i)=\mathrm{PROXW}_{ES1_e}(i)/\mathrm{PROXS}_{ES1_b}(i)$ $[1h]$ $\mathrm{PROXR}_{ES1i}(i)=\mathrm{PROXW}_{ES1_f}(i)/\mathrm{PROXS}_{ES1_c}(i)$ $[1i]$	Regional $\mathrm{PROXR}_{ES1z}(R_n)=\mathrm{PROXW}_{ES1_y}(R_n)/\mathrm{PROXS}_{ES1_x}(R_n)$ National $\mathrm{PROXR}_{ES1z}(ES)=\mathrm{PROXW}_{ES1_y}(ES)/\mathrm{PROXS}_{ES1_x}(ES)$ $(z, y, x)=\begin{cases}(g, d, a)\\(h, e, b)\\(i, f, c)\end{cases}$	Based on Galster criteria (Galster, G. et al (2001)).[54,55] While we have quantified PROXS_{ES1x} and PROXW_{ES1y} in absolute terms (level), PROXR_{ES1z} is a relative magnitude. It relates population proximity to geographical. High values of this indicator mean that populated SE are closer to each other than the entire set of SE locations: population proximity is higher than geographical proximity. This indicator approaches 0 as a minimum when the population concentrate in far from each other SE while SE locations as a whole are located on average close to each other. Thus, highly disperse population would be associated with low values of the indicator. On the contrary, if population were mostly concentrated in one SE the indicator could be extremely high. The maximum value is undefined.

[54] Please notice that the difference between this indicator and the one proposed by Galster lies in the weights used to calculate the weighted average distance. Galster's formulation uses $\frac{\pi_{ik_1}\pi_{ik_2}}{\pi_i^2}$ as loads for the distances, whose sum is not one. In our opinion, this option makes its interpretation and aggregation at the national and regional levels somewhat less intuitive.

[55] Galster, G. et al. (2001). Wrestling sprawl to the ground: Defining and measuring an elusive concept. Housing Policy Debate, 12, 681–717.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$\text{PROXN}_{\text{SE1j}}$	***Normalised geographical proximity (SE & straight-line distance)***	$\text{PROXN}_{\text{SE1}j}(i) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}a}(i) \times D_{adj}(i)}$ $[1j]$	Regional $\text{PROXN}_{\text{SE1}j}(iR_n) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}a}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$ National $\text{PROXN}_{\text{SE1}j}(ES) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}a}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$	The simple average of straight-line distances between SE of province i ($\bar{d}(i)$) is rescaled and expressed as units of the diagonal of the province, which is set as the standard. Then we calculate: $1 - \frac{\bar{d}(i)}{D_{adj}(i)}$ Low values of the indicator (low proximity) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that the population is located in one single centre (this centre does not necessarily match the geometric centre). If the indicator is close to zero, SE are, on average, as distant to each other as the diagonal of the province.
$\text{PROXN}_{\text{SE1k}}$	***Normalised geographical proximity (SE & travel distance)***	$\text{PROXN}_{\text{SE1}k}(i) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}b}(i) \times D_{adj}(i)}$ $[1k]$	Regional $\text{PROXN}_{\text{SE1}k}(R_n) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}b}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$ National $\text{PROXN}_{\text{SE1}k}(ES) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}b}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$	The simple average of travel distances between SE of province i ($\bar{\bar{d}}(i)$) is rescaled and expressed as units of the diagonal of the province, which is set as the standard. Then we calculate: $1 - \frac{\bar{\bar{d}}(i)}{D_{adj}(i)}$ Low values of the indicator (low proximity) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in one single centre (this centre does not necessarily match the geometric centre). If the indicator is close to zero, SE are, on average, as distant to each other as the diagonal of the province.
$\text{PROXN}_{\text{SE1l}}$	***Normalised population proximity (SE & straight-line distance)***	$\text{PROXN}_{\text{SE1}l}(i) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}d}(i) \times D_{adj}(i)}$ $[1l]$	Regional $\text{PROXN}_{\text{SE1}l}(R_n) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}d}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$ National $\text{PROXN}_{\text{SE1}l}(ES) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}d}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$	The weighted average of straight-line distances between SE of province i ($\tilde{d}$) is rescaled and expressed as units of the diagonal of the province, which is set as the standard. Then we calculate: $1 - \frac{\tilde{d}(i)}{D_{adj}(i)}$ Low values of the indicator (low proximity) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in one single centre (this centre does not necessarily match the geometric centre). If the indicator is close to zero, people are, on average, as distant as the diagonal of the province.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$\text{PROXN}_{\text{SE1m}}$	***Normalised population proximity (SE & travel distance)***	$\text{PROXN}_{\text{SE1}m}(i) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}e}(i) \times D_{adj}(i)}$ $[1m]$	Regional $\text{PROXN}_{\text{SE1}m}(R_n) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}e}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$ National $\text{PROXN}_{\text{SE1}m}(ES) = 1 - \frac{1}{\text{PROXS}_{\text{SE1}e}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$	The weighted average of travel distances between SE of province i ($\tilde{\tilde{d}}(i)$) is rescaled and expressed as units of the diagonal of the province, which is set as the standard. $1 - \frac{\tilde{\tilde{d}}(i)}{D_{adj}(i)}$ Low values of the indicator (low proximity) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in one single centre (this centre does not necessarily match the geometric centre). If the indicator is close to zero, people are, on average, as distant as the diagonal of the province.
$\text{PROXS}_{\text{MUN2a}}$	***Inverse of the simple average of straight-line distances*** between municipalities within province i.	$\text{PROXS}_{\text{MUN2}a}(i) = \frac{\mu_i(\mu_i - 1)}{2} [\mathbf{1}' \times T(i) \times \mathbf{1}]^{-1}$ $[2a]$ Where: $\mathbf{1}$ is a column vector $\mu_i \times 1$ with 1 in all entries. $\boldsymbol{T(i)}$ is a triangular distance matrix whose entry d_{j1j2} is the straight-line distance between the municipalities j_1 and j_2 within province i for under-diagonal entries and zero for the diagonal and upper-diagonal entries.	Regional $\text{PROXS}_{\text{MUN2}a}(R_n) = \left[\sum_{i \in R_n} \frac{\mu_i(\mu_i - 1)}{2}\right]\left[\sum_{i \in R_n} \mathbf{1}' \times T(i) \times \mathbf{1}\right]^{-1}$ National $\text{PROXS}_{\text{MUN2}a}(\text{ES}) = \left[\sum_{i=1}^{50} \frac{\mu_i(\mu_i - 1)}{2}\right]\left[\sum_{i=1}^{50} \mathbf{1}' \times T(i) \times \mathbf{1}\right]^{-1}$	Same as $\text{PROXS}_{\text{ES1a}}$ referred to municipalities.
$\text{PROXS}_{\text{MUN2b}}$	***Inverse of the simple average of travel distances*** between municipalities within province i.	$\text{PROXS}_{\text{MUN2}b}(i) = \frac{\mu_i(\mu_i - 1)}{2} \left[\mathbf{1}' \times \overline{T}(i) \times \mathbf{1}\right]^{-1}$ $[2b]$ Where $\mathbf{1}$ is a column vector $\mu_i \times 1$ with 1 in all entries. $\boldsymbol{\overline{T}(i)}$ is a $\mu_i \times \mu_i$ triangular distance matrix whose entry td_{j1j2} is the travel distance between the municipalities j_1 and j_2 within province i for under-diagonal entries and zero for the diagonal and upper-diagonal entries of province i. $\boldsymbol{td_{j_1,j_2}}$ is our estimated travel distance between the municipalities j_1 and j_2 within province i. We calculated it using the ratio ρ_{1i} "travel distance to straight-line distance" for province i as follows: $tdr_{j_1,j_2} = \rho_{1i} d_i[j_1, j_2]$. We have estimated ρ_{1i} with a random sample of observation points, as described in point 5 of this paper.	Regional $\text{PROXS}_{\text{MUN2}b}(R_n) = \left[\sum_{i \in R_n} \frac{\mu_i(\mu_i - 1)}{2}\right]\left[\sum_{i \in R_n} \mathbf{1}' \times \overline{T}(i) \times \mathbf{1}\right]^{-1}$ National $\text{PROXS}_{\text{MUN2}b}(\text{ES}) = \left[\sum_{i=1}^{50} \frac{\mu_i(\mu_i - 1)}{2}\right]\left[\sum_{i=1}^{50} \mathbf{1}' \times \overline{T}(i) \times \mathbf{1}\right]^{-1}$	Same as $\text{PROXS}_{\text{ES1b}}$ referred to municipalities.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$PROXS_{MUN2c}$	***Inverse of the simple average of travel durations*** between municipalities within province i	$$PROXS_{MUN1c}(i) = \frac{\mu_i(\mu_i - 1)}{2}\left[\mathbf{1}' \times \overline{\overline{T}}(i) \times \mathbf{1}\right]^{-1} \quad [2c]$$ Where $\mathbf{1}$ is a column vector $\mu_i \times 1$ with 1 in all entries. $\overline{\overline{\boldsymbol{T}}}(\boldsymbol{i})$ is a $\mu_i \times \mu_i$ triangular distance matrix whose entry tdr_{j1j2} is the travel duration between the municipalities j_1 and j_2 within province i for under-diagonal entries and zero for the diagonal and upper-diagonal entries of province i. $\boldsymbol{tdr_{j_1 j_2}}$ is our estimated travel duration between the municipalities j_1 and j_2 within province i. We calculated it using the ratio ρ_{2i} "travel distance to straight-line distance" for province i as follows: $tdr_{j_1,j_2} = \rho_{2i} d_i[j_1, j_2]$. We have estimated ρ_{2i} with a random sample of observation points, as described in point 5 of this paper.	Regional $$PROXS_{MUN2c}(R_n) = \left[\sum_{i \in R_n} \frac{\mu_i(\mu_i - 1)}{2}\right]\left[\sum_{i \in R_n} \mathbf{1}' \times \overline{\overline{T}}(i) \times \mathbf{1}\right]^{-1}$$ National $$PROXS_{MUN2c}(ES) = \left[\sum_{i=1}^{50} \frac{\mu_i(\mu_i - 1)}{2}\right]\left[\sum_{i=1}^{50} \mathbf{1}' \times \overline{\overline{T}}(i) \times \mathbf{1}\right]^{-1}$$	Same as $PROXS_{ES1c}$ referred to municipalities.
$PROXW_{MUN2d}$	***Inverse of the weighted average of straight-line distances*** between municipalities within province i.	$$PROXW_{MUN2d}(i) = \left[\frac{s'(i) \times T(i) \times s(i)}{s'(i) \times T(1) \times s(i)}\right]^{-1} \quad [2d]$$ Where $\boldsymbol{s(i)}$ is a column vector $\mu_i \times 1$ of municipalities populations: $$s(i)' = (\pi_{i1}, \pi_{i2}, \ldots, \pi_{i\mu_i})$$ $\boldsymbol{T(1)}$ is a $\mu_i \times \mu_i$ triangular distance matrix whose entry d_{j1j2} is 1 for under-diagonal entries and zero for the diagonal and upper-diagonal entries. We weight the distance between two different municipalities of province i by the product of their respective populations. We notice that for indicator $PROXW_{MUN2d}$, weights for the distance between municipalities j_1 and j_2 ($j_1 < j_2$) of province i are: $$\frac{\pi_{ij_1}\pi_{ij_2}}{s'(i) \times T(1) \times s(i)}$$ Which are a proper set of weights as their sum is 1.	Regional $$PROXW_{MUN2d}(R_n) = \left[\frac{\sum_{i \in R_n} s'(i) \times T(i) \times s(i)}{\sum_{i \in R_n} s'(i) \times T(1) \times s(i)}\right]^{-1}$$ National $$PROXW_{MUN2d}(ES) = \left[\frac{\sum_{n=1}^{17}\sum_{i \in R_n} s'(i) \times T(i) \times s(i)}{\sum_{n=1}^{17}\sum_{i \in R_n} s'(i) \times T(1) \times s(i)}\right]^{-1}$$	Same as $PROXW_{ES1d}$ referred to municipalities.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$PROXW_{MUN2e}$	***Inverse of the weighted average of travel distances*** between municipalities within province i.	$$PROXW_{MUN2e}(i) = \left[\frac{s'(i) \times \overline{T}(i) \times s(i)}{s'(i) \times T(1) \times s(i)}\right]^{-1} \quad [2e]$$ We weight the distance between two different municipalities of province i by the product of their respective populations.	Regional $$PROXW_{MUN2e}(R_n) = \left[\frac{\sum_{i \in R_n} s'(i) \times \overline{T}(i) \times s(i)}{\sum_{i \in R_n} s'(i) \times T(1) \times s(i)}\right]^{-1}$$ National $$PROXW_{MUN2e}(ES) = \left[\frac{\sum_{n=1}^{17} \sum_{i \in R_n} s'(i) \times \overline{T}(i) \times s(i)}{\sum_{n=1}^{17} \sum_{i \in R_n} s'(i) \times T(1) \times s(i)}\right]^{-1}$$	Same as $PROXW_{ES1e}$ referred to municipalities.
$PROXW_{MUN2f}$	***Inverse of the weighted average of travel durations*** between municipalities within province i.	$$PROXW_{MUN2f}(i) = \left[\frac{s'(i) \times \overline{\overline{T}}(i) \times s(i)}{s'(i) \times T(1) \times s(i)}\right]^{-1} \quad [2f]$$ We weight the distance between two different municipalities of province i by the product of their respective populations.	Regional $$PROXW_{MUN2f}(R_n) = \left[\frac{\sum_{i \in R_n} s'(i) \times \overline{\overline{T}}(i) \times s(i)}{\sum_{i \in R_n} s'(i) \times T(1) \times s(i)}\right]^{-1}$$ National $$PROXW_{MUN2f}(ES) = \left[\frac{\sum_{n=1}^{17} \sum_{i \in R_n} s'(i) \times \overline{\overline{T}}(i) \times s(i)}{\sum_{n=1}^{17} \sum_{i \in R_n} s'(i) \times T(1) \times s(i)}\right]^{-1}$$	Same as $PROXW_{ES1f}$ referred to municipalities.
$PROXR_{MUN2g}$ $PROXR_{MUN2h}$ $PROXR_{MUN2i}$	***Ratio population proximity to geographical proximity***.	$$PROXR_{MUN2g}(i) = PROXW_{MUN2d}(i)/PROXS_{MUN2a}(i) \quad [2g]$$ $$PROXR_{MUN2h}(i) = PROXW_{MUN2e}(i)/PROXS_{MUN2b}(i) \quad [2h]$$ $$PROXR_{MUN2i}(i) = PROXW_{MUN2f}(i)/PROXS_{MUN2c}(i) \quad [2i]$$	Regional $$PROXR_{MUN2z}(R_n) = PROXW_{MUN2y}(R_n)/PROXS_{MUN2x}(R_n)$$ National $$PROXR_{MUN2z}(ES) = PROXW_{MUN2y}(ES)/PROXS_{MUN2x}(ES)$$ (z, y, x) = (g, d, a); or (h, e, b); or (i, f, c)	Same as $PROXS_{ES1g}$, $PROXS_{ES1h}$ and $PROXS_{ES1i}$ referred to municipalities.
$PROXN_{MUN2j}$	***Normalised geographical proximity (MUN & straight-line distance)***	$$PROXN_{MUN2j}(i) = 1 - \frac{1}{PROXS_{MUN2a}(i) \times D_{adj}(i)} \quad [2j]$$	Regional $$PROXN_{MUN21j}(R_n) = 1 - \frac{1}{PROXS_{MUN2a}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$$ National $$PROXN_{MUN2j}(ES) = 1 - \frac{1}{PROXS_{MUN2a}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$$	The simple average of straight-line distances between municipalities of province i ($\overline{dd}(i)$) is rescaled and expressed as units of the diagonal of the province, which is set as the measurement standard. $$1 - \frac{\overline{dd}(i)}{D_{adj}(i)}$$ Low values of the indicator (low proximity) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in one single centre (this centre does not necessarily match the geometric centre). If the indicator is close to zero, municipalities are, on average, as distant to each other as the diagonal of the province.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
PROXN_{MUN2k}	***Normalised geographical proximity (MUN & travel distance)***	$\text{PROXN}_{MUN2k}(i) = 1 - \frac{1}{\text{PROXS}_{MUN2b}(i) \times D_{adj}(i)}$ [2k]	Regional $\text{PROXN}_{MUN2k}(R_n) = 1 - \frac{1}{\text{PROXS}_{MUN2b}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$ National $\text{PROXN}_{MUN2k}(ES) = 1 - \frac{1}{\text{PROXS}_{MUN2b}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$	The simple average of travel distances between municipalities of province i ($\overline{\overline{dd}}(i)$) is rescaled and expressed as units of the diagonal of the province, which is set as the measurement standard. $1 - \frac{\overline{\overline{dd}}(i)}{D_{adj}(i)}$ Low values of the indicator (low proximity) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in one single centre (this centre does not necessarily match the geometric centre). If the indicator is close to zero, municipalities are, on average, as distant to each other as the diagonal of the province.
PROXN_{MUN2l}	***Normalised population proximity (MUN & straight-line distance)***	$\text{PROXN}_{MUN2l}(i) = 1 - \frac{1}{\text{PROXS}_{MUN2d}(i) \times D_{adj}(i)}$ [2l]	Regional $\text{PROXN}_{MUN2l}(R_n) = 1 - \frac{1}{\text{PROXS}_{MUN2d}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$ National $\text{PROXN}_{MUN2l}(ES) = 1 - \frac{1}{\text{PROXS}_{MUN2d}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$	The weighted average of straight-line distances between municipalities of province i ($\widetilde{dd}(i)$) is rescaled and expressed as units of the diagonal of the province, which is set as the measurement standard. $1 - \frac{\widetilde{dd}(i)}{D_{adj}(i)}$ Low values of the indicator (low proximity) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in one single centre (this centre does not necessarily match the geometric centre). If the indicator is close to zero, people are, on average, as distant as the diagonal of the province.
PROXN_{MUN2m}	***Normalised population proximity (MUN & travel distance)***	$\text{PROXN}_{MUN2m}(i) = 1 - \frac{1}{\text{PROXS}_{MUN2e}(i) \times D_{adj}(i)}$ [2m]	Regional $\text{PROXN}_{MUN2m}(R_n) = 1 - \frac{1}{\text{PROXS}_{MUN2e}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$ National $\text{PROXN}_{MUN2m}(ES) = 1 - \frac{1}{\text{PROXS}_{MUN2e}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$	The weighted average of travel distance between municipalities of province i ($\widetilde{\widetilde{dd}}(i)$) is rescaled and expressed as units of the diagonal of the province, which is set as the measurement standard. $1 - \frac{\widetilde{\widetilde{dd}}(i)}{D_{adj}(i)}$ Low values of the indicator (low proximity) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in one single centre (this centre does not necessarily match the geometric centre). If the indicator is close to zero, people are, on average, as distant as the diagonal of the province.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$PROXV_{MUN2n}$ $PROXV_{MUN2o}$ $PROXV_{MUN2p}$	***Standardised Proximity Index*** for province i based on Venables Index.	$$PROXV_{MUN2n}(i) = 1 - \frac{V(i)}{V_{max}(i)} \qquad [2n]$$ Where **V** is the Venables index, calculated as follows: $\boldsymbol{V(i)} = \tilde{s}\,'(i) \times T(i) \times \tilde{s}\,(i)$ $\boldsymbol{V_{max}(i)}$ is the maximum attainable value of the Venables index. $\boldsymbol{\tilde{s}(i)}$ is a column vector of population weights of the municipalities: $$\tilde{s}'(i) = \left(\frac{\pi_{i1}}{\pi_i}, \quad \frac{\pi_{i2}}{\pi_i}, \dots, \frac{\pi_{i\mu_i}}{\pi_i}\right)$$ The Venables index is a sort of weighted average of straight-line distances between any two municipalities within the same province where each distance is given a different load depending on the population of the municipalities, specifically on the product of the two population weights. Please notice that the system of products of the population weights of any two municipalities is not a proper weight system as they don't sum 1. Nonetheless, we kept this formulation following Pereira, R.H.M. et al. (2013) and Pereira, R.H.M. et al. (2015), our original source, because our final use of the Venables index is to build a standardised proximity indicator whose formulations using both the Venables index and the one based on the proper system of population weights are equivalent. The estimation of V_{max} is not trivial, because it has no closed-form solution. In a region forming a perfect circle, the maximum value of V occurs when all the population is evenly distributed along the external edge (Pereira et al. 2013 and 2015), where evenly means equidistant and with the same population (equal population weights). In our work, for each province I, we start with: $$VI_{max}(i) = \tilde{s}\,'(i) \times T_{max}(i) \times \tilde{s}\,(i)$$ Where $T_{max}(i)$ is a triangular distance matrix whose entry d_{ij1j2}, j1 < j2, is the distance between municipalities j1 and j2 of province i after having distributed them evenly along the external edge of the axes-	Regional $$\mathrm{PROXN}_{\mathrm{MUN}2n}(R_n) = 1 - \frac{\sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} \times \mathrm{V}(i)}{\sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} \times \mathrm{V}_{max}(i)}$$ National $$\mathrm{PROXN}_{\mathrm{MUN}2n}(ES) = 1 - \frac{\sum_{i=1}^{50} \frac{\pi_i}{\pi} \times \mathrm{V}(i)}{\sum_{i=1}^{50} \frac{\pi_i}{\pi} \times \mathrm{V}_{max}(i)}$$ ***$PROXV_{MUN2n}$*** is formulated for straight-line distance using the matrix T(i) ***$PROXV_{MUN2o}$*** is formulated for travel distance using the matrix $\overline{\boldsymbol{T}}(\boldsymbol{i})$ ***$PROXV_{MUN2p}$*** is formulated for travel duration using the matrix $\overline{\overline{\boldsymbol{T}}}(\boldsymbol{i})$	Based on the spatial separation index originally proposed by Midelfart-Knarvik et al. (Pereira et al. 2013 and 2015) to evaluate changes in the spatial distribution of economic activity across European regions. When all population is located in just one spatial unit, V reaches its minimum value that is zero. However, the index has no maximum value and we lack benchmarks for comparisons. A way to overcome this limitation is to calculate the maximum attainable value of V. This normalisation procedure makes comparisons of provinces of different shapes and sizes possible. The interpretation of $PROXV_{MUN2n}$ is the opposite of V. Its theoretical range is (0, 1). Values of $PROXV_{MUN2n}$ close to 1 mean that population is located in one single centre (this centre does not necessarily match the geometric centre). If $PROXV_{MUN2n}$ is zero, population is as spatially separated as possible. In other words, population is distributed in a way that maximises the distances between them.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
		aligned 2-dimensional bounding box of province i; starting in the upper left corner. We notice that placing the population on the edge of the axes-aligned bounding box is further away from placing it on the edge of the province and the gap depends on the surface area of the province and the degree of fit of that surface to the rectangle. Therefore, to improve comparability, we have controlled VI_{max} by this gap measured through the percentage surface of the axes-aligned bounding box covered by the surface area of the province. To this end, we have used the following equation that captures the association between VI_{max} and its relevant drivers, including the mentioned control factor: $$y = m_1 x_1 + m_2 x_2 + m_3 x_3 + b$$ Where, for each province: $y = VI_{max}$ (Km) x_1 = Square root of the surface area of the province (Km). x_2 = Surface coverage within the axes-aligned bounding box (in parts per unit). x_3 = Ratio simple average of straight-lines distance to V (dimensionless). The rationale behind is that the VI_{max} is driven by the axes-aligned bounding box extension (captured by x_1), the shape of the province (capture by x_2), and the population loads given to each distance between any two municipalities within the province (capture by x_3). Based on this association, we define our final V_{max} as follows: $$V_{max} = \hat{m}_1 x_1 + \hat{m}_2 x_2 + \hat{m}_3 x_3 + \hat{b}$$ Where we have determine $\hat{m}_1,\ \hat{m}_2,\ \hat{m}_3\ and\ \hat{b}$ based on the least squares criterion for each year of the analysed period. As for peninsular provinces, we exclude the islands to determine $\hat{m}_1, \hat{m}_2, \hat{m}_3, \hat{b}$. As for the islands, we include all provinces to determine these parameters so that we control by the gap between the surface area of each province and that of the axes-aligned bounding box, but we retain the effect of inter-island distances.		

Centrality indicators

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$CBDdS_{SE3a}$	***Inverse of the simple average of straight-line distances from SE to CBD.***	$$\mathrm{CBDdS_{SE3a}}(i) = (\#_i - 1)[\mathbf{1}' \times \tau_{cbd}(i)]^{-1} \quad [3a]$$ Where: **1** is a column vector (#i -1) × 1 with 1 in all entries. $\boldsymbol{\tau_{cbd}(i)}$ is a column vector (#i -1) × 1 whose entry $d_{cbd,k}$ is the straight-line distance from SE k to the CBD within province i. It excludes the CBD.	Regional $$\mathrm{CBDdS_{SE3a}}(R_n) = \left[\sum_{i \in R_n}(\#_i - 1)\right]\left[\sum_{i \in R_n}\mathbf{1}' \times \tau_{cbd}(i)\right]^{-1}$$ National $$\mathrm{CBDdS_{SE3a}}(\mathrm{ES}) = \left[\sum_{i=1}^{50}(\#_i - 1)\right]\left[\sum_{i=1}^{50}\mathbf{1}' \times \tau_{cbd}(i)\right]^{-1}$$	It captures centrality through the geographical proximity of the SE to the CDB. Its formulation is based on the inverse of the average of straight-line distances from SE k to the CBD within province i. Thus, the lowest the centrality the highest the dispersion.
$CBDdS_{SE3b}$	***Inverse of the simple average of the travel distances from SE to CBD.***	$$\mathrm{CBDdS_{SE3b}}(i) = (\#_i - 1)[\mathbf{1}' \times \overline{\tau}_{cbd}(i)]^{-1} \quad [3b]$$ Where: $\boldsymbol{\overline{\tau}_{cbd}(i)}$ is a column vector (#i -1) × 1 whose entry $td_{cbd,k}$ is the travel distance from SE k to the CBD within province i. It excludes the CBD. $\boldsymbol{td_{cbd,k}}$ is our estimated travel distance between the CBD and SE k within province i. We calculated it using the ratio ρ_{1i} "travel distance to straight-line distance" for province i as follows: $td_{cbd,k} = \rho_{1i}d_i[CBD,k]$. We have estimated ρ_{1i} with a random sample of observation points, as described in point 5 of this paper.	Regional $$\mathrm{CBDdS_{SE3b}}(R_n) = \left[\sum_{i \in R_n}(\#_i - 1)\right]\left[\sum_{i \in R_n}\mathbf{1}' \times \overline{\tau}_{cbd}(i)\right]^{-1}$$ National $$\mathrm{CBDdS_{SE3b}}(\mathrm{ES}) = \left[\sum_{i=1}^{50}(\#_i - 1)\right]\left[\sum_{i=1}^{50}\mathbf{1}' \times \overline{\tau}_{cbd}(i)\right]^{-1}$$	It captures centrality through the geographical proximity of the SE to the CDB. Its formulation is based on the inverse of the average of travel distances from SE k to the CBD within province i. Thus, the lowest the centrality the highest the dispersion.
$CBDdS_{SE3c}$	***Inverse of the simple average of the travel durations from SE to CBD.***	$$\mathrm{CBDdS_{SE3c}}(i) = (\#_i - 1)\left[\mathbf{1}' \times \overline{\overline{\tau}}_{cbd}(i)\right]^{-1} \quad [3c]$$ Where: $\boldsymbol{\overline{\overline{\tau}}_{cbd}(i)}$ is a column vector (#i -1) × 1 whose entry $tdr_{cbd,k}$ is the travel duration from SE k to the CBD within province i. It excludes the CBD. $\boldsymbol{tdr_{cbd,k}}$ is our estimated travel duration between the CBD and SE k within province i. We calculated it using the ratio ρ_{2i} "travel duration to straight-line distance" for province i as follows: $tdr_{cbd,k} = \rho_{2i}d_i[CBD,k]$. We have estimated ρ_{2i} with a random sample of observation points, as described in point 5 of this paper.	Regional $$\mathrm{CBDdS_{SE3c}}(R_n) = \left[\sum_{i \in R_n}(\#_i - 1)\right]\left[\sum_{i \in R_n}\mathbf{1}' \times \overline{\overline{\tau}}_{cbd}(i)\right]^{-1}$$ National $$\mathrm{CBDdS_{SE3c}}(\mathrm{ES}) = \left[\sum_{i=1}^{50}(\#_i - 1)\right]\left[\sum_{i=1}^{50}\mathbf{1}' \times \overline{\overline{\tau}}_{cbd}(i)\right]^{-1}$$	It captures centrality through the geographical proximity of the SE to the CDB. Its formulation is based on the inverse of the average of travel durations from SE k to the CBD within province i. Thus, the lowest the centrality the highest the dispersion.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$\text{CBDdW}_{\text{SE3d}}$	***Inverse of the weighted average of the straight-line distances from SE to CBD.***	$$\text{CBDdW}_{\text{SE3d}}(i) = \left[\frac{\sigma(i)' \times \tau_{cbd}(i)}{\sigma(i)' \times \mathbf{1}}\right]^{-1} \quad [3d]$$ Where: $\boldsymbol{\sigma(i)}$ is a column vector $\#_i \times 1$ of populations of province i singular entities: $$\sigma(i)' = (\pi_{i1}, \pi_{i2}, \dots, \pi_{i\#_i})$$ It includes the CBD. We weight the distance between each singular entity of province i and the CBD by the population of the SE. Therefore, the set of weights for the distances between SE k_1 and the CBD of province i is defined as follows: $$\frac{\pi_{ik_1}}{\sum_{k=1}^{\#_i} \pi_{ik}}$$ In detail: $$\text{PROXW}_{\text{SE1}d}(i) = \left[\frac{\sum_k^{\#_i} \pi_{ik} d_i[k, CBD]}{\pi_I}\right]^{-1}$$	Regional $$\text{CBDdW}_{\text{SE3d}}(R_n) = \left[\frac{\sum_{i \in R_n} \sigma(i)' \times \tau_{cbd}(i)}{\sum_{i \in R_n} \sigma(i)' \times \mathbf{1}}\right]^{-1}$$ National $$\text{CBDdW}_{\text{SE3d}}(\text{ES}) = \left[\frac{\sum_{i=1}^{50} \sigma(i)' \times \tau_{cbd}(i)}{\sum_{i=1}^{50} \sigma(i)' \times \mathbf{1}}\right]^{-1}$$	It captures centrality through the population proximity to the CDB. Its formulation is based on the inverse of the weighted average of straight-line distances from SE k to the CBD within province i. Thus, that the lowest the centrality the highest the dispersion.
$\text{CBDdW}_{\text{SE3e}}$	***Inverse of the weighted average of the travel distances from SE to CBD.***	$$\text{CBDdW}_{\text{SE3e}}(i) = \left[\frac{\sigma(i)' \times \overline{\tau}_{cbd}(i)}{\sigma(i)' \times \mathbf{1}}\right]^{-1} \quad [3e]$$ We weight the distance between each singular entity of province i and the CBD by the population of the SE. Therefore, the set of weights for the distances between SE k_1 and the CBD of province i is defined as follows: $$\frac{\pi_{ik_1}}{\sum_{k=1}^{\#_i} \pi_{ik}}$$ In detail: $$\text{PROXW}_{\text{SE1}d}(i) = \left[\frac{\sum_k^{\#_i} \pi_{ik} \rho_{1i} d_i[k, CBD]}{\pi_i}\right]^{-1}$$	Regional $$\text{CBDdW}_{\text{SE3e}}(R_n) = \left[\frac{\sum_{i \in R_n} \sigma(i)' \times \overline{\tau}_{cbd}(i)}{\sum_{i \in R_n} \sigma(i)' \times \mathbf{1}}\right]^{-1}$$ National $$\text{CBDdW}_{\text{SE3e}}(\text{ES}) = \left[\frac{\sum_{i=1}^{50} \mathbf{1}' \times \overline{\tau}_{cbd}(i)}{\sum_{i=1}^{50} \mathbf{1}' \times \sigma(i)}\right]^{-1}$$	It captures centrality through the population proximity to CDB. Its formulation is based on the inverse of the weighted average of travel distances from SE k to the CBD within province i. Thus, the lowest the centrality the highest the dispersion.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$\text{CBDdW}_{\text{SE3f}}$	***Inverse of the weighted average of the travel durations from SE to CBD.***	$\text{CBDdW}_{\text{SE3f}}(i) = \left[\frac{\sigma(i)' \times \bar{\bar{\tau}}_{cbd}(i)}{\sigma(i)' \times \mathbf{1}}\right]^{-1}$ $[3f]$ We weight the distance between each singular entity of province i and the CBD by the population of the SE. Therefore, the set of weights for the distances between SE k_1 and the CBD of province i is defined as follows: $\frac{\pi_{ik_1}}{\sum_{k=1}^{\#_i} \pi_{ik}}$ In detail: $\text{PROXW}_{\text{SE1}d}(i) = \left[\frac{\sum_{k}^{\#_i} \pi_{ik}\rho_{2i}d_i[k, CBD]}{\pi_i}\right]^{-1}$	Regional $\text{CBDdW}_{\text{SE3f}}(R_n) = \left[\frac{\sum_{i \in R_n} \sigma(i)' \times \bar{\bar{\tau}}_{cbd}(i)}{\sum_{i \in R_n} \sigma(i)' \times \mathbf{1}}\right]^{-1}$ National $\text{CBDdW}_{\text{SE3f}}(\text{ES}) = \left[\frac{\sum_{i=1}^{50} \sigma(i)' \times \bar{\bar{\tau}}_{cbd}(i)}{\sum_{i=1}^{50} \sigma(i)' \times \mathbf{1}}\right]^{-1}$	It captures centrality through the population proximity to CDB. Its formulation is based on the inverse of the weighted average of travel durations from SE k to the CBD within province i. Thus, the lowest the centrality the highest the dispersion.
$\text{CBDdR}_{\text{SE3g}}$ $\text{CBDdR}_{\text{SE3h}}$ $\text{CBDdR}_{\text{SE3i}}$	***Ratio population centrality to geographical centrality.***[56]	$\text{CBDdR}_{\text{SE3g}}(i) = \text{CBDdW}_{\text{ES}_{3d}}(i)/\text{CBDdS}_{\text{ES}_{3a}}(i)$ $[3g]$ $\text{CBDdR}_{\text{SE3h}}(i) = \text{CBDdW}_{\text{ES}_{3e}}(i)/\text{CBDdS}_{\text{ES}_{3b}}(i)$ $[3h]$ $\text{CBDdR}_{\text{SE3i}}(i) = \text{CBDdW}_{\text{ES}_{3f}}(i)/\text{CBDdS}_{\text{ES}_{3c}}(i)$ $[3i]$	Regional $\text{CBDdR}_{\text{ES3z}}(R_n) = \text{CBDdR}_{\text{ES}_{3y}}(R_n)/\text{CBDdR}_{\text{ES}_{3x}}(R_n)$ National $\text{CBDdR}_{\text{ES3z}}(ES) = \text{CBDdR}_{\text{ES}_{3y}}(ES)/\text{CBDdR}_{\text{ES}_{3x}}(ES)$ (z, y, x) = $\begin{cases} (\text{g, d, a}) \\ (\text{h, e, b}) \\ (\text{i, f, c}) \end{cases}$	While we have quantified $\text{CBDdS}_{\text{ES3x}}$ and $\text{CBDdW}_{\text{ES3y}}$ in absolute terms (level), $\text{CBDdR}_{\text{ES3z}}$ is a relative magnitude. It relates population centrality to geographical centrality. High values of this indicator signify that populated SE are closer to CBD than the entire set of SE locations: population centrality is higher than locations centrality. This indicator approaches 0 as a minimum when populated SE tend to be farther from the centre. On the other hand, it increases as the whole population tends to be concentrated in the CBD to a greater extent than locations themselves. The maximum value is undefined. Lower values reflect greater dispersion.
$\text{CBDdN}_{\text{SE3j}}$	***Normalised geographical centrality (SE & straight-line distance)***	$\text{CBDdN}_{\text{SE3j}}(i) = 1 - \frac{1}{\text{CBDdS}_{\text{SE3a}}(i) \times D_{adj}(i)}$ $[3j]$	Regional $\text{CBDdN}_{\text{SE3j}}(R_n) = 1 - \frac{1}{\text{CBDdS}_{\text{SE3a}}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$ National $\text{CBDdN}_{\text{SE3j}}(ES) = 1 - \frac{1}{\text{CBDdS}_{\text{SE3a}}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$	It captures centrality through the geographical proximity of the SE to CDB. Its formulation is based on the simple average of straight-line distances from SE k to the CBD within province i, rescaled and expressed as units of the diagonal of the province, which is set as the standard. Low values of the indicator (low centrality) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in the CBD. If the indicator is close to zero, SE are, on average, as distant to the CDB as the diagonal of the province.

[56] The ratio of weighted to unweighted average distance. These equations has been adapted from Lee, S. (2015).

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$CBDdN_{SE3k}$	***Normalised geographical centrality (SE & travel distance)***	$$\text{CBDdN}_{\text{SE3k}}(i) = 1 - \frac{1}{\text{CBDdS}_{\text{SE3b}}(i) \times D_{adj}(i)} \quad [3k]$$	Regional $$\text{CBDdN}_{\text{SE3k}}(R_n) = 1 - \frac{1}{\text{CBDdS}_{\text{SE3b}}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$$ National $$\text{CBDdN}_{\text{SE3k}}(ES) = 1 - \frac{1}{\text{CBDdS}_{\text{SE3b}}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$$	It captures centrality through the geographical proximity of the SE to CDB. Its formulation is based on the simple average of travel distances from SE k to the CBD within province i, rescaled and expressed as units of the diagonal of the province, which is set as the standard. Low values of the indicator (low centrality) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in the CBD. If the indicator is close to zero, SE are, on average, as distant to the CDB as the diagonal of the province.
$CBDdN_{SE3l}$	***Normalised population centrality (SE & straight-line distance)***	$$\text{CBDdN}_{\text{SE3l}}(i) = 1 - \frac{1}{\text{CBDdW}_{\text{SE3d}}(i) \times D_{adj}(i)} \quad [3l]$$	Regional $$\text{CBDdN}_{\text{SE3l}}(R_n) = 1 - \frac{1}{\text{CBDdW}_{\text{SE3d}}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$$ National $$\text{CBDdN}_{\text{SE3l}}(ES) = 1 - \frac{1}{\text{CBDdW}_{\text{SE3d}}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$$	It captures centrality through the geographical proximity of the SE to CDB. Its formulation is based on the weighted average of straight-line distances from SE k to the CBD within province i, rescaled and expressed as units of the diagonal of the province, which is set as the standard. Low values of the indicator (low centrality) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in the CBD. If the indicator is close to zero, people are, on average, as distant to the CDB as the diagonal of the province.
$CBDdN_{SE3m}$	***Normalised geographical centrality (SE & travel distance)***	$$\text{CBDdN}_{\text{SE3m}}(i) = 1 - \frac{1}{\text{CBDdW}_{\text{SE3e}}(i) \times D_{adj}(i)} \quad [3m]$$	Regional $$\text{CBDdN}_{\text{SE3m}}(R_n) = 1 - \frac{1}{\text{CBDdW}_{\text{SE3e}}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$$ National $$\text{CBDdN}_{\text{SE3m}}(ES) = 1 - \frac{1}{\text{CBDdW}_{\text{SE3e}}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$$	It captures centrality through the population proximity to CDB. Its formulation is based on the weighted average of travel distances from SE k to the CBD within province i, rescaled and expressed as units of the diagonal of the province, which is set as the standard. Low values of the indicator (low centrality) point out high dispersion. Its theoretical range is (0, 1). Values close to one mean that population is located in the CBD. If the indicator is close to zero, people are, on average, as distant to the CDB as the diagonal of the province.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$CBDdS_{MUN4a}$	***Inverse of the simple average of the straight-line distances from municipalities to CBD***	$$\text{CBDdS}_{\text{MUN4a}}(i) = (\mu_i - 1)[\mathbf{1}' \times T_{cbd}(i)]^{-1} \quad [4a]$$ Where: $\mathbf{1}$ is a column vector (μ_i - 1) × 1 with 1 in all entries. $\boldsymbol{T_{cbd}(i)}$ is a column vector (μ_i - 1) × 1 whose entry $d_{cbd,j}$ is the straight-line distance between the municipality j and the CBD of province i. It excludes the CBD.	Regional $$\text{CBDdS}_{\text{MUN4a}}(R_n) = \left[\sum_{i \in R_n} (\mu_i - 1)\right]\left[\sum_{i \in R_n} \mathbf{1}' \times T_{cbd}(i)\right]^{-1}$$ National $$\text{CBDdS}_{\text{MUN4a}}(\text{ES}) = \left[\sum_{i=1}^{50} (\mu_i - 1)\right]\left[\sum_{i=1}^{50} \mathbf{1}' \times T_{cbd}(i)\right]^{-1}$$	Same as $CBDdS_{SE3a}$ referred to municipalities.
$CBDdS_{MUN4b}$	***Inverse of the simple average of the travel distances from SE municipalities to CBD***	$$\text{CBDdS}_{\text{MUN4b}}(i) = (\mu_i - 1)\left[\mathbf{1}' \times \overline{T}_{cbd}(i)\right]^{-1} \quad [4b]$$ Where: $\boldsymbol{\overline{T}_{cbd}(i)}$ is a column vector (μ_i - 1) × 1 whose entry $td_{cbd,j}$ is the travel distance between the municipality j and the CBD of province i. It excludes the CBD. $\boldsymbol{td_{cbd,j}}$ is our estimated travel distance between the CBD and municipality j within province i. We calculated it using the ratio ρ_{1i} "travel distance to straight-line distance" for province i as follows: $td_{cbd,j} = \rho_{1i} d_i[CBD, j]$. We have estimated ρ_{1i} with a random sample of observation points, as described previously.	Regional $$\text{CBDdS}_{\text{MUN4b}}(R_n) = \left[\sum_{i \in R_n} (\mu_i - 1)\right]\left[\sum_{i \in R_n} \mathbf{1}' \times \overline{T}_{cbd}(i)\right]^{-1}$$ National $$\text{CBDdS}_{\text{MUN4b}}(\text{ES}) = \left[\sum_{i=1}^{50} (\mu_i - 1)\right]\left[\sum_{i=1}^{50} \mathbf{1}' \times \overline{T}_{cbd}(i)\right]^{-1}$$	Same as $CBDdS_{SE3b}$ referred to municipalities.
$CBDdS_{MUN4c}$	***Inverse of the simple average of the travel durations from municipalities to CBD***	$$\text{CBDdS}_{\text{MUN4c}}(i) = (\mu_i - 1)\left[\mathbf{1}' \times \overline{\overline{T}}_{cbd}(i)\right]^{-1} \quad [4c]$$ Where: $\boldsymbol{\overline{\overline{T}}_{cbd}(i)}$ is a column vector (μ_i - 1) × 1 whose entry $tdr_{cbd,j}$ is the travel duration between the municipality j and the CBD of province i. It excludes the CBD. $\boldsymbol{tdr_{cbd,j}}$ is our estimated travel duration between the CBD and municipality j within province i. We calculated it using the ratio ρ_{2i} "travel duration to straight-line distance" for province i as follows: $tdr_{cbd,j} = \rho_{2i} d_i[CBD, j]$. We have estimated ρ_{2i} with a random sample of observation points, as described previously.	Regional $$\text{CBDdS}_{\text{MUN4c}}(R_n) = \left[\sum_{i \in R_n} (\mu_i - 1)\right]\left[\sum_{i \in R_n} \mathbf{1}' \times \overline{\overline{T}}_{cbd}(i)\right]^{-1}$$ National $$\text{CBDdS}_{\text{MUN4c}}(\text{ES}) = \left[\sum_{i=1}^{50} (\mu_i - 1)\right]\left[\sum_{i=1}^{50} \mathbf{1}' \times \overline{\overline{T}}_{cbd}(i)\right]^{-1}$$	Same as $CBDdS_{SE3c}$ referred to municipalities.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$\text{CBDdW}_{\text{MUN4d}}$	***Inverse of the weighted average of the straight-line distances from municipalities to CBD***	$\text{CBDdW}_{\text{MUN4d}}(i) = \left[\frac{s'(i) \times T_{cbd}(i)}{s'(i) \times \mathbf{1}}\right]^{-1} \quad [4d]$ Where: $\boldsymbol{s'(i)}$ is a column vector $\mu_i \times 1$ of populations of province i singular entities: $s'(i) = (\pi_{i1}, \pi_{i2}, \dots, \pi_{i\mu_i})$ It includes the CBD.	Regional $\text{CBDdW}_{\text{MUN4d}}(R_n) = \left[\frac{\sum_{i \in R_n} \mathbf{1}' \times T_{cbd}(i)}{\sum_{i \in R_n} \mathbf{1}' \times s(i)}\right]^{-1}$ National $\text{CBDdW}_{\text{MUN4d}}(\text{ES}) = \left[\frac{\sum_{i=1}^{50} \mathbf{1}' \times T_{cbd}(i)}{\sum_{i=1}^{50} \mathbf{1}' \times s(i)}\right]^{-1}$	Same as $\text{CBDdW}_{\text{SE3d}}$ referred to municipalities.
$\text{CBDdW}_{\text{MUN4e}}$	***Inverse of the weighted average of the travel distances from municipalities to CBD***	$\text{CBDdW}_{\text{MUN4e}}(i) = \left[\frac{s'(i) \times \overline{T}_{cbd}(i)}{s'(i) \times \mathbf{1}}\right]^{-1} \quad [4e]$	Regional $\text{CBDdW}_{\text{MUN4e}}(R_n) = \left[\frac{\sum_{i \in R_n} \mathbf{1}' \times \overline{T}_{cbd}(i)}{\sum_{i \in R_n} \mathbf{1}' \times s(i)}\right]^{-1}$ National $\text{CBDdW}_{\text{MUN4e}}(\text{ES}) = \left[\frac{\sum_{i=1}^{50} \mathbf{1}' \times \overline{T}_{cbd}(i)}{\sum_{i=1}^{50} \mathbf{1}' \times s(i)}\right]^{-1}$	Same as $\text{CBDdW}_{\text{SE3e}}$ referred to municipalities.
$\text{CBDdW}_{\text{MUN4f}}$	***Inverse of the weighted average of the travel durations from municipalities to CBD***	$\text{CBDdW}_{\text{MUN4f}}(i) = \left[\frac{s'(i) \times \overline{\overline{T}}_{cbd}(i)}{s'(i) \times \mathbf{1}}\right]^{-1} \quad [4f]$	Regional $\text{CBDdW}_{\text{MUN4f}}(R_n) = \left[\frac{\sum_{i \in R_n} \mathbf{1}' \times \overline{\overline{T}}_{cbd}(i)}{\sum_{i \in R_n} \mathbf{1}' \times s(i)}\right]^{-1}$ National $\text{CBDdW}_{\text{MUN4f}}(\text{ES}) = \left[\frac{\sum_{i=1}^{50} \mathbf{1}' \times \overline{\overline{T}}_{cbd}(i)}{\sum_{i=1}^{50} \mathbf{1}' \times s(i)}\right]^{-1}$	Same as $\text{CBDdW}_{\text{SE3f}}$ referred to municipalities.
$\text{CBDdR}_{\text{MUN4g}}$ $\text{CBDdR}_{\text{MUN4h}}$ $\text{CBDdR}_{\text{MUN4i}}$	***Ratio population centrality to geographical centrality***[57]	$\text{CBDdR}_{\text{MUN4g}}(i) = \text{CBDdW}_{MUN4d}(i)/\text{CBDdS}_{\text{MUN4}_a}(i) \quad [4g]$ $\text{CBDdR}_{\text{MUN4h}}(i) = \text{CBDdW}_{MUN4e}(i)/\text{CBDdS}_{\text{MUN4}_b}(i) \quad [4h]$ $\text{CBDdR}_{\text{MUN4i}}(i) = \text{CBDdW}_{MUN4f}(i)/\text{CBDdS}_{\text{MUN4}_c}(i) \quad [4i]$	Regional $\text{CBDdR}_{\text{MUN4z}}(R_n) = \text{CBDdW}_{MUN4y}(R_n)/\text{CBDdS}_{\text{MUN4}_x}(R_n)$ National $\text{CBDdR}_{\text{MUN4z}}(ES) = \text{CBDdW}_{MUN4y}(ES)/\text{CBDdS}_{\text{MUN4}_x}(ES)$ $(z, y, x) = \begin{cases} (\text{g}, \text{d}, \text{a}) \\ (\text{h}, \text{e}, \text{b}) \\ (\text{i}, \text{f}, \text{c}) \end{cases}$	Same as $\text{CBDdR}_{\text{SE3g}}$ to $\text{CBDdR}_{\text{SE3i}}$ referred to municipalities.

[57] The ratio of weighted to unweighted average distance. These equations has been adapted from Lee, S. (2015).

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$CBDdN_{MUN4j}$	***Normalised geographical centrality (MUN & straight-line distance)***	$$\mathrm{CBDdN_{MUN4j}}(i) = 1 - \frac{1}{\mathrm{CBDdS_{MUN4a}}(i) \times D_{adj}(i)} \quad [4j]$$	Regional $$\mathrm{CBDdN_{SE4j}}(R_n) = 1 - \frac{1}{\mathrm{CBDdS_{MUN4a}}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$$ National $$\mathrm{CBDdN_{SE4j}}(ES) = 1 - \frac{1}{\mathrm{CBDdS_{MUN4a}}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$$	Same as $CBDdN_{SE3j}$ referred to municipalities.
$CBDdN_{MUN4k}$	***Normalised geographical centrality (MUN & travel distance)***	$$\mathrm{CBDdN_{MUN4k}}(i) = 1 - \frac{1}{\mathrm{CBDdR_{SE4b}}(i) \times D_{adj}(i)} \quad [4k]$$	Regional $$\mathrm{CBDdN_{MUN4k}}(R_n) = 1 - \frac{1}{\mathrm{CBDdS_{SE4b}}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$$ National $$\mathrm{CBDdN_{MUN4k}}(ES) = 1 - \frac{1}{\mathrm{CBDdS_{SE4b}}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$$	Same as $CBDdN_{SE3k}$ referred to municipalities.
$CBDdN_{MUN4l}$	***Normalised population centrality (MUN & straight-line distance)***	$$\mathrm{CBDdN_{MUN4l}}(i) = 1 - \frac{1}{\mathrm{CBDdW_{SE4d}}(i) \times D_{adj}(i)} \quad [4l]$$	Regional $$\mathrm{CBDdN_{MUN4l}}(R_n) = 1 - \frac{1}{\mathrm{CBDdW_{SE4d}}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$$ National $$\mathrm{CBDdN_{MUN4l}}(ES) = 1 - \frac{1}{\mathrm{CBDdW_{SE4d}}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$$	Same as $CBDdN_{SE3l}$ referred to municipalities.
$CBDdN_{MUN4m}$	***Normalised population centrality (MUN & travel distance)***	$$\mathrm{CBDdN_{MUN4m}}(i) = 1 - \frac{1}{\mathrm{CBDdW_{SE4e}}(i) \times D_{adj}(i)} \quad [4m]$$	Regional $$\mathrm{CBDdN_{MUN4m}}(R_n) = 1 - \frac{1}{\mathrm{CBDdW_{SE4e}}(R_n) \times \frac{\sum_{i \in R_n} D_{adj}(i)}{\#R_n}}$$ National $$\mathrm{CBDdN_{MUN4m}}(ES) = 1 - \frac{1}{\mathrm{CBDdW_{SE4e}}(ES) \times \frac{\sum_{i=1}^{50} D_{adj}(i)}{50}}$$	Same as $CBDdN_{SE3m}$ referred to municipalities.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$CBDdCR_{MUN4n}$	***Centralisation ratio (CR)***	$$CBDdCR_{MUN4n}(i) = 1 - \frac{\bar{s}_i}{\bar{s}_{Ui}} = \frac{\bar{s}_{Ui} - \bar{s}_i}{\bar{s}_{Ui}} \quad [4n]$$ Where: $$\bar{s}_i = \frac{\sum_{j=1}^{\mu_i} \pi_{ij} d_i[CBD,j]}{\sum_{j=1}^{\mu_i} \pi_{ij}}$$ $$\bar{s}_{Ui} = \frac{\sum_{j=1}^{\mu_i} \alpha_{ij} d_i[CBD,j]}{\sum_{j=1}^{\mu_i} \alpha_{ij}}$$	Regional $$CBDdCR_{MUN4n}(R_n) = \frac{\sum_{i \in R_n} \pi_i \times CBDdCR_{MUN4n}(i)}{\sum_{i \in R_n} \pi_i}$$ National $$CBDdCR_{MUN4n}(ES) = \frac{\sum_{i=1}^{50} \pi_i \times CBDdCR_{MUN4n}(i)}{\pi}$$	$\frac{\bar{s}}{\bar{s}_U}$ compares the mean distance population is located from the centre to the mean distance to the centre if population were uniformly distributed across the province with the same density in each municipality (Please refer to Ottensmann, J.R. (2017) for further details). If the population is actually distributed uniformly, the value of the ratio will be 1. As centralisation increases, the actual mean distance to the centre decreases and the ratio will decline to a minimum value of 0 if all units are located at the centre. This ratio is thus a measure of decentralisation, increasing as units are located farther from the centre. To create a measure of centralisation, this ratio is subtracted from 1, giving the centralisation ratio CR. The centralisation ratio is 0 for a uniform distribution of population and 1 if all the people are located in the centre. If the population were actually more decentralised than a uniform distribution, the centralisation ratio can be negative. Lower values reflect greater dispersion.
$CBDdACI_{MUN4o}$	***Centralisation index (ACI)***	It is computed as follows for a province: one draws a series of concentric rings from the CBD (2 Km in our case). Then one computes the accumulation of observations of land use from the innermost ring and working progressively outward. This cumulative distribution is compared to the corresponding distribution of land area as a baseline. $$CBDdACI_{MUN4o}(i) = \sum_{h=2}^{H_i} \tilde{\pi}_{h-1}\tilde{\alpha}_h - \sum_{h=2}^{H_i} \tilde{\pi}_h \tilde{\alpha}_{h-1} \quad [4o]$$ Where: h represents the h^{th} concentric ring (length 2 Km) H_i the total number of concentric rings For each province i: $$H_i^* = \frac{maximum\ distance\ from\ a\ Municipality\ to\ the\ CBD}{2}$$ $\tilde{\alpha}_h$ the accumulated proportion of land area of municipalities within ring h $\tilde{\pi}_h$ The accumulated proportion of population of municipalities within ring h * The maximum distance from a municipality to the CBD has been rounded to the nearest even number.	Regional $$CBDdACI_{MUN4o}(R_n) = \frac{\sum_{i \in R_n} \pi_i \times CBDdACI_{MUN4o}(i)}{\sum_{i \in R_n} \pi_i}$$ National $$CBDdACI_{MUN4o}(ES) = \frac{\sum_{i=1}^{50} \pi_i \times CBDdACI_{MUN4o}(i)}{\pi}$$	The centralisation index (ACI) measures how rapidly population in land uses accumulate relative to land area as one moves progressively outward in concentric rings from the CBD. It ranges between -1 and 1, with a larger value indicating a higher degree of centrality. Thus, lower values reflect greater dispersion. If all the population resides in the CBD, the ACI will be 1. On the contrary, if few people are located near the centre but most are instead near the edge, land area will accumulate faster that the particular population moving outward from the centre and centrality will have a low (even negative) value. A 0 value means that the population is uniformly distributed across the province. The index gives the proportion of people that require to change residence to achieve an uniform distribution around the CBD.

Nuclearity indicators

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$NUNoN_{SE5a}$	***Number of nuclei***	$\mathrm{NUNoN_{SE5a}(i)} = \#\mathrm{N}^{-1} \quad [5a]$ *Where #N is the number of nuclei in province i (SE-based).*	Regional $\mathrm{NUNoN_{SE5a}}(R_n) = \frac{\sum_{i \in R_n} \pi_i \times \mathrm{NUNoN_{SE5a}}(i)}{\sum_{i \in R_n} \pi_i}$ National $\mathrm{NUNoN_{SE5a}}(ES) = \frac{\sum_{i=1}^{50} \pi_i \times \mathrm{NUNoN_{SE5a}}(i)}{\pi}$	Nuclearity is maximised when the province has a mononuclear pattern of residential development: the CBD is the only nucleus. When the number of nuclei increases, this would point out more dispersion. Thus, the indicator for nuclearity is $\#N^{-1}$, which is low when nuclearity is low and dispersion is high. A nucleus is defined as an urban singular entity.
$NUSCBD_{SE5b}$	***Share of the population in the CBD over the population in nuclei***	$\mathrm{NUSCBD_{SE5b}(i)} = \frac{\pi_{iCBD}}{\sum_{k\ is\ a\ nucleus\ (SE-based)} \pi_{ik}} \quad [5b]$	Regional $\mathrm{NUSCBD_{SE5b}}(R_n) = \frac{\sum_{i \in R_n} \pi_i \times \mathrm{NUSCBD_{SE5b}}(i)}{\sum_{i \in R_n} \pi_i}$ National $\mathrm{NUSCBD_{SE5b}}(ES) = \frac{\sum_{i=1}^{50} \pi_i \times \mathrm{NUSCBD_{SE5b}}(i)}{\pi}$	The lower the share of the CBD in the total population of nuclei, the lower the nuclearity and the higher the dispersion.
$NUNoN_{MUN6a}$	***Number of nuclei***	$\mathrm{NUNoN_{MUN6a}(i)} = \#\mathrm{N}^{-1} \quad [6a]$ *Where #N is the number of nuclei in province i (MUN-based).*	Regional $\mathrm{NUNoN_{MUN6a}}(R_n) = \frac{\sum_{i \in R_n} \pi_i \times \mathrm{NUNoN_{MUN6a}}(i)}{\sum_{i \in R_n} \pi_i}$ National $\mathrm{NUNoN_{MUN6a}}(ES) = \frac{\sum_{i=1}^{50} \pi_i \times \mathrm{NUNoN_{MUN6a}}(i)}{\pi}$	Nuclearity is maximised when the province has a mononuclear pattern of residential development: the CBD is the only nucleus. When the number of nuclei increases, this would point out more dispersion. Thus, the indicator for nuclearity is $\#N^{-1}$, which is low when nuclearity is low and dispersion is high. A nucleus is defined as an urban municipality.
$NUSCBD_{MUN6b}$	***Share of the population in the CBD over the population in nuclei***	$\mathrm{NUSCBD_{MUN6b}(i)} = \frac{\pi_{iCBD}}{\sum_{j\ is\ a\ nucleus\ (MUN-based)} \pi_{ij}} \quad [6b]$	Regional $\mathrm{NUSCBD_{MUN6b}}(R_n) = \frac{\sum_{i \in R_n} \pi_i \times \mathrm{NUNoN_{MUN6b}}(i)}{\sum_{i \in R_n} \pi_i}$ National $\mathrm{NUSCBD_{MUN6b}}(ES) = \frac{\sum_{i=1}^{50} \pi_i \times \mathrm{NUNoN_{MUN6b}}(i)}{\pi}$	The lower the share of the CBD in the total population of nuclei, the lower the nuclearity and the higher the dispersion.

Density indicators

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$\mathrm{DEPWD}_{\mathrm{MUN7a}}$ $\mathrm{DEPWD}_{\mathrm{MUN7b}}$ $\mathrm{DEPWD}_{\mathrm{MUN7c}}$	***Population-weighted density (PWD)***	$\mathrm{DEPWD}_{\mathrm{MUN7a}}(i) = \sum_j \delta_{ij}^{0} \frac{\pi_{ij}}{\pi_i} \quad [7a]$ $\mathrm{DEPWD}_{\mathrm{MUN7b}}(i) = \sum_j \delta_{ij}^{1} \frac{\pi_{ij}}{\pi_i} \quad [7b]$ $\mathrm{DEPWD}_{\mathrm{MUN7c}}(i) = \sum_j \delta_{ij}^{2} \frac{\pi_{ij}}{\pi_i} \quad [7c]$	Regional $\mathrm{DEPWD}_{\mathrm{MUN7x}}(R_n) = \sum_{i\in R_n} \sum_j \delta_{ij}^{\nu} \frac{\pi_{ij}}{\sum_{i\in R_n} \pi_i}$ National $\mathrm{DEPWD}_{\mathrm{MUN7x}}(ES) = \sum_{i=1}^{50} \sum_j \delta_{ij}^{\nu} \frac{\pi_{ij}}{\pi}$ x = a, b, c for $\nu = 0,1,2$ respectively	Most populated municipalities have higher influence on the average density. $\mathrm{DEPWD}_{\mathrm{MUN7x}}$ is higher than δ^{ν}_{i} when the most populated municipalities within the province tend to be denser. $\mathrm{DEPWD}_{\mathrm{MUN7}x}(i) = \delta_i^{\nu}$ when all municipalities are equally dense. Please notice that, even if the population share of all the municipalities within a province is the same, $\mathrm{DEPWD}_{\mathrm{MUN7}x}(i) \neq \delta_i^{\nu}$. The lower the value of $\mathrm{DEPWD}_{\mathrm{MUN7x}}$ the higher the dispersion.
$\mathrm{DENMAX}_{\mathrm{MUN7d}}$ $\mathrm{DENMAX}_{\mathrm{MUN7e}}$ $\mathrm{DENMAX}_{\mathrm{MUN7f}}$	***Maximum density***	$DENMAX_{\mathrm{MUN7d}}(i) = \max_j\{\delta_{ij}^{0}\} \quad [7d]$ $DENMAX_{\mathrm{MUN7e}}(i) = \max_j\{\delta_{ij}^{1}\} \quad [7e]$ $DENMAX_{\mathrm{MUN7f}}(i) = \max_j\{\delta_{ij}^{2}\} \quad [7f]$	Regional $DENMAX_{\mathrm{MUN7y}}(R_n) = \frac{\sum_{i\in R_n} \pi_i \times DENMAX_{\mathrm{MUN7y}}(i)}{\sum_{i\in R_n} \pi_i}$ National $DENMAX_{\mathrm{MUN7y}}(ES) = \frac{\sum_{i=1}^{50} \pi_i \times DENMAX_{\mathrm{MUN7y}}(i)}{\pi}$ y = d, e, f for $\nu = 0,1,2$ respectively	Low values of the maximum density would be associated with overall low density over the province, thus, with high dispersion. However, high values of this indicator cannot be associated with low dispersion. We have included this indicator in our list just for descriptive purposes.
$\mathrm{DENMIN}_{\mathrm{MUN7g}}$ $\mathrm{DENMIN}_{\mathrm{MUN7h}}$ $\mathrm{DENMIN}_{\mathrm{MUN7i}}$	***Minimum density***	$DENMIN_{\mathrm{MUN7g}}(i) = \min_j\{\delta_{ij}^{0}\} \quad [7g]$ $DENMIN_{\mathrm{MUN7h}}(i) = \min_j\{\delta_{ij}^{1}\} \quad [7h]$ $DENMIN_{\mathrm{MUN7i}}(i) = \min_j\{\delta_{ij}^{2}\} \quad [7i]$	Regional $DENMIN_{\mathrm{MUN7z}}(R_n) = \frac{\sum_{i\in R_n} \pi_i \times DENMIN_{\mathrm{MUN7z}}(i)}{\sum_{i\in R_n} \pi_i}$ National $DENMIN_{\mathrm{MUN7z}}(ES) = \frac{\sum_{i=1}^{50} \pi_i \times DENMIN_{\mathrm{MUN7z}}(i)}{\pi}$ z = g, h, i for $\nu = 0,1,2$ respectively	High values of the minimum density would be associated with overall high density over the province, thus, with low dispersion. On the other hand, low values of this indicator cannot be associated with high dispersion. We have included this indicator in our list to complement the maximum density just for descriptive purposes although it does not follow the criterion selection "Low values of the indicator would be associated with high dispersion."
$\mathrm{DENHIGH}_{\mathrm{MUN7j}}$ $\mathrm{DENHIGH}_{\mathrm{MUN7k}}$ $\mathrm{DENHIGH}_{\mathrm{MUN7l}}$	***Share of the population living in high density municipalities***	$DENHIGH_{\mathrm{MUN7j}}(i) = \frac{\sum_{\delta_{ij}^{0} > \delta_0^{0}} \pi_{ij}}{\pi_i} \quad [7j]$ $DENHIGH_{\mathrm{MUN7k}}(i) = \frac{\sum_{\delta_{ij}^{1} > \delta_0^{1}} \pi_{ij}}{\pi_i} \quad [7k]$ $DENHIGH_{\mathrm{MUN7l}}(i) = \frac{\sum_{\delta_{ij}^{2} > \delta_0^{2}} \pi_{ij}}{\pi_i} \quad [7l]$	Regional $DENHIGH_{\mathrm{MUN7u}}(R_n) = \frac{\sum_{i\in R_n} \pi_i \times DENHIGH_{\mathrm{MUN7u}}(i)}{\sum_{i\in R_n} \pi_i}$ National $DENHIGH_{\mathrm{MUN7u}}(ES) = \frac{\sum_{i=1}^{50} \pi_i \times DENHIGH_{\mathrm{MUN7u}}(i)}{\pi}$ u = j, k, l for $\nu = 0,1,2$ respectively	δ_0^{ν} Has been determine as the mean value at the national level of the corresponding population-weighted densities: $\delta_0^{\nu} = \begin{cases} \delta_0^{0} = 2{,}478\ inh./Km^2 \\ \delta_0^{1} = 8{,}475\ inh./Km^2 \\ \delta_0^{2} = 12{,}379\ inh./Km^2 \end{cases}$ The lower the value the higher the dispersion.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$DENCBD_{MUN7m}$ $DENCBD_{MUN7n}$ $DENCBD_{MUN7o}$	***Density of land use in the CBD***	$DENCBD_{\mathrm{MUN7m}}\,(i) = \delta_{iCBD}^{0} \qquad [7m]$ $DENCBD_{\mathrm{MUN7n}}\,(i) = \delta_{iCBD}^{1} \qquad [7n]$ $DENCBD_{\mathrm{MUN7o}}\,(i) = \delta_{iCBD}^{2} \qquad [7o]$	Regional $DENCBD_{\mathrm{MUN7w}}\,(i) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} \delta_{iCBD}^{\nu}$ National $DENCBD_{\mathrm{MUN7w}}\,(i) = \sum_{i=1\ldots 50} \frac{\pi_i}{\pi} \delta_{iCBD}^{\nu}$ w = m, n, o for $\nu = 0{,}1{,}2$ respectively	The lower the value the higher the dispersion.

Concentration Indicators

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
CNGINI$_{SE8a}$	***Gini index for SE***	$CNGINI_{SE8a}(i) = I_{Gi}(\#) = 1 - \sum_{l=1}^{L} p_l \left[\sum_{h=1}^{l-1} 2q_h + q_l\right] \quad [8a]$ Where: $p_s = \frac{n_s}{\sum_{h=1}^{L} n_h} = \frac{n_s}{\#_i}$ is the share of SE with population π_{is} (relative frequency of the value π_{is}). $q_s = \frac{\pi_{is} n_s}{\sum_{h=1}^{L} \pi_{ih} n_h} = \frac{\pi_{is} n_s}{\pi_i}$ is the share of population in SE with population π_{is}. π_{il} with $l = 1, \ldots, L$ stands for the series of different values of the population in the singular entities of province i sorted in ascending order: $\pi_{il} \neq \pi_{is} \; for \; l \neq s$ $\pi_{il} < \pi_{is} \; for \; l < s$ n_l is the absolute frequency of singular entities with population π_{il} ($\sum_{l=1}^{L} n_l = \#_i$). The maximum value of $I_{Gi}(\#)$ is equal to 1 - p_L (Lemelin, A. (2004)). Please notice that when L = 1 then Gini index is 0, indicating even distribution across the just one only existing land use. This theoretical situation does not occur in Spain's provinces. Should it happen we would impute 1 as Gini index in the understanding that there is infinite empty singular entities (#) and just one concentrating all the population, in which case, with this operationalisation of the index, it will be 1 ($p_L = 1/\# = 0$).	Regional $CNGINI_{SE8a}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} CNGINI_{SE8a}(i)$ National $CNGINI_{SE8a}(ES) = \sum_{i=1 \ldots 50} \frac{\pi_i}{\pi} CNGINI_{SE8a}(i)$	This Gini index measures the concentration of the population of singular entities by comparing the actual distribution of population across SE with an even distribution, where all the singular entities have the same share of population. Therefore, it measures departure from evenness. Usually, the definition of the Gini index is geometric. It defines said index as the quotient between the area delimited by the Lorenz curve together with the representative diagonal of total dispersion and the area between the respective representations of total dispersion and total concentration. There are alternative ways to express and calculate this index that approximate the mentioned geometric calculation. In this work, we use an exact method proposed by Ferreira, E. et al. (1997). Theoretically, it ranges from 0 to 1. However, to reach the theoretical maximum it is necessary that the number of SE tends towards infinity in such a way that p_L tends towards 0. Otherwise, when p_L> 0, the maximum value is equal to 1 - p_L (Lemelin, A. (2004)). If there were an even distribution of the population, the Gini index would be 0, indicating absence of concentration. In the other extreme case, if there were a situation where the entire population of the province resides in one single singular entity, the Gini index would tend to 1, indicating total concentration. The indicator increases as the distribution is more heterogeneous, concentrating more population in some territories than in others. Therefore, lower values reflect greater dispersion. We highlight that, when used as a measure of spatial concentration, the Gini index does not take into account the proximity between the different geographical units where people reside.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$CNSTHEI_{SE8b}$	***Standardised Theil entropy index (SE)***	$$CNSTHEI_{SE8b}(i) = \frac{1}{ln\#_i}\sum_{k=1}^{\#_i}\frac{\pi_{ik}}{\pi_i}ln\frac{\#_i \times \pi_{ik}}{\pi_i} \quad [8b]$$ Please notice that this algorithm cannot be applied to data with a population share of zero. In our work, this case does exist as we only work with inhabited land uses. We only retained with zero population those having disappeared in 2017 and they have been excluded of the calculation of $CNSTHEI_{SE8b}(i)$.	Regional $$\mathrm{CNSTHEI_{SE8b}}(R_n) = \sum_{i \in R_n}\frac{\pi_i}{\sum_{i \in R_n}\pi_i}\mathrm{CNSTHEI_{SE8b}}(i)$$ National $$\mathrm{CNSTHEI_{SE8b}}(ES) = \sum_{i=1\ldots50}\frac{\pi_i}{\pi}\mathrm{CNSTHEI_{SE8b}}(i)$$	Based on the concept of entropy as the degree of disorder or homogeneity, this indicator measures departure from homogeneity (evenness) of the distribution of the population throughout singular entities. Where homogeneity is the situation where all the singular entities have the same share of population. The value of the index ranges from 0 (minimum geographic concentration), when all the singular entities have the same share of population to, 1 (maximum concentration), when all the population resides in only one singular entity. Thus, intermediate values capture varying degrees of spatial concentration: the greater the *Standardised Theil entropy index*, the higher the spatial concentration of the population. Please notice that the Theil entropy index $E(i) = \sum_{k=1}^{\#_i}\frac{\pi_{ik}}{\pi_i}ln\frac{\pi_i}{\pi_{ik}}$ ranges from 0 (when all the population resides in only one singular entity and $E(i) = -\ln(1) = 0$) to $\ln\#_i$ (when all the singular entities have the same share of population $\frac{\pi_{ik}}{\pi_i} = \frac{1}{\#_i}$) (Aiginger, K. et al. (2004)). To adapt it to our evaluation criteria: The lower the value, the higher the dispersion and obtein values between 0 to 1, we have done the adjustment adopted by Combes, P.P. et al. (2008) and Atienza, M. et al. (2012). The main criticism to this indicator according to the literature reviewed refers to the fact that it is not appropriate when the number of SE is small (Castañeda, C. (2007)).

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$CNSHHI_{SE8c}$	***Standardised Herfindahl[58] Index (SE)***	$$CNSHHI_{SE8c}(i) = \frac{\#_i H_{SE}(i) - 1}{\#_i - 1} \quad [8c]$$ Where: $$H_{SE}(i) = \sum_{k=1}^{\#_i} \left(\frac{\pi_{ik}}{\pi_i}\right)^2$$ Is the Herfindahl index SE-based.	Regional $$\mathrm{CNSHHI_{SE8c}}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} \mathrm{CNSHHI_{SE8c}}(i)$$ National $$\mathrm{CNSHHI_{SE8c}}(ES) = \sum_{i=1\dots50} \frac{\pi_i}{\pi} \mathrm{CNSHHI_{SE8c}}(i)$$	The value of the index ranges from 0 (minimum concentration/maximum dispersion), when all the singular entities have the same share of population, to 1 (maximum concentration). Please notice that the Herfindahl index $H_{SE}(i)$ ranges from $1/\#_i$ to 1. To normalise it and make it independent of the number of singular entities we have done the following transformation: $$\frac{H_{SE}(i) - 1/\#_i}{1 - (1/\#_i)}$$ $H_{SE}(i)$ thus shows whether the population of the SE is concentrated in a small number of them, giving more relevance to the largest SE by "square weighting." Some orientations concerning the scale of the Herfindahl index, stemming from the economic concentration field, would be (Lis-Gutiérrez, J.P. (2013); Zurita, J. (2014)): $H_{SE}(i) < 0.01$ very deconcentrated. $0.01 < H_{SE}(i) < 0.15$ deconcentrated. $0.15 < H_{SE}(i) < 0.25$ moderate concentration $0.25 < H_{SE}(i)$ high concentration.

[58] Also Herfindahl- Hirschman index. See PUCC (2009); Lis-Gutiérrez, J.P. (2013).

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$CNDCV_{MUN9a}$	***Coefficient of variation of densities***	$$CNDCV_{MUN9a}(i) = \frac{\underset{j}{\sigma}(\delta_{ij}^0)}{\delta_i^0} \qquad [9a]$$ Where σ_j stands for standard deviation of δ_{ij}^0 over the municipalities. $$\underset{j}{\sigma}(\delta_{ij}^0) = \left[\sum_{j=1}^{\mu_i} \frac{\alpha_{ij}}{\alpha_i}\left(\delta_{ij}^0 - \sum_{j=1}^{\mu_i} \frac{\alpha_{ij}}{\alpha_i}\delta_{ij}^0\right)^2\right]^{1/2}$$	Regional $$CNDCV_{MUN9a}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} CNDCV_{MUN9a}(i)$$ National $$CNDCV_{MUN9a}(ES) = \sum_{i=1\ldots50} \frac{\pi_i}{\pi} CNDCV_{MUN9a}(i)$$	If all the municipalities have the same population density, the standard deviation is zero; if only one municipality attracts all the population, the coefficient of variation tends to 1 as the number of municipalities increases. Nonetheless, we cannot say that this is the maximum attainable by the CV. There is not an upper bound for this indicator.
$CNHGD_{MUN9b}$	***Share of the population living in high density municipalities based on built-up land area***	$$CNHGD_{MUN9b}(i) = \frac{\sum_j \pi_{ij} I_{\delta_{ij}^2}}{\pi_i} \qquad [9b]$$ $$I_{\delta_{ij}^2} = \begin{cases} 1 \ \forall j \ \vert \delta_{ij}^2 > \delta_0^2 \\ 0 \ otherwise \end{cases}$$	Regional $$CNHGD_{MUN9b}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} CNHGD_{MUN9b}(i)$$ National $$CNHGD_{MUN9b}(ES) = \sum_{i=1\ldots50} \frac{\pi_i}{\pi} CNHGD_{MUN9b}(i)$$	This concentration rate measures the population shares of the most densely populated municipalities. Lower values reflect greater dispersion. δ_0^2 Has been stablished as the mean value at national level of the population-weighted residential densities: $$\boldsymbol{\delta_0^2} = 12{,}379\ inh./Km^2$$
$CNPDG_{MUN9c}$	***Population Density Gradient***	$$CNPDG_{MUN9c}(i) = \phi \qquad [9c]$$ Where: $$\delta_{ij}^0(d_{ij}) = \delta_{CBD} e^{-\phi d_{ij} \varepsilon}$$ $\delta_{ij}^0(d_{ij})$ Population density of municipality j of province i at distance $d_{ij} = d_i[j, CBD]$ from the CBD. δ_{CBD} Density at the CBD. e Base of natural logarithms. ϕ *Population Density Gradient.* ε Error term Which we estimate from the equation: $$Ln(\delta_{ij}^0(d_{ij})) = Ln(\delta_{CBD}) - \phi d_{ij} + \varepsilon$$	Regional $$CNPDG_{MUN9c}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} CNPDG_{MUN9c}(i)$$ National $$CNPDG_{MUN9c}(ES) = \sum_{i=1\ldots50} \frac{\pi_i}{\pi} CNPDG_{MUN9c}(i)$$	*"A readily grasped measure of population concentration."* (Ottensmann, J.R. (2017). The gradient ϕ is the rate at which density falls from the centre. A high value of ϕ means that density will decline sharply with increasing distance from the CBD, i.e., the population is concentrated in the CBD (Clark, C. (1951)). Lower density gradients reflect greater degrees of dispersion.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$CNGINI_{MUN9d}$	***Gini index for MUN based on population***[59]	$$CNGINI_{MUN9d}(i) = I_{Gi}(\mu) = 1 - \sum_{l=1}^{L} p_l \left[\sum_{h=1}^{l-1} 2q_h + q_l \right] \quad [9d]$$ Where: $$p_s = \frac{n_s}{\sum_{h=1}^{L} n_h} = \frac{n_s}{\pi_i}$$ is the share of municipalities with population π_{is} (relative frequency of the value π_{is}). $$q_s = \frac{\pi_{is} n_s}{\sum_{h=1}^{L} \pi_{ih} n_h} = \frac{\pi_{is} n_s}{\pi_i}$$ is the share of population in municipalities with population π_{is}. π_{il} with $l = 1, \dots, L$ stands for the series of different values of the population in the municipalities of province i sorted in ascending: $$\pi_{il} \neq \pi_{is} \; for \; l \neq s$$ $$\pi_{il} < \pi_{is} \; for \; l < s$$ n_l is the absolute frequency of municipalities with population π_{il} ($\sum_{l=1}^{L} n_l = \mu_i$)	Regional $$CNGINI_{MUN9d}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} CNGINI_{MUN9d}(i)$$ National $$CNGINI_{MUN9d}(ES) = \sum_{i=1 \dots 50} \frac{\pi_i}{\pi} CNGINI_{MUN9d}(i)$$	Same as $CNGINI_{SE8a}$ referred to municipalities. Please refer to the formulation of $CNGINI_{SE8a}(i)$.
$CNGINI_{MUN9e}$	***Gini index for MUN based on land areas***[60]	$$CNGINI_{MUN9e}(i) = I_{Gi}(\alpha) = 1 - \sum_{l=1}^{L} p_l \left[\sum_{h=1}^{l-1} 2q_h + q_l \right] \quad [9e]$$ Where: $$p_s = \frac{n_s}{\sum_{h=1}^{L} n_h} = \frac{n_s}{\alpha_i}$$ is the share of km² (surface) with population density δ_{is}^0 (relative frequency of the value δ_{is}^0). $$q_s = \frac{\delta_{is}^0 n_s}{\sum_{h=1}^{L} \delta_{ih}^0 n_h} = \frac{\delta_{is}^0 n_s}{\pi_i}$$ is the share of population in municipalities with population density δ_{is}^0. δ_{il}^0 with $l = 1, \dots, L$ stands for the series of different values of the population density in the municipalities of province i sorted in ascending: $$\delta_{il}^0 \neq \delta_{is}^0 \; for \; l \neq s$$ $$\delta_{il}^0 < \delta_{is}^0 \; for \; l < s$$ n_l is the absolute frequency in terms of km² of the municipalities with density δ_{il}^0 ($\sum_{l=1}^{L} n_l = \alpha_i$)	Regional $$CNGINI_{MUN9e}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} CNGINI_{MUN9e}(i)$$ National $$CNGINI_{MUN9e}(ES) = \sum_{i=1 \dots 50} \frac{\pi_i}{\pi} CNGINI_{MUN9e}(i)$$	This Gini index measures the concentration of the population of municipalities by comparing the actual distribution of population across MUN with an even distribution: all the municipalities have the same population density. Thus, the benchmark for the comparison is the distribution of the areas of the municipalities. The Gini index ranges from 0 to 1. The indicator takes the value 0 when the population is evenly distributed in the territory and increases as said distribution is more heterogeneous concentrating more population in some territories than others. It takes the value 1 when the entire population is concentrated in a single municipality. Lower values reflect greater dispersion. Please refer to the formulation of $CNGINI_{MUN9d}(i)$

[59] Referred to the variable: MUN population size.
[60] Referred to the variable: MUN population density.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
CNSTHEI$_{MUN9f}$	***Standardised Theil Entropy index (MUN)***	$CNSTHEI_{MUN9f}(i) = \frac{1}{ln\mu_i}\sum_{j=1}^{\mu_i}\frac{\pi_{ij}}{\pi_i}ln\frac{\mu_i \times \pi_{ij}}{\pi_i}$ [9f]	Regional $\mathrm{CNSTHEI}_{\mathrm{MUN9f}}(R_n) = \sum_{i\in R_n}\frac{\pi_i}{\sum_{i\in R_n}\pi_i}\mathrm{CNSTHEI}_{\mathrm{MUN9f}}(i)$ National $\mathrm{CNSTHEI}_{\mathrm{MUN9f}}(ES) = \sum_{i=1\ldots50}\frac{\pi_i}{\pi}\mathrm{CNSTHEI}_{\mathrm{MUN9f}}(i)$	Based on the concept of entropy as the degree of disorder or homogeneity, this indicator measures departure from homogeneity (evenness), which is the situation where all the municipalities have the same share of population. The value of the index ranges from 0 (minimum geographic concentration), when all the singular entities have the same share of population to, 1 (maximum concentration), when all the population resides in only one municipality. Thus, intermediate values capture varying degrees of spatial concentration: the greater the Standardised Theil entropy index, the higher the spatial concentration of the population. Please notice that the Theil entropy index $E(i) = \sum_{j=1}^{\mu_i}\frac{\pi_{ij}}{\pi_i}ln\frac{\pi_i}{\pi_{ij}}$ ranges from 0 (maximum geographic concentration) to $\ln\mu_i$ (minimum spatial concentration). To adapt it to our evaluation criteria: The lower the value, the higher the dispersion and obtein values between 0 to 1, we have done the adjustment adopted by Atienza, M. et al. (2012). The main criticism to this indicator according to the literature reviewed refers to the fact that it is not appropriate when the number of MUN is small (Castañeda, C. (2007)).
CNTHI$_{MUN9g}$	***Theil index***	$\mathrm{CNTHI}_{\mathrm{MUN9g}}\ (\mathrm{i}) = ln\frac{\eta}{\tilde{\eta}}$ [9g] *Where*: $\eta = \sum_{j=1}^{\mu_i}\frac{\alpha_{ij}}{\alpha_i}\delta_{ij}^0$ $ln\tilde{\eta} = \sum_{j=1}^{\mu_i}\frac{\alpha_{ij}}{\alpha_i}\ln\delta_{ij}^0$ *Also*: $\mathrm{CNTHI}_{\mathrm{MUN9g}}\ (\mathrm{i}) = \sum_{j=1}^{\mu_i}\frac{\alpha_{ij}}{\alpha_i}\ln\left(\frac{\alpha_{ij}/\alpha_i}{\pi_{ij}/\pi_i}\right)$	Regional $\mathrm{CNTHI}_{\mathrm{MUN9g}}\ (R_n) = \sum_{i\in R_n}\frac{\pi_i}{\sum_{i\in R_n}\pi_i}\mathrm{CNTHI}_{\mathrm{MUN9g}}\ (i)$ National $\mathrm{CNTHI}_{\mathrm{MUN9f}}\ (ES) = \sum_{i=1\ldots50}\frac{\pi_i}{\pi}\mathrm{CNTHI}_{\mathrm{MUN9g}}\ (i)$	Named also Theil information measure and mean logarithmic deviation, it is a measure of dispersion of the densities of the municipalities. It is an alternative to the Gini index based on land areas. The index ranges from zero onwards. The mean logarithmic deviation has a minimum value of zero (if there is an even spatial distribution of population: equal densities in each municipality), but has no upper limit.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
CNSHHI_{MUN9h}	***Standardised Herfindahl index (MUN)***	$$\text{CNSHHI}_{MUN9h}(i) = \frac{\mu_i H_{MUN}(i) - 1}{\mu_i - 1} \qquad [9h]$$ Where: $$H_{MUN}(i) = \sum_{j=1}^{\mu_i} \left(\frac{\pi_{ij}}{\pi_i}\right)^2$$ Is the Herfindahl index MUN-based.	Regional $$\text{CNSHHI}_{MUN9h}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} \text{CNSHHI}_{MUN9h}(i)$$ National $$\text{CNSHHI}_{MUN9h}(ES) = \sum_{i=1\ldots50} \frac{\pi_i}{\pi} \text{CNSHHI}_{MUN9h}(i)$$	For province i, the value of the index ranges from 0 (minimum concentration/maximum dispersion), when all the municipalities have the same share of population, to 1 (maximum concentration). Please notice that the Hirschman-Herfindahl index $H_{MUN}(i)$ ranges from $1/\mu_i$ to 1. To standardise it and make it independent of the number of municipalities we have done the following transformation: $\frac{H_{MUN}(i) - (1/\mu_i)}{1 - (1/\mu_i)}$. $H_{MUN}(i)$ thus shows whether the population of the MUN is concentrated in a small number of them, giving more relevance to the largest MUN by square weighting. Some orientations concerning the scale of the Herfindahl index, stemming from the economic concentration field, would be (Lis-Gutiérrez, J.P. (2013); Zurita, J. (2014)): $H_{SE}(i) < 0.01$ very deconcentrated. $0.10 < H_{SE}(i) < 0.15$ deconcentrated. $0.15 < H_{SE}(i) < 0.25$ moderate concentration. $0.25 < H_{SE}(i)$ high concentration. The Herfindahl index is inferior to the Gini coefficient as an indicator for measuring geographic concentration in the sense that the former does not take into account the differences among the areas of the regions (Aso, Y. (2008)).
CNPCI_{MUN9i}	***Raw geographic concentration Index***	$$\text{CNPCI}_{\text{MUN9i}}(i) = G(i) = \sum_{j=1}^{\mu_i} \left(\frac{\pi_{ij}}{\pi_i} - \frac{\alpha_{ij}}{\alpha_i}\right)^2 \qquad [9i]$$	Regional $$\text{CNPCI}_{\text{MUN9i}}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} \text{CNPCI}_{\text{MUN9i}}(i)$$ National $$\text{CNPCI}_{\text{MUN9i}}(ES) = \sum_{i=1\ldots50} \frac{\pi_i}{\pi} \text{CNPCI}_{\text{MUN9i}}(i)$$	$G(i) \geq 0$. The indicator takes the value 0 when the population is homogeneously distributed in the territory and increases as the distribution is more heterogeneous, concentrating more population in some territories than in others. Homogeneously distributed in the territory means that population shares are equal to surface shares (equal densities in each municipality), in which case the index is 0. On the contrary, an index greater than 0 indicates the existence of other agglomeration-generating factors that go beyond the surface area of the municipality.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$\text{CNEG}_{\text{MUN9j}}$	***Ellison and Glaesser***[61]	$$\text{CNEG}_{\text{MUN9j}} = \gamma = \frac{\bar{G}(i) - H(i)}{1 - H(i)} \qquad [9j]$$ Where: $$\bar{G}(i) = \frac{G(i)}{1 - \sum_{j=1}^{\mu_i} \left(\frac{\alpha_{ij}}{\alpha_i}\right)^2}$$ $$H(i) = \frac{1}{\pi_i}$$ In detail: $$\gamma = \frac{\sum_{j=1}^{\mu_i} \left(\frac{\pi_{ij}}{\pi_i} - \frac{\alpha_{ij}}{\alpha_i}\right)^2 - \left(1 - \sum_{j=1}^{\mu_i} \left(\frac{\alpha_{ij}}{\alpha_i}\right)^2\right)\left(\frac{1}{\pi_i}\right)}{\left(1 - \sum_{j=1}^{\mu_i} \left(\frac{\alpha_{ij}}{\alpha_i}\right)^2\right)\left(1 - \frac{1}{\pi_i}\right)}$$ Please notice that H(*i*) in this formula is the translation to the population concentration ambit of the Herfindahl index used by Ellison and Glaesser in the context of industrial concentration. We set a parallelism between plants location decisions and population settlement decisions. This formulation is different from H_{MUN}(*i*).	Regional $$\text{CNEG}_{\text{MUN9j}}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} \text{CNEG}_{\text{MUN9j}}\,(i)$$ National $$\text{CNEG}_{\text{MUN9j}}\,(ES) = \sum_{i=1\ldots50} \frac{\pi_i}{\pi} \text{CNEG}_{\text{MUN9j}}(i)$$	Translated to population concentration, Ellison and Glaesser index is a normalised comparison between the distributions of the people against the benchmark of land area distribution. The index is a measure of excess-concentration with respect to land areas concentration. The scale of γ is such that one can interpret a value of 0 as indicating a complete lack of concentration. Indeed, the index is scaled so that it takes on a value of 0 if the population is as concentrated as it would be expected to be had the people chosen settlement locations randomly (Ellison, G. et al. (1997)). In turn, the random location process is defined in such a way that it would lead the population to reproduce the spatial distribution patterns of municipal areas within the province: $\frac{\pi_{ij}}{\pi_i} = \frac{\alpha_{ij}}{\alpha_i}$. (Callejon, M. (1998)). Positive values can be interpreted as a measure of the degree of the population concentration, as they reflect the excess of spatial population concentration above what would be expected if the population were randomly settle in the territory (uniform density across municipalities). Lower values reflect greater dispersion. It can be seen that γ can take negative values. We interpret negative values as a greater de-concentration of the population than the random location ($\gamma = 0$), which could point out a more marked tendency to polarisation. The index is comparable across territories and over time regardless of differences in the level of geographic aggregation. On the other hand, The literature review has shown that the degree of concentration increases with the size of the chosen spatial units. The fact that concentration measures are sensitive to the size and shape of territories is

[61] Also Maurel and Sedillot index. Ellison and Glaesser proposed an index for industrial geographic concentration that we can also use for population concentration (Jurado, I. et al. (2013)).

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
				widely addressed as the Modifiable Area Problem (MAUP). Likewise for the Gini and the Herfindahl indices, there seems to be a certain consensus about the choice of local units as the most appropriate (Santa Maria, M.J. et al. (2005)). In this work, we use local units: singular entity and municipalities. In general, care should be taken when the population size is smaller than the number of spatial units (Bertinelli, L. et al. (2005)), which is not the case in our analysis. According to Alonso, O. (2006) and Van Egeraat, C. (2016), it would be: $\gamma < 0.02$ low concentration $0.02 < \gamma < 0.05$ moderate concentration $0.05 < \gamma$ high concentration
$\mathrm{CNDI_{MUN9k}}$	***Delta index (also Hoover index)***	$\mathrm{CNDI_{MUN9k}}\,(i) = DDI_i = \frac{1}{2}\sum_{j=1}^{\mu_i}\left\lvert\frac{\pi_{ij}}{\pi_i} - \frac{\alpha_{ij}}{\alpha_i}\right\rvert \quad [9k]$	Regional $\mathrm{CNDI_{MUN9k}}\,(R_n) = \sum_{i\in R_n}\frac{\pi_i}{\sum_{i\in R_n}\pi_i}\mathrm{CNDI_{MUN9k}}\,(i)$ National $\mathrm{CNDI_{MUN9k}}\,(ES) = \sum_{i=1\ldots 50}\frac{\pi_i}{\pi}\mathrm{CNDI_{MUN9k}}\,(i)$	Delta is a specific application of the more general index of dissimilarity. Delta index computes the proportion of population residing in municipalities with above average density of population that would have to move in order to achieve a perfectly even distribution: one with uniform density (equal densities in each municipality). It ranges from 0 to 1 and lower values reflect greater dispersion. An even distribution of the population (maximum dispersion) across the municipalities is assessed with DDI = 0, when the population shares and the area shares in the comparison are exactly equal in all the municipalities. On the contrary, full concentration of the population in one single point is assessed with DDI = 1. A value of 0 means that the population is perfectly dispersed across the province (i.e., all municipalities have the same population density) and a value of 1 indicates that all the population resides in one point of an infinitely small unit of analysis.

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$\mathrm{CNMDDI_{MUN9l}}$	***Massey and Denton dissimilarity index for urban land***	$\mathrm{CNMDDI_{MUN9l}}\ (\mathrm{i}) = IOD_i[1] = \sum_{j=1}^{\mu_i} \frac{\alpha_{ij}\left\lvert \frac{\Delta_{ij}^1}{\alpha_{ij}} - \frac{\Delta_i^1}{\alpha_i} \right\rvert}{2\alpha_i \left(\frac{\Delta_i^1}{\alpha_i}\right)\left(1 - \frac{\Delta_i^1}{\alpha_i}\right)} \quad [9l]$	Regional $\mathrm{CNDCV_{MUN9l}}\ (R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} \mathrm{CNMDDI_{MUN9l}}$ National $\mathrm{CNDCV_{MUN9l}}(ES) = \sum_{i=1\ldots50} \frac{\pi_i}{\pi} \mathrm{CNMDDI_{MUN9l}}$	Adapted to measure the dissimilarity between the distribution of urban land and evenness, where evenness means equal share of urban land across all the municipalities. The index stands for the proportion of urban land that would have to relocate itself to achieve an even distribution. The value of this IOD measures the divergence from evenness of urban land developments relative to those of total land by taking the weighted mean of the absolute differences between urban land proportion in each municipality and the urban land proportion at the provincial level (P=$\frac{\Delta_i^1}{\alpha_i}$). This mean is rescaled with P(1-P), its maximum attainable value, so that the IOD ranges from 0, the most even distribution of urban land, to 1, the maximum value, which reflects uneven distribution. The degree of population dispersion of the province is maximised when the IOD takes a cero value: when every municipality has the same relative amount of urban land as in the province as a whole. The degree of population dispersion is minimised with an IOD value of 1, when urban land concentrates in one municipality. Lower values reflect greater dispersion.
$\mathrm{CNMDDI_{MUN9m}}$	***Massey and Denton dissimilarity index for built-up land***	$\mathrm{CNMDDI_{MUN9m}}(\mathrm{i}) = IOD_i[2] = \sum_{j=1}^{\mu_i} \frac{\alpha_{ij}\left\lvert \frac{\Delta_{ij}^2}{\alpha_{ij}} - \frac{\Delta_i^2}{\alpha_i} \right\rvert}{2\alpha_i \left(\frac{\Delta_i^2}{\alpha_i}\right)\left(1 - \frac{\Delta_i^2}{\alpha_i}\right)} \quad [9m]$	Regional $\mathrm{CNDCV_{MUN9m}}\ (R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} \mathrm{CNMDDI_{MUN9m}}\ (\mathrm{i})$ National $\mathrm{CNDCV_{MUN9m}}\ (ES) = \sum_{i=1\ldots50} \frac{\pi_i}{\pi} \mathrm{CNMDDI_{MUN9m}}\ (\mathrm{i})$	Adapted to measure the dissimilarity between the distribution of built-up land and evenness, where evenness means equal share of built-up land across all the municipalities. The index stands for the proportion of built-up land that would have to relocate itself to achieve an even distribution. The value of this IOD measures the divergence from evenness of urban land developments relative to those of total land by taking the weighted mean of the absolute differences of between urban land proportion in each municipality and the urban land

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
				proportion at the provincial level (P=$\frac{\Delta_i^1}{\alpha_i}$). This mean is rescaled with P(1-P), its maximum attainable value, so that the IOD ranges from 0, the most even distribution of built-up land, to 1, the maximum value, which reflects uneven distribution. The degree of population dispersion of the province is maximised when the IOD takes a cero value: when every municipality has the same relative amount of built-up land as in the province as a whole. The degree of population dispersion is minimised with an IOD value of 1, when built-up land concentrates in one municipality. Lower values reflect greater dispersion.

Continuity Indicators

INDICATOR	DEFINITION	BASIC FORMULATION (Province i)	AGREGATION TO REGIONAL (R) AND NATIONAL (ES) LEVEL	RATIONALE
$CNTRUT_{PROV10a}$	***Ratio urban land area to total land area***	$CNTRUT_{PROV10a} = \frac{\Delta_i^1}{\alpha_i} \quad [10a]$	Regional $CNTRUT_{PROV10a}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} CNTRUT_{PROV10a}(i)$ National $CNTRUT_{PROV10a}(ES) = \sum_{i=1...50} \frac{\pi_i}{\pi} CNTRUT_{PROV10a}(i)$	The lower the ratio the lower the continuity and the greater the dispersion.
$CNTRBT_{PROV10b}$	***Ratio built-up land area to total land area***	$CNTRBT_{PROV10b} = \frac{\Delta_i^2}{\alpha_i} \quad [10b]$	Regional $CNTRBT_{PROV10b}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} CNTRBT_{PROV10b}(i)$ National $CNTRBT_{PROV10b}(ES) = \sum_{i=1...50} \frac{\pi_i}{\pi} CNTRBT_{PROV10b}(i)$	The lower the ratio the lower the continuity and the greater the dispersion.
$CNTR2_{PROV10c}$	***R-square of the exponential density function***	$CNTR2_{PROV10c}(i) = R^2 \quad [10c]$ Where: R^2 is the determination coefficient resulting from applying OLS regression to the equation $\ln \delta_{ij}^0(d_{ij}) = \delta_0 - \phi d_{ij} + \varepsilon$	Regional $CNTR2_{PROV10c}(R_n) = \sum_{i \in R_n} \frac{\pi_i}{\sum_{i \in R_n} \pi_i} CNTR2_{PROV10c}(i)$ National $CNTR2_{PROV10c}(ES) = \sum_{i=1...50} \frac{\pi_i}{\pi} CNTR2_{PROV10c}(i)$	The lower the R^2 the lower the continuity and the greater the dispersion.

ANNEX V: APPROXIMATION FOR THE MAXIMUM SPATIAL SEPARATION ATTAINABLE

In this annex, we explain the way we have proceeded for providing further elements that would support our approach to approximate the maximum spatial separation attainable by the municipalities within a province.

We start with the concept by Venables of Spatial Separation as defined in point 3: a population-based indicator that measures population separation instead of that of locations, which formulated using distances between municipalities together with the population shares of the municipalities.

According to Pereira, R.H.M. et al. (2013, 2015), in a region forming a perfect circle, the maximum spatial separation occurs when the population is evenly distributed along the external edge. [62] Where evenly means equidistant and with the same population:

> "*In the very simple square grids presented in the next section, Vmax is obtained when each corner has one-fourth of the total employment. In a region forming a perfect circle, the maximum value of V occurs when all employment is evenly distributed along the external edge*."
>
> *"Thus, we have chosen to consider the 'opposite of maximum proximity' as a homogeneous distribution of values along the edge of a map. Although this solution is not the global maximum of V, it may be considered a satisfactory solution."*

Thus, an upper bound for the maximum spatial separation in a province would be the Spatial Separation Index when its municipalities are evenly distributed on the external edge of the bounding circle of the province. For the Spanish provinces, the percentage of the circle's surface covered by the province's land area of the bounding circle may

[62] As for the circle, when the number of municipalities (#MUN) is 2, we have checked that equidistant locations of the municipalities on the edge of the bounding circle produce maximum values of the average distance between them (simple average). For #MUN = 3 and 4 we have checked just the first order conditions. We have verified that a maximum occurs in the equation:

$$2r\left[\sum_{j=1}^{\#MUN-1} sin\left(\frac{\varphi_j}{2}\right)+\sum_{j=1}^{\#MUN-2} sin\left(\frac{\varphi_j+\varphi_{j+1}}{2}\right)+\sum_{j=1}^{\#MUN-3} sin\left(\frac{\varphi_j+\varphi_{j+1}+\varphi_{j+2}}{2}\right)+\cdots+\sin\left(\frac{\varphi_1+\varphi_2+\cdots+\varphi_{\#mun-1}}{2}\right)\right]$$

conditioned to $\varphi_j \neq \pi \; for \; \#MUN > 2$ when $\varphi_j = \frac{2\pi}{\#MUN} \; \forall j$. This equation is the sum of the chords (distances) between the municipalities on the edge of the circle; and φ_j is the angle of the circle's chord between municipalities j and j+1 (j = 1 to #MUN-1). It goes beyond of the scope of this work to explore the possibility of a general demonstration.

differ considerably between provinces depending on the shape of the province and might produce a lack of comparability of the mentioned upper bound.

To improve the comparability, we use the bounding box instead of the bounding circle. That is to say, we hold that an adequate proxy of the maximum spatial separation attainable is the Spatial Separation Index between the municipalities within a province when they are evenly distributed on the edge of the bounding box, where evenly means equidistant and with the same population (equal population weights).

Having moved from the circle to the rectangular model, to check this extreme in practical terms would require that we verify that for any set of population weights (ω) of the municipalities within a province, and for any equidistant ordination of them along the edge of the bounding box (σ), its Spatial Separation Index $V_{edge}(\omega, \sigma)$ is lower than or practically equal to VI_{max}. Please notice that $VI_{max} = V_{edge}(\omega_1, \sigma)$ where ω_1 is the set of population weights all equal for all municipalities, $\forall\ \sigma$.

To this end, we have developed simulations as follows:

- We have designed 100 stylised square provinces Si, $i = 1\ to\ 100$, with different sizes l and different numbers of municipalities μ, $Si = S(l, \mu)$. The size and the number of municipalities that characterise each province is in Annex V. Table 1. For example, $S1$ is a province with 48 municipalities and a square shape of side $l = 50\ Km$
- For each province i, we have calculated the triangular distances matrix $S(i)$ with the location distances between municipalities when they are placed equidistantly on the edge of the province
- We have created 350 sets of stylised population weights ω_s, $s =$ 1 to 350 for the municipalities of a province. These sets ω_s are structurally the same no matter the number of municipalities of the province $\omega_s(\mu) = \omega_s\ \forall \mu, s = 1\ to\ 350$. The set of population weights of a province's municipalities depends on the number of municipalities but we have designed them with a structural approach to simplify, without loss of generality. The sets of stylised population

weights have the same "*root*" components and differ on the negligible population weights allocated to the "*tail*". With our design, what depends on the number of municipalities is the "*tail*". The set of weights ω_1 correspond to a distribution of the population among the municipalities with all population weights being equal; simulations ω_s, $s = 301$ to 350, reflect in a stylised way the weights of the municipalities for the 50 Spanish provinces

- For each province i and for each set of weights ω_s, we have started with the municipalities in decreasing order of population share, placing the first one on the upper-left corner of the square and the rest from left to right equidistantly
- We have produced 2,000 ordinations σ of the set ω_s, $\sigma = 1\ to\ 2{,}000$ [63]
- We have calculated the Spatial Separation Indices $V_{edge}(l, \mu, \omega_s, \sigma)$
- As the simulation ω_1 corresponds to an even distribution of the population among the municipalities $VI_{max}(l, \mu) = V_{edge}(l, \mu, \omega_1, \sigma)$, ∀ σ
- We have represented each set ω_s by its Gini Index of population concentration $GI\omega_s$ to facilitate the analysis through graphic representations; please see Annex V Chart 1.

Our conclusions are:

- For each province i, the set of points ($GI\omega_s$;$V_{edge}(l, \mu, \omega_s, \sigma)$) is a bounded region whose shape and size are independent of the number of municipalities
- The bounded regions seems to be homothetic with size depending on the square side l (province size) but not on μ. The larger the size of the province, the larger the bounded region
- We noticed that the variability of the observations $V_{edge}(l, \mu, \omega_s, \sigma)$ given $l, \mu\ and\ \omega_s$ depends on the population concentration: the higher the Gini index of the set ω_s, the higher the variability
- When the concentration of the population tends to cero the observations $V_{edge}(l, \mu, \omega_s, \sigma)$ tend to $VI_{max}(l, \mu)$ and present very low variability

[63] 2,000 for $S(1)\ to\ S(40)$) and 500 for $S(41)\ to\ S(100)$), because of time processing reasons.

- The upper bound of that region ($\underset{\omega,\sigma}{\mathrm{Max}}(V_{edge}(l,\mu,\omega_s,\sigma))$ does not depend on the number of municipalities. It only depends on the size of the square l. Moreover, $VI_{max}(l,\mu) = VI_{max}(l)$. Please refer to Annex V Table 2
- We have verified that, for a given province i and a given ω_s, $\underset{\sigma}{\mathrm{Max}}(V_{edge}(l,\mu,\omega_s,\sigma))$ may overpass $VI_{max}(l)$. We estimate that $\underset{\omega,\sigma}{\mathrm{Max}}(V_{edge}(l,\mu,\omega_s,\sigma))$ overpasses $VI_{max}(l)$ with an average difference of around 6%, ranging from 3% to 9%, independently of the side l
- Normally, the weights for which $\underset{\sigma}{\mathrm{Max}}(V_{edge}(l,\mu,\omega_s,\sigma))$ overpass $VI_{max}(l)$ more intensively have high Gini index
- We have also verified that, for a given province i and a given set of weights ω_s, the difference between $VI_{max}(l)$ and the upper limit (I_U) of the 95% confidence interval for the values of $V_{edge}(l,\mu,\omega_s,\sigma)$ is on average around 3%, ranging from 0% to 4%, independently of the size l and the number of municipalities μ
- Therefore, we consider that the Spatial Separation Index between the municipalities within a province when they are equidistant on the edge of the bounding box and the population is evenly distributed among them, VI_{max}, is an adequate proxy of the maximum spatial separation attainable.

We notice that VI_{max} is not an absolute maximum. Indeed, as shown in the simulations, there are some situations where specific distributions of the population among municipalities present spatial separations greater than VI_{max}. We have proceeded to verify through general simulations that the empirical evidence supports that VI_{max} is "typically" an adequate proxy of the maximum spatial separation attainable.

These results are corroborated by the real life specific data of Spain's provinces.

For each of Spain's province, we have calculated the Spatial Separation Indices VI_{max} and V_{edge} corresponding to the actual distribution of its population among municipalities when they are equidistant on the edge of the bounding box of the province, starting at the upper left corner and following the INE's code. In addition, we have generated 4000 permutations of the municipalities and their associated V_{edge}.

We present a summary of our findings in Annex V Table 3. They show that VI_{max} is very close to the maximum value of the 4000 spatial separations calculated and to the upper limit (I_U) of the 99.9% confidence interval for those values, with some exceptions.

For Álava, Zaragoza, Valladolid, and Madrid VI_{max} is very high compared to the maximum spatial separation of the 4000 permutations and to I_U. We have checked with the general simulations that VI_{max} is as expected for provinces of that size and μ. Thus, we conclude that, in fact, the specific distribution of the population (population weights) among the municipalities of these provinces yields low spatial separation indices. We understand that in this case VI_{max} is a proper estimate for the maximum attainable spatial separation.

On the other hand, for Asturias VI_{max} is very low compared to the maximum spatial separation of the 4000 permutations and to I_U. This is because both real life upper bounds for this province are very high. This points out that the specific distribution of Asturias' population among its municipalities may yield high spatial separation on the edge, well overpassing VI_{max}. Indeed, for all Si, the simulations based on the stylised weights of Asturias' municipalities (ω_{312}) yield values of $\underset{\sigma}{\mathrm{Max}}(\mathrm{V}_{edge}(\omega_{312}, \sigma))$ systematically greater than VI_{max}.

In this particular case, VI_{max} could not be a proper upper bound for the maximum spatial separation attainable and using it might underestimate proximity. Therefore, we propose to correct VI_{max} for Asturias with the factor:

$$\frac{\text{Maximum spatial separation of the 4000 permutations}}{VI_{max}} = 1.11$$

Annex V. Table 1

Parameters for the simulations on V_{edge} **It continues**

	Number of municipalities μ	Length of the Square (Km) l	Number of Weights	Number of Ordinations
S1	48	50	350	2000
S2	48	100	350	2000
S3	48	150	350	2000
S4	48	200	350	2000
S5	48	250	350	2000
S6	48	300	350	2000
S7	48	350	350	2000
S8	48	400	350	2000
S9	48	450	350	2000
S10	48	500	350	2000
S11	100	50	350	2000
S12	100	100	350	2000
S13	100	150	350	2000
S14	100	200	350	2000
S15	100	250	350	2000
S16	100	300	350	2000
S17	100	350	350	2000
S18	100	400	350	2000
S19	100	450	350	2000
S20	100	500	350	2000
S21	148	50	350	2000
S22	148	100	350	2000
S23	148	150	350	2000
S24	148	200	350	2000
S25	148	250	350	2000
S26	148	300	350	2000
S27	148	350	350	2000
S28	148	400	350	2000
S29	148	450	350	2000
S30	148	500	350	2000
S31	200	50	350	2000
S32	200	100	350	2000
S33	200	150	350	2000
S34	200	200	350	2000
S35	200	250	350	2000
S36	200	300	350	2000
S37	200	350	350	2000
S38	200	400	350	2000
S39	200	450	350	2000
S40	200	500	350	2000
S41	248	50	350	500
S42	248	100	350	500
S43	248	150	350	500
S44	248	200	350	500
S45	248	250	350	500
S46	248	300	350	500
S47	248	350	350	500
S48	248	400	350	500
S49	248	450	350	500
S50	248	500	350	500
S51	300	50	350	500
S52	300	100	350	500
S53	300	150	350	500
S54	300	200	350	500
S55	300	250	350	500
S56	300	300	350	500
S57	300	350	350	500
S58	300	400	350	500
S59	300	450	350	500
S60	300	500	350	500
S61	348	50	350	500
S62	348	100	350	500
S63	348	150	350	500
S64	348	200	350	500
S65	348	250	350	500
S66	348	300	350	500
S67	348	350	350	500
S68	348	400	350	500
S69	348	450	350	500
S70	348	500	350	500
S71	400	50	350	500
S72	400	100	350	500
S73	400	150	350	500
S74	400	200	350	500
S75	400	250	350	500
S76	400	300	350	500
S77	400	350	350	500

	Number of municipalities μ	Length of the Square (Km) l	Number of Weights	Number of Ordinations
S78	400	400	350	500
S79	400	450	350	500
S80	400	500	350	500
S81	448	50	350	500
S82	448	100	350	500
S83	448	150	350	500
S84	448	200	350	500
S85	448	250	350	500
S86	448	300	350	500
S87	448	350	350	500
S88	448	400	350	500
S89	448	450	350	500
S90	500	500	350	500
S91	500	50	350	500
S92	500	100	350	500
S93	500	150	350	500
S94	500	200	350	500
S95	500	250	350	500
S96	500	300	350	500
S97	500	350	350	500
S98	500	400	350	500
S99	500	450	350	500
S100	500	500	350	500

Source: Authors' own work.

Notes:

This table provides information on the number of spatial separations calculated for analysing the maximum values attained. The authors have set these parameters.

Each province is represented by a square with side l (in Km) and its number of municipalities (#MUN). For example S1 is a province with 48 municipalities and a square shape of side l = 50 Km.

The municipalities are equidistantly distributed on the edge of the square.

For each province we have simulated 350 sets of weights (#Weights) representing the municipalities' population shares.

Each ordination of the municipalities with their corresponding population has associated a spatial separation value, measured by the Spatial Separation Index that we name $V_{edge}(\omega,\sigma)$.

The differences in the number of ordinations are due to processing times.

The maximum relative error for the estimates of $\underset{\sigma}{\text{Average}}(V_{edge}(l,\ \mu,\ \omega_s,\sigma))$ with a sample size of 500 is 0.0358%. As for a sample size of 2000 it is 0.0179%.

Annex V. Table 2

Summary of the results for spatial separation in the simulated provinces S1 to S100. SIMULATIONS

Each cell of the tables reflects the variable indicated in the upper left corner for a given Si. Si is identified according to the parameters in the row and column headers, as describes in Annex V. Table 1

$\underset{\omega}{\text{Max}}(\underset{\sigma}{\text{Average}}(V_{edge}(l,\mu,\omega_s,\sigma)))$	μ									
	48	**100**	**148**	**200**	**248**	**300**	**348**	**400**	**448**	**500**
l =50	17.67	18.38	18.21	18.38	18.24	18.25	18.23	18.24	18.22	18.36
l =100	35.33	36.77	36.32	36.76	36.23	36.34	36.56	36.76	36.64	36.74
l =150	53.00	55.15	54.40	55.14	54.60	55.13	54.73	55.06	55.02	54.92
l =200	70.67	73.53	72.49	73.51	73.38	72.95	72.70	73.51	73.05	73.31
l =250	88.33	91.92	90.31	91.89	91.16	91.23	91.65	91.63	91.80	91.89
l =300	106.00	110.30	108.81	110.27	108.86	110.27	109.04	110.27	109.72	109.90
l =350	123.67	128.68	126.76	128.98	126.93	128.60	127.91	127.95	128.03	128.27
l =400	141.33	147.07	144.85	147.03	146.76	146.93	145.82	147.02	146.84	146.51
l =450	159.00	165.45	163.21	165.41	164.08	165.40	165.32	166.13	164.66	164.84
l =500	176.67	183.83	178.99	183.79	182.31	182.46	182.59	183.78	182.69	183.77

VI_{max}	μ									
	48	**100**	**148**	**200**	**248**	**300**	**348**	**400**	**448**	**500**
l =50	17.67	18.38	18.21	18.38	18.24	18.25	18.23	18.24	18.22	18.36
l =100	35.33	36.77	36.32	36.76	36.23	36.34	36.56	36.76	36.64	36.74
l =150	53.00	55.15	54.40	55.14	54.60	55.13	54.73	55.05	55.02	54.92
l =200	70.67	73.53	72.49	73.51	73.38	72.95	72.70	73.51	73.05	73.31
l =250	88.33	91.92	90.31	91.89	91.16	91.23	91.65	91.63	91.80	91.89
l =300	106.00	110.30	108.81	110.27	108.86	110.27	109.04	110.27	109.72	109.90
l =350	123.67	128.68	126.76	128.65	126.93	128.60	127.91	127.95	128.03	128.27
l =400	141.33	147.07	144.85	147.03	146.76	146.93	145.82	147.02	146.84	146.51
l =450	159.00	165.45	163.21	165.41	164.08	165.40	165.32	166.13	164.66	164.84
l =500	176.67	183.83	178.99	183.79	182.31	182.46	182.59	183.78	182.69	183.77

$\underset{\omega_{301}-\omega_{350}}{\text{Maximun}}(\underset{\sigma}{\text{Average}}(V_{edge}(l,\mu,\omega_s,\sigma)))$	μ									
	48	**100**	**148**	**200**	**248**	**300**	**348**	**400**	**448**	**500**
l =50	17.25	17.83	17.61	17.76	17.57	17.65	17.57	17.56	17.53	17.71
l =100	34.53	35.64	35.19	35.51	34.87	35.08	35.32	35.44	35.29	35.46
l =150	51.73	53.43	52.64	53.32	52.63	53.29	52.82	53.26	53.01	52.89
l =200	69.05	71.34	70.14	70.99	70.83	70.42	70.24	70.95	70.57	70.76
l =250	86.35	89.14	87.27	88.83	87.92	88.08	88.45	88.39	88.59	88.36
l =300	103.49	106.93	105.24	106.69	104.85	106.55	104.91	106.12	105.76	106.23
l =350	120.72	124.82	122.58	127.98	122.34	123.94	123.58	123.51	123.92	123.72
l =400	138.04	142.54	140.04	142.15	141.75	142.09	140.40	141.92	141.60	141.38
l =450	155.30	160.47	157.89	159.83	158.64	159.57	159.34	160.28	158.29	159.17
l =500	172.63	178.27	173.16	177.63	175.88	176.31	176.41	177.57	176.07	177.61

Source: Authors' own work.

Annex V Table 3

Summary of the results for REAL LIFE Spatial Separation Index between municipalities of Spain's provinces in 2016

It continues

PROVINCES	V Spatial Separation Index (Actual) (a)	ON THE EDGE: V_{edge} and VI_{max}						Ratio MAX4000/ VImax (***)	Ratio IU/VImax (***)
		INE 's order starting at the left upper corner (b)	Average of the 4000 permutations (c) (*)	Maximum of the 4000 permutations (d) (*)	I_L(**) (e)	I_U(**) (f)	VI_{max} (g)		
TOTAL	**17.70**	**50.66**	**43.75**	**50.76**	**35.93**	**51.93**	**51.03**	*1.01*	*0.98*
Almería	20.77	41.78	42.41	49.99	34.29	51.39	48.09	1.04	1.07
Cádiz	23.92	38.79	38.94	43.62	34.31	43.58	41.14	1.06	1.06
Córdoba	21.75	51.11	45.02	53.40	33.32	56.53	55.63	0.96	1.02
Granada	18.39	58.81	55.63	61.69	50.48	62.66	61.11	1.01	1.03
Huelva	18.32	48.18	45.05	51.74	35.18	55.62	51.16	1.01	1.09
Jaén	23.06	51.22	49.54	54.19	44.52	54.47	51.87	1.04	1.05
Málaga	17.32	45.89	40.77	48.86	33.60	51.13	49.13	0.99	1.04
Sevilla	15.90	49.27	47.79	55.18	38.83	59.85	57.90	0.95	1.03
Huesca	27.33	57.52	52.42	58.53	44.19	59.80	56.35	1.04	1.06
Teruel	32.32	53.22	58.42	65.58	49.17	65.75	62.78	1.04	1.05
Zaragoza	12.43	42.32	37.54	46.16	29.42	47.08	75.95	**0.61**	**0.62**
Asturias	16.13	44.89	46.91	58.23	32.08	60.52	52.39	**1.11**	**1.16**
Balears	33.56	73.50	64.14	74.70	54.11	76.73	75.19	0.99	1.02
Palmas	38.69	70.16	62.78	73.56	54.88	74.80	73.65	1.00	1.02
SC Tenerife	24.32	58.63	52.43	61.91	39.82	66.33	59.04	1.05	1.12
Cantabria	13.59	38.44	35.01	41.85	29.56	42.45	40.09	1.04	1.06
Ávila	20.30	37.81	40.99	46.48	33.83	45.42	45.79	1.02	0.99
Burgos	22.93	47.16	46.95	58.96	32.11	59.23	62.35	0.95	0.95
León	24.48	59.24	52.32	61.21	41.38	63.16	58.36	1.05	1.08
Palencia	18.48	33.08	33.35	40.65	24.06	41.29	43.22	0.94	0.96
Salamanca	18.57	40.73	38.92	46.93	32.63	47.10	49.35	0.95	0.95
Segovia	15.88	40.48	36.23	43.03	31.34	42.80	42.24	1.02	1.01
Soria	18.20	42.62	40.93	49.17	30.81	47.17	48.67	1.01	0.97
Valladolid	9.12	35.55	28.30	35.78	23.03	40.01	47.92	**0.75**	**0.83**
Zamora	20.86	44.17	44.08	50.89	36.07	50.31	49.87	1.02	1.01
Albacete	24.07	47.07	47.38	58.16	35.36	56.66	57.73	1.01	0.98
Ciudad Real	29.10	62.56	59.76	66.18	52.38	66.58	62.19	1.06	1.07
Cuenca	30.86	61.15	56.42	62.61	48.06	61.97	60.23	1.04	1.03
Guadalajara	13.72	44.52	45.44	54.95	35.79	57.35	54.28	1.01	1.06
Toledo	28.81	56.95	57.19	63.17	51.32	63.69	59.47	1.06	1.07
Barcelona	12.76	42.61	39.37	44.92	33.91	46.10	44.43	1.01	1.04
Girona	18.47	41.90	39.37	42.95	34.98	43.74	41.04	1.05	1.07
Lleida	21.74	50.35	49.07	55.01	40.85	56.81	55.17	1.00	1.03
Tarragona	19.66	41.49	40.85	46.10	35.24	47.02	43.91	1.05	1.07
Alicante	21.20	40.78	38.32	43.78	32.39	44.54	41.29	1.06	1.08
Castellón	15.33	35.17	36.87	43.89	28.88	45.47	42.20	1.04	1.08
Valencia	14.84	51.05	47.48	53.33	41.13	55.51	54.18	0.98	1.02
Badajoz	32.45	71.76	68.10	75.32	59.59	75.70	71.93	1.05	1.05
Cáceres	31.02	70.00	64.52	71.89	53.82	73.44	69.07	1.04	1.06
Coruña	22.74	49.19	47.04	52.13	42.06	52.15	50.05	1.04	1.04
Lugo	24.34	43.12	41.02	46.31	30.70	49.41	45.26	1.02	1.09
Ourense	17.27	34.28	35.02	39.70	29.27	39.69	39.01	1.02	1.02
Pontevedra	13.64	30.34	31.60	36.06	27.32	36.07	35.24	1.02	1.02

Source: Authors' own work.

Annex V Table 3

Summary of the results for REAL LIFE Spatial Separation Index between municipalities of Spain's provinces in 2016

Conclusion

PROVINCES	V Spatial Separation Index (Actual) (a)	ON THE EDGE: V_{edge} and VI_{max} INE 's order starting at the left upper corner (b)	Average of the 4000 permutations (c) (*)	Maximum of the 4000 permutations (d) (*)	I_L(**) (e)	I_U(**) (f)	VI_{max} (g)	Ratio MAX4000/ VImax (***)	Ratio IU/VImax (***)
Madrid	9.26	41.20	34.53	42.92	27.76	46.24	49.54	**0.87**	**0.93**
Murcia	20.74	52.67	45.73	55.08	32.02	60.03	53.81	1.02	1.12
Navarra	19.62	43.14	44.22	50.43	39.38	50.48	49.97	1.01	1.01
Álava	6.78	13.46	12.07	16.71	6.17	18.33	29.32	**0.57**	**0.63**
Bizkaia	7.46	23.71	21.58	25.89	16.82	26.73	24.73	1.05	1.08
Gipuzkoa	11.95	22.22	20.53	23.16	17.80	23.37	22.37	1.04	1.04
La Rioja	13.91	29.10	27.11	35.01	20.40	34.31	35.87	0.98	0.96

Source: Authors' own work.

(*) Total values calculated as a simple average of provincial values. As for real life data on Spain's provinces, we have used 4000 ordinations of their municipalities population weights.

(**) I_L and I_U are the lower and upper limits of the confidence interval at 99.9% for V_{edge}.

(***)

Ratio << 1 VI_{max} is very high compared to the maximum of the 4000 permutations (MAX4000) or I_U

Álava
Zaragoza
Valladolid
Madrid

We have checked with the general simulations that VI_{max} is as expected for a Province of that size and μ.

Thus, this is because MAX4000 and I_U for that province are actually very low.

Also, in the general simulations, the V_{edge} for the stylised weights based on these provinces are low and the ratio V_{edge} /VI_{max} low.

Therefore, this occurs of the specific distribution of the population among the municipalities of those provinces, which produces low spatial separation indices.

We understand that in these cases VI_{max} is a proper estimate for the maximum attainable spatial separation.

Ratio >>1 VI_{max} is very low compared to MAX4000 or I_U

Asturias

We have checked with the general simulations that VI_{max} is as expected for a Province of that size and #MUN

Thus, this is because MAX4000 and I_U for that Province are very high

Also, in the general simulations, the V_{edge} for the stylised weights based on these provinces are high and the ratio V_{edge} /VI_{max} high.

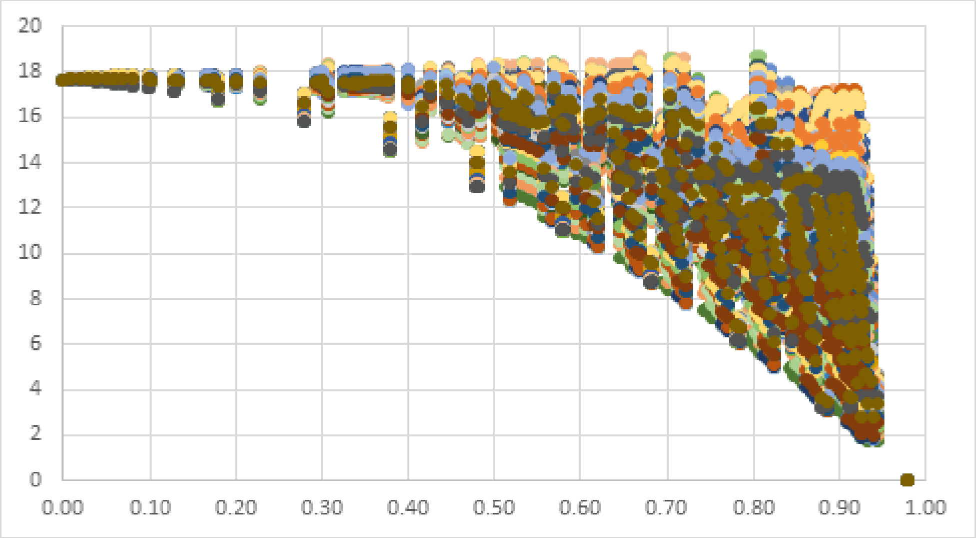
Therefore, this occurs because of the specific distribution of the population among the municipalities of those provinces, which produces high spatial separation indices.

In this case, VI_{max} could not be a proper upper bound for the maximum spatial separation attainable and using it might underestimate proximity.

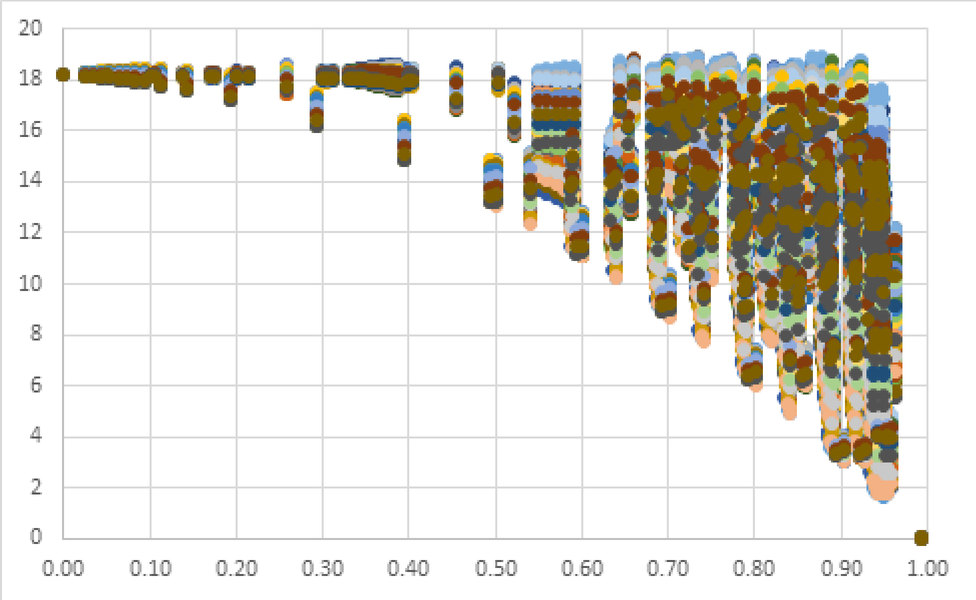
Therefore, we have corrected VI_{max} for Asturias with the factor: MAX4000/VI_{max} = 1.11.

Annex V Chart 1

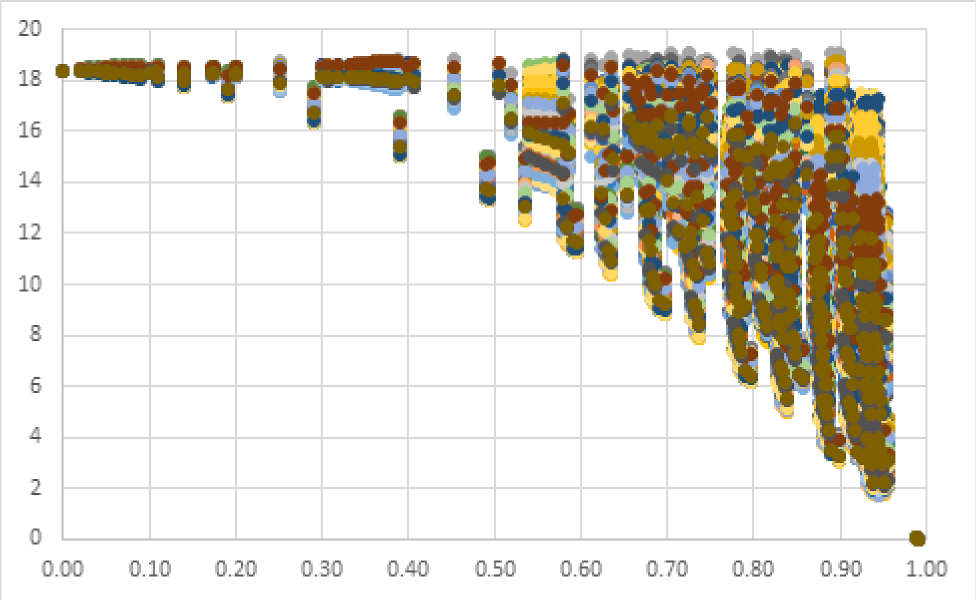
Scatter plot for $(GI\omega_s; V_{edge}(\omega_s, \sigma))$ in S1 (l=50Km; μ=48; ω=1-350; σ=1-2000)



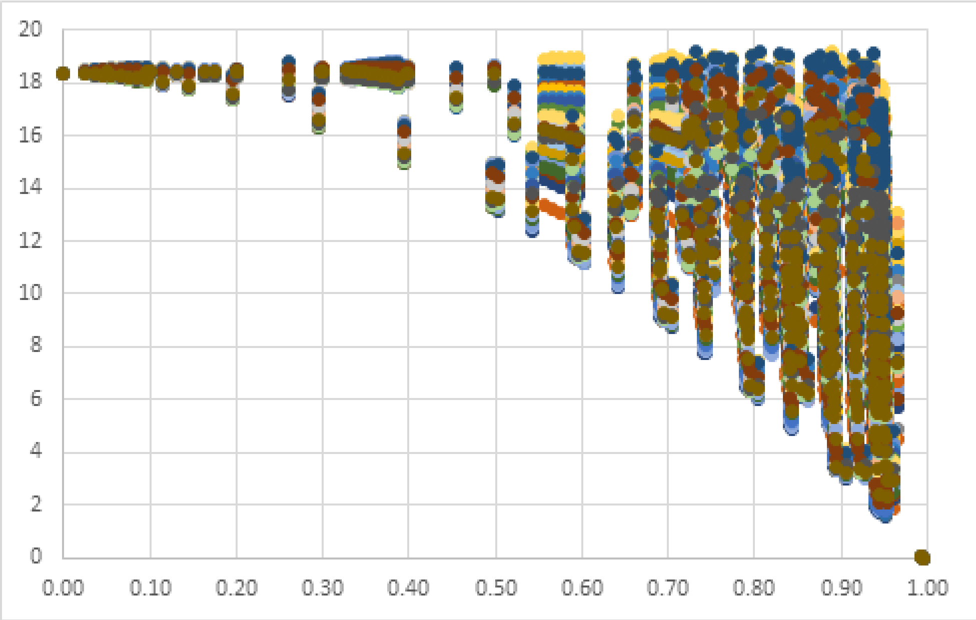
Scatter plot for $(GI\omega_s; V_{edge}(\omega_s, \sigma))$ in S21 (l=50Km; μ =148; ω=1-350; σ=1-2000)



Scatter plot for $(GI\omega_s; V_{edge}(\omega_s, \sigma))$ in S11 (l=50Km; μ=100; ω=1-350; σ=1-2000)

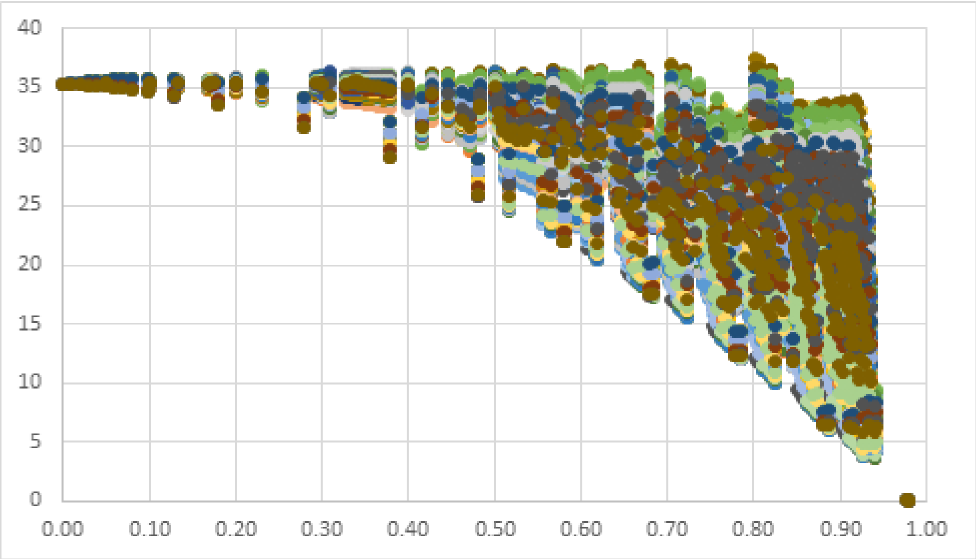


Scatter plot for $(GI\omega_s; V_{edge}(\omega_s, \sigma))$ in S31 (l=50Km; μ =200; ω=1-350; σ=1-2000)

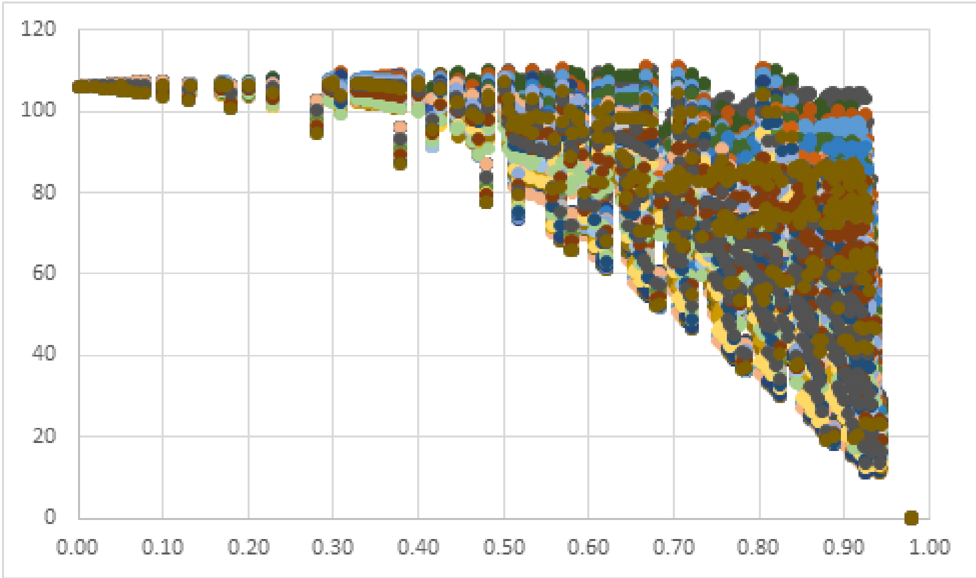


Source: authors' own work

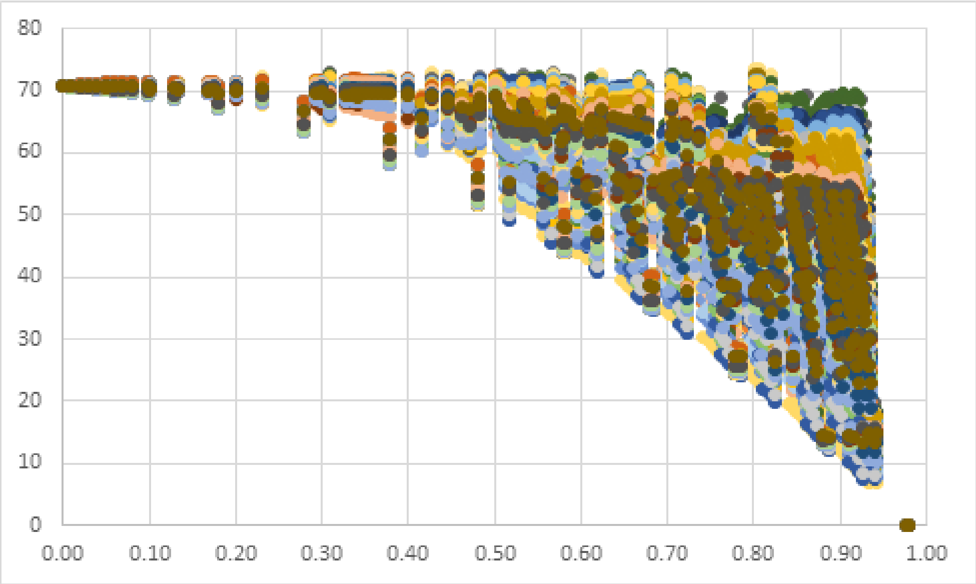
Scatter plot for $(GI\omega_s;\ V_{edge}(\omega_s,\sigma))$ in S2 (l=100Km; μ=48; ω=1-350; σ=1-2000)



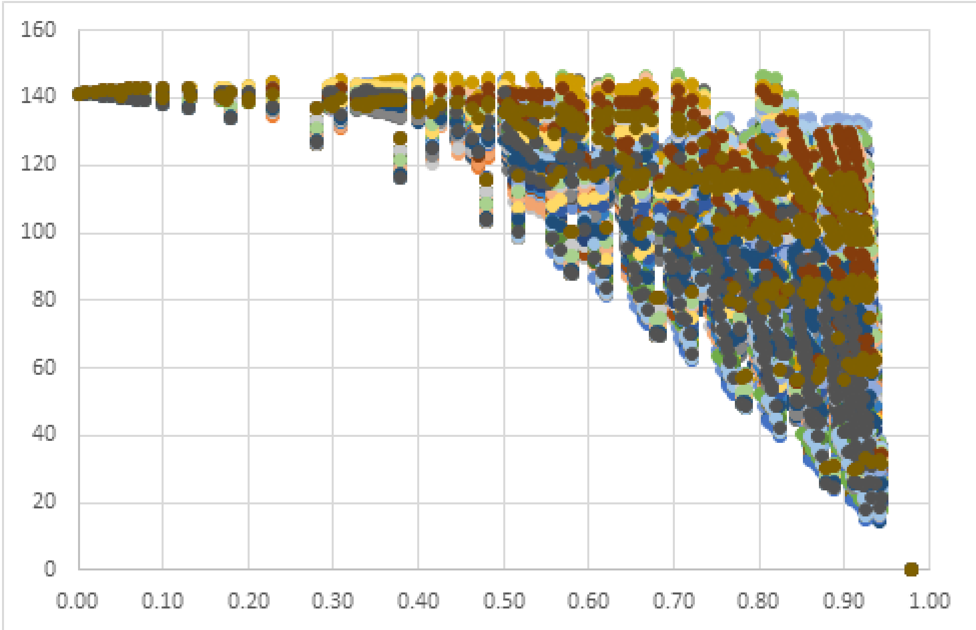
Scatter plot for $(GI\omega_s;\ V_{edge}(\omega_s,\sigma))$ in S6 (l=300Km; μ=48; ω=1-350; σ=1-2000)



Scatter plot for $(GI\omega_s;\ V_{edge}(\omega_s,\sigma))$ in S4 (l=200Km; μ=48; ω=1-350; σ=1-2000)



Scatter plot for $(GI\omega_s;\ V_{edge}(\omega_s,\sigma))$ in S8 (l=400Km; μ=48; ω=1-350; σ=1-2000)



Source: authors' own work

BOE (1985). Ley 7/1985, de 2 de abril, Reguladora de las Bases del Régimen Local. Retrieved from: https://www.boe.es/buscar/pdf/1985/BOE-A-1985-5392-consolidado.pdf

BOE (2015). Resolución de 16 de marzo de 2015, de la Subsecretaría, por la que se publica la Resolución de 30 de enero de 2015, de la Presidencia del Instituto Nacional de Estadística y de la Dirección General de Coordinación de Competencias con las Comunidades Autónomas y las Entidades Locales, sobre instrucciones técnicas a los Ayuntamientos sobre gestión del padrón municipal. Retrieved from: BOE-A-2018-12265. Retrieved from:
https://www.boe.es/buscar/act.php?id=BOE-A-2018-12265&p=20181012&tn=2

BOE (2018). Real Decreto 1113/2018, de 7 de septiembre, por el que se desarrolla la estructura orgánica básica del Ministerio de Hacienda. Ministerio de Política Territorial y Función Pública «BOE» núm. 218, de 8 de septiembre de 2018. Referencia: BOE-A-2018-12265. Retrieved from: https://www.boe.es/buscar/act.php?id=BOE-A-2018-12265&p=20181012&tn=2

Cadastre. Statistics. Retrieved from: http://www.catastro.meh.es/esp/estadisticas.asp

Cadastre. Statistics. Retrieved from:
http://www.catastro.meh.es/documentos/estadisticas_Metodologia_Catastro_2015.pdf

Cochran, W. G. (1984). Sampling Techniques. John Wiley and Sons, Inc. Mexico.

De la Fuente, A. (2017). Simulaciones del modelo propuesto por la Comisión de Expertos para la revisión del sistema de financiación de las Comunidades Autónomas de régimen común. FEDEA e Instituto de Análisis Económico (CSIC). Retrieved from:
http://documentos.fedea.net/pubs/eee/eee2017-19.pdf

Delgado, M. et al. (2016). La financiación de las comunidades autónomas más allá del sistema de financiación autonómica. Fedea Policy Papers - 2016/22. Retrieved from:
https://documentos.fedea.net/pubs/fpp/2016/10/FPP2016-22.pdf

Delgado, M., Lledó, V., Pérez, J. (2016). On the determinants of fiscal non-compliance: an empirical analysis of Spain's regions. Retrieved from:
https://www.bde.es/f/webbde/SES/Secciones/Publicaciones/PublicacionesSeriadas/DocumentosTrabajo/16/Fich/dt1632e.pdf

European Committee of the Regions. Division of Powers. Retrieved from: https://portal.cor.europa.eu/divisionpowers/Pages/Spain-intro.aspx

EUROSTAT (2020). Statistical regions in the European Union and partner countries — NUTS and statistical regions 2021. Retrieved from: https://ec.europa.eu/eurostat/documents/3859598/10967554/KS-GQ-20-092-EN-N.pdf/9d57ae79-3ee7-3c14-da3e-34726da385cf?t=1591285035000

EUROSTAT (a). Population density. Retrieved from:
https://ec.europa.eu/eurostat/databrowser/view/TPS00003/default/table
https://appsso.eurostat.ec.europa.eu/nui/show.do?dataset=demo_r_d3dens&lang=en

EUROSTAT (b). Land Use and Cover Area frame Survey. Retrieved from:
https://ec.europa.eu/eurostat/web/lucas
https://ec.europa.eu/eurostat/documents/4031688/8503684/KS-01-17-069-EN-N.pdf/91e45d7a-ee8c-47ea-a666-f49600d1ee6c?t=1520237929000

EUROSTAT (c). Glossary: Population grid cell. Retrieved from:
https://ec.europa.eu/eurostat/statistics-explained/index.php/Glossary:Population_grid_cell#:~:text=Short%20definition%3A%20a%20population%20grid,to%20the%20number%20of%20inhabitants.&text=Eurostat%20gives%20preference%20to%20the,overlaid%20across%20the%20EU%20territory

EUROSTAT (d). Population grids. Retrieved from: https://ec.europa.eu/eurostat/statistics-explained/index.php/Population_grids

Goerlich, F.J. *et al.* (Mimeo). Financiación Autonómica: Superficie, dispersión y accesibilidad a los servicios públicos. Ivie.

IGN. Gazetteer of municipalities and population entities. Retrieved from: http://centrodedescargas.cnig.es/CentroDescargas/catalogo.do?Serie=NGMEN#
http://centrodedescargas.cnig.es/CentroDescargas/locale?request_locale=en

INE. Continuous Municipal Register. Definitions. Retrieved from: http://www.ine.es/nomen2/Metodologia.do?L=1

INE. Continuous Municipal Register. Microdata provided by the INE under petition.

INE. Yearbook. Several years.

Ivis, F. (2006). Calculating Geographic Distance: Concepts and Methods. Canadian Institute for Health Information, Toronto, Ontario, Canada. Available at: https://www.lexjansen.com/nesug/nesug06/dm/da15.pdf

Jošić, H. et al. (2019). Measurement of Export Market Concentration for the Largest European Economic Integrations. Business Systems Research | Vol. 10 No. 2 |2019. Retrieved from: https://hrcak.srce.hr/file/327462

Mapanet. Distancia entre dos puntos / coordenadas. Retrieved from: http://www.mapanet.eu/Resources/Script-Distance.htm

Microsoft. Bing Maps Distance Matrix API. Retrieved from: https://www.microsoft.com/en-us/maps/distance-matrix

Ministry of Development of Spain (2018). Áreas urbanas en España 2018. Constitución, Cuarenta años de las ciudades españolas. Retrieved from: https://apps.fomento.gob.es/CVP/handlers/pdfhandler.ashx?idpub=BAW058

Ministry of Development of Spain (2018). Sistema de Información Urbana (SIU). Retrieved from: https://apps.fomento.gob.es/CVP/handlers/pdfhandler.ashx?idpub=BAW055

OECD. Dataset: Built-up area and built-up area change in countries and regions. Retrieved from: https://stats.oecd.org/viewhtml.aspx?datasetcode=BUILT_UP&lang=en

OECD/EU (2020). Cities in the World A new perspective on urbanisation. Retrieved from: https://www.oecd.org/regional/regional-statistics/Cities-in-the-World-Policy-Highlights.pdf

Ooi, B.C. Sacks-Davis and R. Han, J. (1993). Indexing in Spatial Databases. Retrieved from: https://www.comp.nus.edu.sg/~ooibc/spatialsurvey.pdf;
https://scholar.google.com/citations?user=9560QjYAAAAJ&hl=en

Sierra, D. et al. (2008). Midiendo el radio real de la tierra. *Rev. Eureka Enseñ. Divul. Cien., 2008, 5(3), pp. 293-301.* Portal de Revistas Electrónicas editadas por la Universidad de Cádiz. Retrieved from: https://revistas.uca.es/index.php/eureka/article/download/3741/3318

Statistics Canada (2017). Dictionary, Census of Population, 2016. Population centre (POPCTR). Retrieved from: https://www12.statcan.gc.ca/census-recensement/2016/ref/dict/geo049a-eng.cfm