A Spatial Analysis of Recent Fertility Patterns in Spain

Alessandra Carioli[1]

Daniel Devolder[2]

Joaquìn Recaño[3]

Spatial demography embodies and important role in detecting fertility behaviour paths. Indeed, national trends are often misleading as they might misrepresent substantial regional deviations from the main national level. Since the onset of the Second Demographic Transition, fertility in Spain has undergone a substantial reduction of fertility during the last three decades, assessing on lowest low levels. Even though we can fairly assert that Spanish fertility is among the lowest in the world, important geographical heterogeneity has always characterized Spanish fertility patterns of tempo and quantum.

Spain is an ideal candidate to study spatial patterns of heterogeneity as its historical and cultural heterogeneity led to highly diverse and unique fertility patterns, which find corroboration in the works from the Princeton European Fertility Project (leasure, 1963; Livi-Bacci 1968a and 1968b).

Spain is the fifth most populated country in the European Union and its history of strong geographical and cultural diversity makes it an ideal candidate to study spatial patterns of fertility trends.

The objective of this paper is to investigate the variability present in fertility across different geographic areas in Spain since the onset of the Second Demographic Transition. The territorial distinctiveness of Spain allows for a spatial analysis of fertility patterns that have always showed unique features throughout time both in terms of tempo and quantum.

Using data from Spanish *municipios* (LAU2), we define approximately 600 territorial units, *comarcas*, in mainland Spain, thus excluding islands and extraterritorial possessions ensuring spatial contiguity.

The indicators used in the analysis are Total Period Fertility Rate, TFR, TFR grouped by 5 years age groups, Mean Age at Childbirth, MAC, Median Age at Childbirth, variance of MAC. All indicators are classified by birth order and by mother and father's country of birth for the calendar years 1986-2011.

The first part of the analysis addresses issues of global and local spatial autocorrelation of the above mentioned indicators through means of descriptive spatial analysis at provincial level (NUTS3).

In a second phase, we introduce the new grouping of municipalities, the *comarcas,* and further investigate patters of geographical dependency. We then apply a cluster analysis to identify which dimensions of fertility can explain the geographical variance in the TFR.

[1] Netherlands Interdisciplinary Demographic Institute and University of Groningen PRC carioli@nidi.nl

[2] CED, Centro d'Estudis Demogràfics, Universitat Autònoma de Barcelona ddevolder@ced.uab.es

[3] CED, Centro d'Estudis Demogràfics, Universitat Autònoma de Barcelona jrecano@ced.uab.es

1. Data and Methodology

1.1 Data

This paper studies the spatial dependency of fertility indicators across Spanish provinces and comarcas for the years 1986-2011.
Indicators for provinces and comarcas are a personal elaboration using data derived from birth INE and the Centre d'Estudis Demografícs at the Universitat Autónoma de Barcelona.
Indicators for comarcas use municipal vital registers with year based information on live births, while data on female population come from municipal register. Given the discontinuous series of population data during years 1981-1997, for years 1981-1985, 1987-1990 and 1992-1997 cohort interpolation methods were necessary to estimate female population numbers.
TFR has been computed using a basis of women aged 15 to 49 years old for total fertility and by 5 years age groups. Mean Age at Childbearing, MAC, employs the same data as TFR and is obtained considering the grouped age specific fertility rates.
The Share of Births by Foreign Mothers at provincial level, SBFM, and the Share of Non Marital Births at provincial level, SNMB come from INE.
All the rates used in the analysis break down at provincial, NUTS3 and at comarcas level. The provinces classification in 50 territorial units follows INE categories, which mirrors EUROSTAT 2004 NUTS3 subdivision.

Table 1. Annual average population by province (males and females), 2008.

NUTS1	Province Name	Population in 1000	NUTS1	Province Name	Population in 1000	NUTS1	Province Name	Population in 1000
	A Coruña	1,121.50		León	483.9		Alicante	1,861.90
NOROESTE	Lugo	346.7		Salamanca	347.6	**(ES5)**	Castellón	583.5
(ES1)	Ourense	327.9	**(ES4)**	Segovia	161	*ES52*	Valencia	2,496.70
ES11	Pontevedra	940.9		Soria	93	*ES53*	Illes Balears	1,057.50
ES12	Asturias	1,059.00		Valladolid	521.2		Almería	670.8
ES13	Cantabria	573.5	*ES41*	Zamora	194.4		Cádiz	1,205.00
	Álava	307.1		Albacete	394.8	**SUR**	Córdoba	784.8
	Guipúzcoa	691.8		Ciudad Real	514.7	**(ES6)**	Granada	898.1
ES21	Vizcaya	1,138.50		Cuenca	213.8		Huelva	499.5
ES22	Navarra	610.4		Guadalajara	231.1		Jaén	654.9
ES23	La Rioja	313.7	*ES42*	Toledo	645.7		Málaga	1,545.70
NORESTE	Huesca	221.4		Badajoz	673.3	*ES61*	Sevilla	1,839.60
(ES2)	Teruel	145.4	*ES43*	Cáceres	406.3	*ES62*	Murcia	1,427.50
ES24	Zaragoza	938.9		Barcelona	5,344.20	**CANARY ISLANDS**	Las Palmas	1063.1
MADRID ES3	Madrid	6,242.20	**ESTE**	Girona	717.2	**(ES7)**	Santa Cruz de Tenerife	995.9
CENTRO	Ávila	169.2	**(ES5)**	Lleida	423.7			
(ES4)	Burgos	365	*ES51*	Tarragona	779			

Source: EUROSTAT

Spain is a large populated country with around 46 million inhabitants and NUTS3 subdivision includes 50[4] provinces which differ quite remarkably both in size and total population.
Table1 depicts the 50 Spanish provinces by population size in year 2008, divided by NUTS1 macro areas and NUTS2, in italics.
Table 2 reports a small summary of descriptive statistics of the variables considered. It interesting to notice the difference in the values reported in the last two columns, which report the minimum and maximum values. All variables have a remarkable difference between min and max values, even though variables standard deviation is rather small (except for MAC and TFR).

[4] With the exception of the extraterritorial possessions of Ceuta and Melilla.

Table 2. Means and standard deviations for variable levels in 2011, across provinces.

Variable	Mean	Standard Deviation	Min	Max
Share of Births by Foreign Mothers	20.65%	0.0895	6%	36%
Share of Non Marital Births	33.2%	0.0665	19%	58%
Mean age at Childbearing	30.83	0.6782	29.31	32.26
Spanish	31.56			
Foreign	28.08			
Total Fertility Rate	1.46	0.2638	0.95	2.42
Spanish	1.38			
Foreign	1.81			

Source: Vital Statistics (MNP) 1996-2008, INE.

2.3 Method

The main question in this study is whether controlling for fertility characteristics reduces or removes the strong spatial patterns of TFR Map.1. Maps are a useful tool to visualize spatial patterns. Map.1 shows how TFR has a clear spatial pattern of lowest-low fertility in the North-Eastern provinces (Galicia, Asturias and Cantabria), while the Southern and Eastern areas show higher fertility levels.

Adjacency between regions can be defined in many ways. In this paper, First Order Queen adjacency is used in order to define neighbouring relations between Spanish provinces and comarcas, so that spatial units are considered neighbours if they share common borders or vertices.

Once the spatial neighbour list has been defined, it is necessary to set the weight matrix for each relationship. The spatial weight matrix has been constructed so that the weights for each areal item sum up to unity.

$$W_{ij} = \begin{bmatrix} 0 & d_{12} & d_{13} \\ d_{21} & 0 & d_{23} \\ d_{31} & d_{32} & 0 \end{bmatrix}$$

A first exploratory measure to evaluate the strength of spatial patterns across the considered variables is Moran's I test (Moran, 1950). In order to measure spatial autocorrelation, Moran's I index is required and is computed on the model's residuals.

Moran's I is the index obtained through the product of the variable considered, let's call it y, and its spatial lag, with the cross product of y and adjusted for the spatial weights considered:

$$\mathrm{I} = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\,(y_i - \overline{y})(y_j - \overline{y})}{\sum_{i=1}^{n}(\mathrm{y_i} - \overline{\mathrm{y}})^2} \quad (1)$$

where n is the number of spatial units i and j, y_i is the i^{th} spatial unit, $\overline{y}$ is the mean of y, and w_{ij} is the spatial weight matrix, where j represents the regions adjacent to i. Moran's I can take on values beween -1 and 1, where – 1 represents strong negative autocorrelation, 0 no spatial autocorrelation and 1, strong positive spatial autocorrelation.

Moran's I test for spatial autocorrelation is a global measure of spatial autocorrelation, meaning that it is computed from the local relationships between the values observed for the geographical unit and its neighbours. It is possible to break down this measure its components in order to identify *clusters* and *hotspots*. Clusters are defined as observations with similar neighbours, while hotspots are observations with very different neighbours (Anselin, 1995). The procedure is knows as *Local Indicators of Spatial Association* or LISA, where the Local Moran's I decomposes Moran's I into its contributions for each location. These indicators detect clusters of either similar or dissimilar values around a given observation. The relationship between global and local indicators is quite simple, as the sum of LISAs for all observations is proportional to Moran's I. Therefore, LISAs can be interpreted both as indicators of local spatial clusters or as pinpointing outliers in global spatial patterns.

The measure for LISAs is defined as:

$$I_i = \frac{(y_i - \bar{y}) \sum_{j=1}^{n} w_{ij}(y_i - \bar{y})}{\frac{\sum_i^n (y_i - \bar{y})}{n}} \quad (2)$$

Where $\bar{y}$, the global mean, is assumed to be an adequate representation of the variable of interest y.

Moran's I and Lisa plots are used to describe both geographical areas classifications, provinces and comarcas. A further step in this study is to apply cluster analysis to identify relevant patterns of fertility.

2. Preliminary Results

The main hypothesis of this study draws from Tobler's first geography rule (Tobler, 1970): "Everything is related to everything else, but near things are more related than distant things". Spatial patterns of heterogeneity in fertility are present in Spain and areas belonging to the same region and sharing the same socio-cultural background (i.e. language) are more likely to show similar patterns of fertility. The period studied, 1986-2011, is very important as substantial transformations of fertility behaviour took place during those years. For instance, Spain started a sharp decrease in fertility, which reached under-replacement levels (TFR<1.3) during the mid-1990s. Also, migration started becoming an important phenomenon very relevant to determine and understand fertility trends during the mid-1990s.

Preliminary analysis of provincial fertility indicators shows clear geographical patterns supporting the aforementioned hypothesis.

Maps 1 and 2 show the geographical distribution of some of the variables considered in this study. TFR in Spain is considered to be one of the lowest in Europe and in the World, having dropped well below 1.3 children per woman during the 1990s (Kohler, Billari, and Ortega, 2002 and 2006), that is to say well below replacement level[5]. The North-Western provinces register the lowest levels of fertility, while some provinces in the South and in the North-East show slightly higher levels of TFR.

Map 2 depicts how much of the total number of children born in each province can be attributed to foreign women. SBFM ranges between 6% and 36%, over a national average of around 20%. There is a clear East-West pattern, with the East of Spain detaining the highest share, especially the region of Catalonia (upper East provinces), with the province of Girona having the highest percentage. The South of Spain and Galicia (upper West provinces) show the lowest share, averaging between 6-14% well below the national level.

The SNMB (Map.2) shows where births outside of wedlock are concentrated. In most of Spain SNMB ranges below 30%, with the Southern provinces scoring the lowest levels. On the other hand, the Balearic and Canary islands, Madrid Autonomous Community and Catalonia are the areas where births outside marriage are highest.

In year 2008, MAC (Map.2) in Spain was almost 31 years old. Northern and North-Eastern provinces represent the area where MAC is highest, with a peak in the Basque Countries (North) of 32 years old. The South shows lower values, even though not much lower than the national average, as the youngest age registered is around 29.5 years old.

The first exploratory analysis to test whether there is spatial autocorrelation is reported in table 3. Moran's I test for the variables considered shows a high and statistically significant autocorrelation, so there is a strong spatial pattern, which needs to be taken into consideration.

[5] Replacement level is considered the level of fertility at which a population is able to replace itself from one generation to the next and is fixed at 2.1 children per woman.

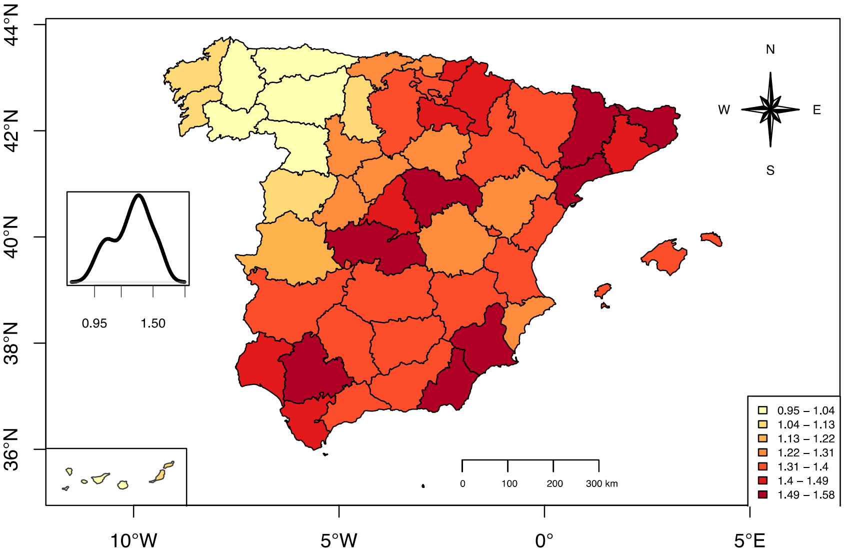
Table3. Moran's I test, year 2011.

	Moran's I statistic	Expectation	Variance	St. Deviation
SBFM	0.5771***	-0.0196	0.0077	6.7962
SNMB	0.5345***	-0.0196	0.0078	6.2631
MAC	0.1831*	-0.0196	0.0068	2.4525
TFR	0.5281***	-0.0196	0.0071	6.4922

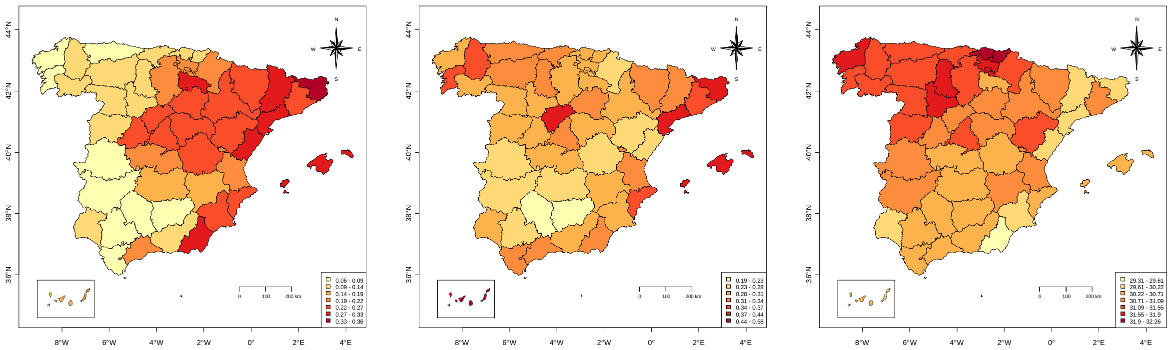
Moran's I statistic is pivotal to assess the general spatial autocorrelation of the variables. In order to measure how covariates behave locally, LISA test is needed. As previously mentioned, the LISA test used in this analysis is based on Moran's I, thus being a proportional to the global indicator of spatial association.
Fig.1 synthesises information about the various variables considered in the scatter plots, where provinces are evidences by colours according to their NUTS1 (macro-areas) classification. The maps depicting significant clusters (maps 3 to 6) illustrate the geographic distribution of four types of value combinations for the 50 Spanish provinces. Red provinces represent *hot-spots*, areas with values above the mean of the phenomenon studied, which also share similar features with their neighbours. On the other hand, blue provinces, *cold-spots*, have below the average values of the studied phenomenon and share no similar characteristics with their neighbours. A cluster appearing among the selected indicators is the region of Galicia (North-West) a province with historically low fertility. Indeed, it is a cold-spot for all variables considered, which is also statistically significant. On the other hand, Catalonia (East) presents features of higher fertility.
This results need to be further investigated as the small number of the provinces of Spain, and thus the high variance of variables across provinces, does not allow for a thorough analysis. The use of comarcas, around 600 geographical units will allow for a better understanding of fertility geographical patterns.

Maps and Graphs:

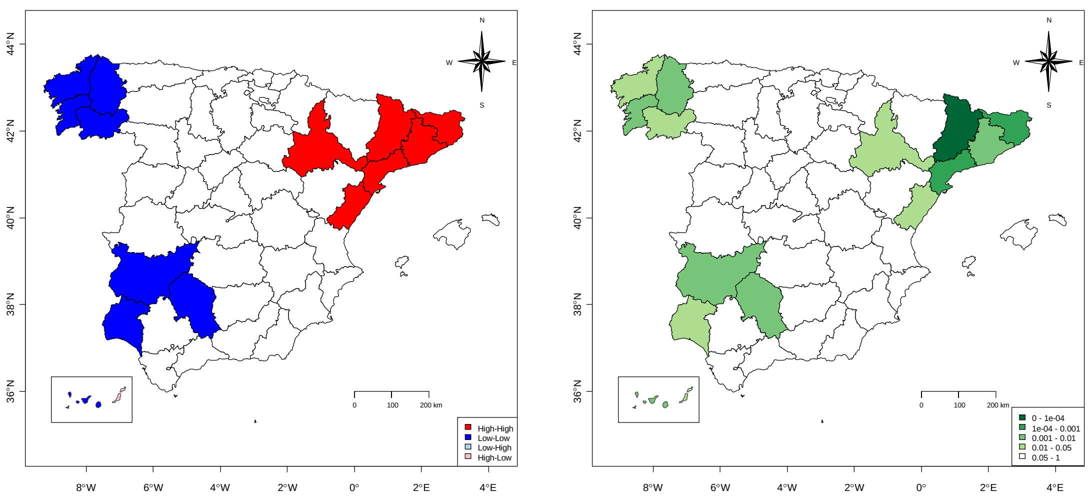
Map1. Total Fertility Rate across provinces in year 2011.



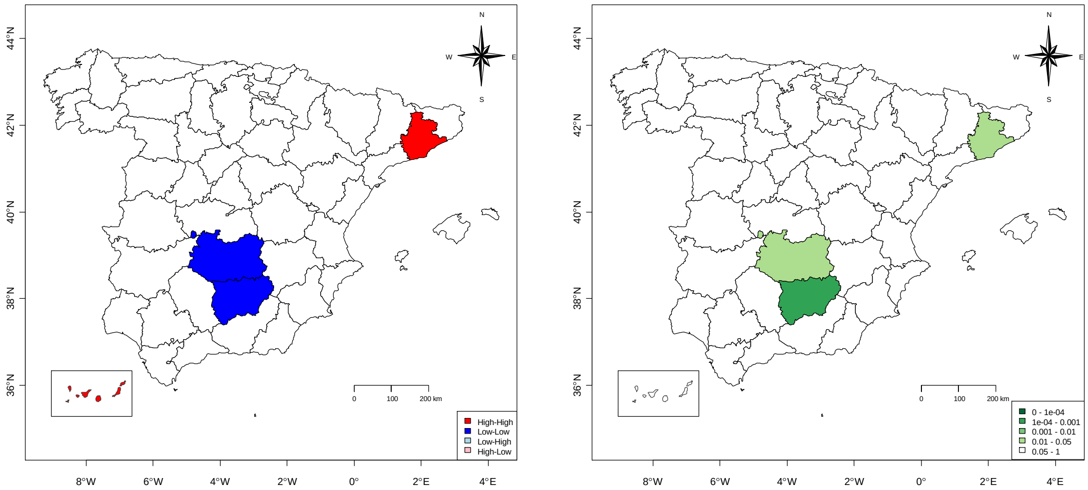
Map2. Share of Births by Foreign Mothers (left) and Share of non-Marital Births (centre) and Mean Age at Childbearing, Spain 2011.



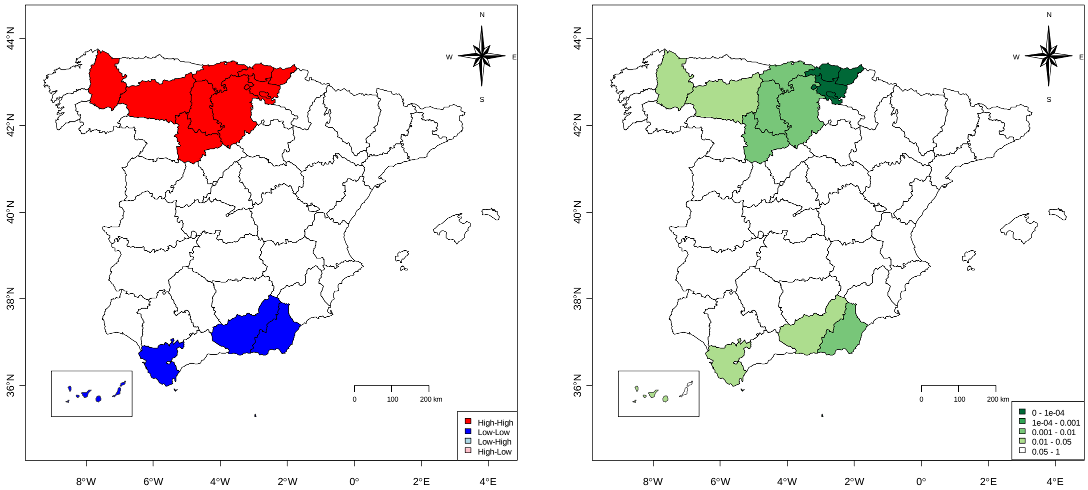
Map3. Significant clusters for SBFM and relative significance plot, 2011.



Map4. Significant clusters for SNMB and relative significance plot, 2011.



Map5. Significant clusters for MAC and relative significance plot, 2011.



Map6. Significant clusters for TFR and relative significance plot, 2011.

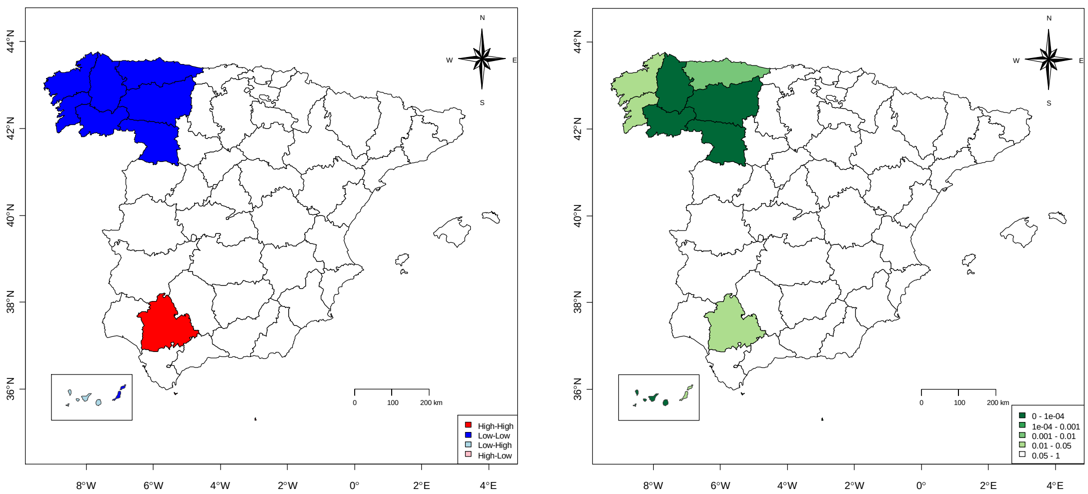
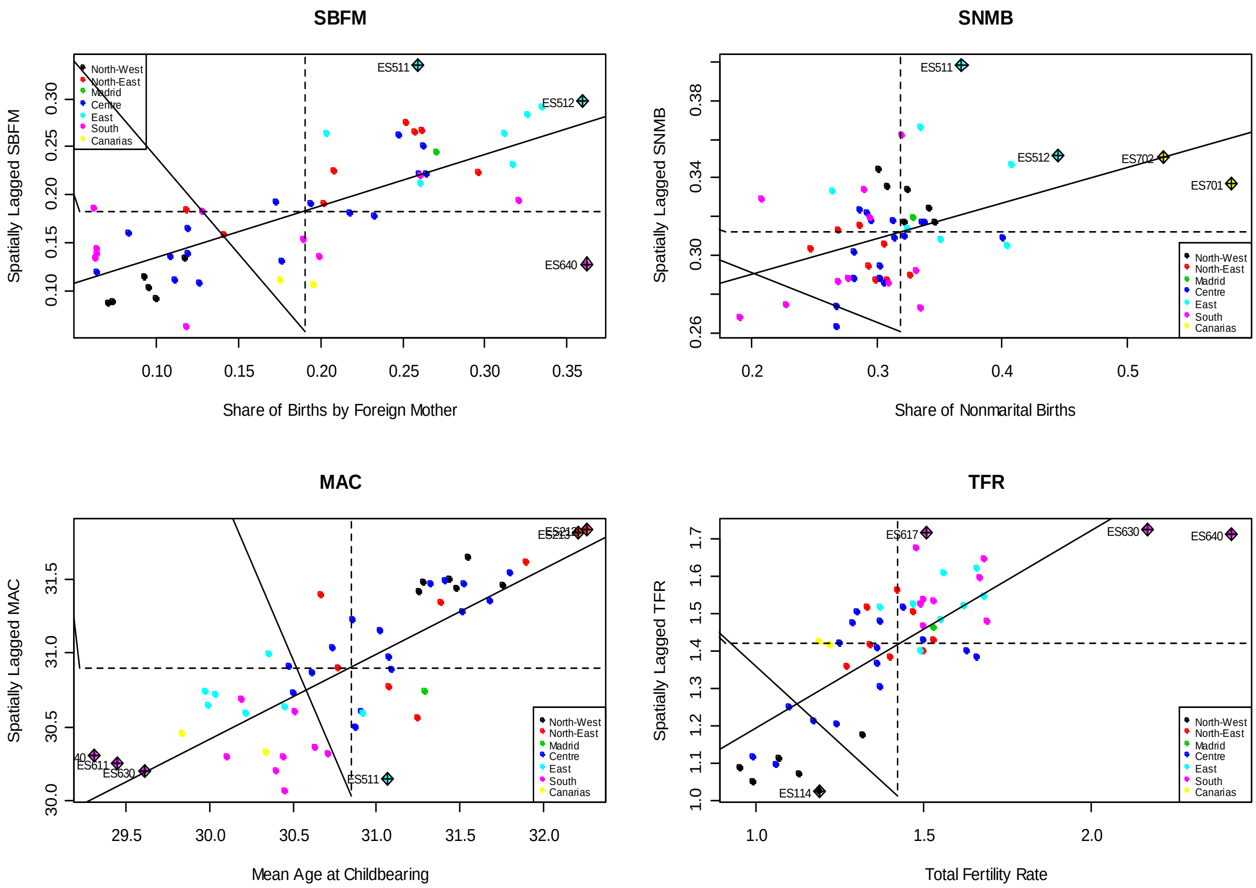


Fig.1 Local Moran's I scatterplots. Provinces evidenced with regard to their NUTS1 matching.



References

Anselin, L. (1995). Local indicators of spatial association-LISA. Geographical Analysis, 27: 93-115.

Leasure, W. J. (1963). Factors involved in the decline of fertility in Spain 1900-1950. Population Studies 16(3): 271-285.

Livi-Bacci, M. (1968). Fertility and Nuptiality Changes in Spain from the Late 18th to the Early 20th Century: Part1. Population Studies, 22(1): 83-102.

Livi-Bacci, M. (1968). Fertility and Nuptiality Changes in Spain from the Late 18th to the Early 20th Century: Part1. Population Studies, 22(2): 211-234.

Kohler, H.-P., F. C. Billari, and J. A. Ortega. 2002. "The emergence of lowest-low fertility in Europe during the 1990s". Population and Development Review 28(4): 641-680.

Kohler, H.-P., F. C. Billari, and J. A. Ortega. 2006. "Low fertility in Europe: Causes, implications and policy options." In F.R. Harris (Ed.) The Baby Bust: Who will do the Work? Who Will Pay the Taxes? Lanham, MD: Rowman & Littlefield Publishers, pp. 48-109.

Moran, P.A.P, 1950. Notes on continuous stochastic phenomena, Biometrika 37:17-23.