UNIVERSITAT ROVIRA I VIRGILI

WORKING PAPERS

Col·lecció "DOCUMENTS DE TREBALL DEL DEPARTAMENT D'ECONOMIA"

"A note on Spatial Autocorrelation at a Local Level"

Josep Maria Arauzo Carod

Document de treball nº -3- 2005

DEPARTAMENT D'ECONOMIA
Facultat de Ciències Econòmiques i Empresarials

UNIVERSITAT
ROVIRA I VIRGILI

Edita:

Departament d'Economia
http://www.fcee.urv.es/departaments/economia/public_html/index.html
Universitat Rovira i Virgili
Facultat de Ciències Econòmiques i Empresarials
Avgda. de la Universitat, 1
432004 Reus
Tel. +34 977 759 811
Fax +34 977 300 661

Dirigir comentaris al Departament d'Economia.

Tots els treballs publicats a la col·lecció *"Documents de Treball del Departament d'Economia"* han superat un procés d'avaluació externa.

Dipòsit Legal: T-1025-2006
ISSN 1576 - 3382

DEPARTAMENT D'ECONOMIA
Facultat de Ciències Econòmiques i Empresarials

A note on Spatial Autocorrelation at a Local Level

Josep-Maria Arauzo-Carod
Grup de Recerca d'Indústria i Territori (GRIT)
Department of Economics (Universitat Rovira i Virgili)
Av. de la Universitat, 1; 43204 – Reus (Catalonia)
Phone: + 0034 977 759 854; Fax: + 0034 977 759 810
E-mail: josepmaria.arauzo@urv.net

Abstract

In this paper we analyze the existence of spatial autocorrelation at a local level in Catalonia using variables such as urbanisation economies, population density, human capital and firm entries. From a static approach, our results show that spatial autocorrelation is weak and diminishes as the distance between municipalities increases. From a dynamic approach, however, spatial autocorrelation increased over the period we analysed. These results are important from a policy point of view, since it is essential to know how economic activities are spatially concentrated or disseminated.

Key words: spatial autocorrelation, municipalities

JEL classification: R110, R120

1. Introduction

In this paper we use Moran's I indicator to analyze spatial autocorrelation at a municipality level in Catalonia. Spatial autocorrelation occurs when the value of a variable at any one point in space depends on values at the neighbouring points. That is, the spatial values of the variable are not random and (from a local level approach) depend on the distances and economic relations between neighbouring municipalities[1]. A lack of spatial autocorrelation, on the other hand, implies that the values of variables are randomly distributed over space, i.e. these values have nothing to do with the position of the municipalities. There are two types of spatial autocorrelation (positive and negative). Positive spatial autocorrelation means that similar values tend to be near each other and negative spatial autocorrelation means that similar values tend to be far from each other.

These kinds of spatial externalities are of great interest and have largely been discussed and empirically verified using spatial econometrics techniques. Recent applications include, for instance, the analysis of spatial dependence at a European regional level (Maza and Villaverde, 2004; Le Gallo et al., 2005), at a Spanish provincial level (Villaverde, 2005), at a Florida counties level (Theil et al.,1996) and at an English municipality level (Revelli, 2002), and highlight the importance of taking spatial aspects into account.

Our aim is not only to test the existence of this phenomenon but also, as Theil et al. (1996) do, to examine the evolution of this spatial dependence over time. Why, though, is spatial autocorrelation important? Perhaps the answer is provided by Waldo Tobler's first law of geography (1970), which states that "everything is related to everything else, but near things are more related than distant things".

From an intuitive point of view, spatial autocorrelation implies that the values of

[1] This implies that the values of a variable in a territory are not only explained by internal conditions, but also by the values of the same variable in neighbouring or nearby regions. However, spatial autocorrelation could also be explained by errors in data collection and data processing, and by the lack of adjustment between the analysed phenomena and the chosen territorial unit.

a variable x in a territory i (x_i) are partially determined by the values of the variable in the neighbouring areas j (x_j, if a $j \neq x$). From an analytical point of view, identifying the j neighbouring areas is a key decision. There are several ways to identify these areas. One involves building a contact matrix in which all the municipalities are considered to be equal in size and shape. This situation, however, is rare because the territorial units (municipalities, counties, regions, etc.) are not equal in size and shape. This implies, for example, that the number of municipalities of which each municipality is a neighbour is never the same. Other measures of spatial contiguity therefore fit better with real data (the distance between municipalities, for instance).

2. Spatial Econometric Analysis

To test the existence of spatial autocorrelation, several indicators exist (Anselin, 1988). One of the oldest and best known is Moran's I indicator (Moran, 1948):

$$Moran's\, I = \frac{\dfrac{\sum_{i \neq j} c_{ij}(x_i - \bar{x})(x_j - \bar{x})}{W}}{\dfrac{\sum_i (x_i - \bar{x})^2}{n}} \qquad \text{where } W = \sum_{i \neq j} c_{ij}$$

The numerator is the covariance between contiguity observations (each contiguity weights c_{ij}/W). This covariance is null if there is no spatial autocorrelation, positive if there is positive spatial autocorrelation and negative if there is negative spatial autocorrelation. The covariance is normalised using the total variance of the series (denominator).

The values of Moran's I are interpreted as follows: if they range from -1 to 0, there is negative spatial autocorrelation; if it is 0, there is a random distribution of the variable; and if they range from 0 to 1, there is positive spatial autocorrelation.

Using this indicator requires previously assuming some null hypothesis about the lack of spatial autocorrelation. To test whether the occurrence of an event in an area follows some kind of systematic spatial pattern, this distribution can be compared with a random pattern distribution[2].

3. Data and variables

We used Moran's I indicator at a local level for Catalan municipalities from 1986 to 2001 for several variables. As it is difficult to identify neighbouring areas only by spatial contiguity[3], we decided, as most scholars do (see, for instance, Le Gallo et al., 2005, and Villaverde, 2005), to use a contiguity matrix based on a distance-based weight[4].

Neighbouring measures always have a certain level of arbitrariness since the researcher must first define the requirements (for a municipality) for being considered as a neighbour of another municipality. Once a distance-based measure is chosen, the second stage is to choose which distance makes two municipalities neighbours. Here we decided on a distance of 30 km, since this is short enough to capture neighbouring effects and long enough to maximise the number of neighbours in each municipality (which allows the use of a sufficient sample of municipalities)[5]. We also used 10 and 20 km distances in order to capture those phenomena of major spatial agglomerations that exist in some of

---

[2] Specifically, the software we used in our calculations (RookCase, version 0.9.6) randomly distributes the values about 20,000 times. In this way all the municipalities have the same probability of receiving one value or another.

[3] Spatial contiguity means that a municipality *j* is adjacent to municipality *i*. Some scholars use a weight matrix in which two areas are neighbours if they have a common border (see Revelli, 2002, for instance).

[4] We made these calculations by considering the cartographical position of each municipality (UTM coordinates) with data obtained from the Catalan Statistical Institute (IDESCAT). These derive from the Universal Transverse Mercator. The central point of the capital of a municipality is determined by the intersection of the *x* and *y* coordinates. The *x* coordinates measure the distance (metres) from this point to a certain meridian of the net (west direction) and the *y* coordinates represent the distance to the equator.

[5] Viladecans (2001), for instance, uses 15 and 30 km. The 30 k distance was selected because it is similar to the size of local labour markets in Catalonia and Valencia. See also Viladecans (2004) for a more detailed analysis.

the *comarques*[6] in our dataset. We therefore consider two municipalities to be neighbours if the distance between them is up to 10, 20 or 30 km.

To test the spatial autocorrelation phenomena we used the following variables obtained from the Catalan Statistical Institute (IDESCAT): *population density*, *urbanisation economies*, *industrial diversity*, *manufacturing jobs*, *service jobs*, *human capital* (*university-level*), *human capital* (*intermediate-level*) and *firm entries*. *Population density* is the total population divided by the area of the municipality, *urbanisation economies* is the number of jobs divided by the area of the municipality, *industrial diversity* is a Hirshmann-Herfindahl index that captures the industrial mix, *manufacturing jobs* is the percentage of jobs in the manufacturing sector, *service jobs* is the percentage of jobs in the services sector, *human capital* (*university-level*) is the percentage of population over 10 years old with a university degree, *human capital* (*intermediate-level*) is the percentage of population over 10 years old with (at least) a high school degree and *firm entries* is the number of entries of manufacturing firms.

4. Results

Our results clearly show that, measured by Moran's I, there is some, though little, positive spatial autocorrelation and that this spatial autocorrelation diminishes when the intermunicipality distance increases[7].

These results mean that the values of variables in each municipality are related to those in neighbouring municipalities and that this relation is closer for the closer municipalities, which implies that polarization is present for Catalan municipalities, as Maza and Villaverde (2004) show at an EU regional level and Villaverde (2005) shows at a Spanish provincial level. For instance, results with the 10 km neighbouring criterion show that manufacturing and service jobs are

[6] The territorial division of Catalonia is based on municipalities and *comarques* (these are made up of municipalities). There are 41 *comarques* in Catalonia with an average area of 781 km^2 and an average population of 145,000. In Catalonia there are 946 municipalities.
[7] Additionally we also performed Geary's C (1954) indicator, which provided very similar results to those obtained by Moran's I.

clustered into small areas. Given this kind of discontinuous agglomerations, this facilitates public policies on manufacturing and services firms.

[INSERT TABLE 1 ABOUT HERE]

However, there are some variables for which this spatial autocorrelation is smaller e.g. *industrial diversity*, *human capital* (*university-level* and *intermediate-level*) and *firm entries*. The values of these variables are less influenced by the values of the same variable in neighbouring municipalities. This more random location pattern needs to be explained specifically for each variable, but in terms of *firm entries*, for instance, it suggests that characteristics such as accessibility to highway or railway networks is a less important determinant of firm location decisions, since accessibility is roughly the same for neighbouring municipalities. Therefore, if firm entries are not clustered in neighbouring municipalities but spread over the territory, decisions regarding supply-side services that could be shared by these entrants should bear in mind this spatial pattern.

Finally, from a dynamic approach the spatial autocorrelation of all the variables has increased during the period analysed (1986 to 2001). This situation is consistent with empirical evidence about firm and population relocation in Catalonia during those years, i.e. the increase of the concentration of manufacturing and services activities in certain areas and some clear trends taken by the municipalities about specialisation in residential or economic activities. We should also insist on major improvements in connectivity among Catalan municipalities in order to explain the rise of similarities in neighbouring municipalities.

5. Conclusion

In this paper we have assumed three criteria with regard to neighbouring: municipalities located 10 km, 20 km and 30 km from another municipality. The differences in the results for each of these neighbouring criteria suggest that the concentration of similar municipalities only exists in small areas, i.e. it is easier to identify smaller clusters rather than larger homogeneous areas. However, we would like to point out that the evidence on spatial autocorrelation is significant but not so strong for some of the variables analysed.

Note that significant and positive values of Moran's I could also be interpreted, from an economic point of view, as the fact that municipalities are perhaps not the most appropriate administrative unit (Viladecans, 2001). This conclusion is clear, for instance, for the more agglomerated activities inside geographical areas (economic regions) that are divided into several administrative units[8].

These results are important from a policy point of view since it is essential to know how economic activities are concentrated or disseminated over the territory. Before designing public policies, public administrations should investigate what the spatial patterns of economic activity are and, according to the characteristics of the territory, thus determine the desired spatial patterns of economic activity. The kind of policies involved in this issue include, for instance, the design of transport infrastructure and the supply of public services.

Despite these conclusions, more work is needed in this area. This is only a preliminary approach into spatial autocorrelation at a local level. Future research should examine the specific determinants of this phenomenon at a more sectoral disaggregated level. It is important to ascertain whether some economic phenomena are clustered at local level since, depending on this spatial relation, public policies differ.

Acknowledgements

This research was partially funded by CICYT: SEJ2004-05860/ECON and CICYT: SEJ2004-07824/ECON. I would like to acknowledge the helpful and supportive comments of Angel Alañón, Miguel C. Manjón and Agustí Segarra. Any errors are, of course, my own.

[8] The metropolitan area of Barcelona is a clear example of this situation.

References

Anselin, L. (1988): *Spatial Econometrics: Methods and Models*, Kluewer Academic Publishers, Dordrecht.

Geary, R. (1954): “The contiguity ratio and statistical mapping”, *The Incorporated Statitian* **5**: 115-145.

Le Gallo, J.; Baumont, C.; Dall’erba, S. and Ertur, C. (2005): “On the property of diffusion in the spatial error model”, *Applied Economics Letters* **12**: 533-536.

Maza, A. and Villaverde, J. (2004): “Regional disparities in the EU: mobility and polarization”, *Applied Economics Letters* **11**: 517-522.

Moran, P. (1948): “The interpretation of statistical maps”, *Journal of the Royal Statistical Society* **59**: 185-193.

Revelli, F. (2002): “Testing the tax mimicking versus expenditure spill-over hypotheses using English data”, *Applied Economics* **14**: 1723-1731.

Theil, H.; Chen, D. and Galvez, J. (1996): “The spatial correlation of the per capita income of Florida’s counties, 1970-1990”, *Applied Economics Letters* **3**: 71-72.

Tobler, W. R. (1970): “A computer movie simulating urban growth in the Detroit region”, *Economic Geography* **46**: 234–40.

Viladecans, E. (2004): "Agglomeration economies and industrial location: city-level evidence", *Journal of Economic Geography* **4 (5)**: 565-582.

Viladecans, E. (2001): "La concentración territorial de las empresas industriales: un estudio sobre el tamaño de las empresas y su proximidad geográfica", *Papeles de Economía Española* **89/90**: 308-319.

Villaverde, J. (2005): “Provincial convergence in Spain: a spatial econometric approach”, *Applied Economics Letters* **12**: 697-700.

Tables

Table 1: Results for Moran's I

	1986			1991			1996			2001		
Variable	10 km	20 km	30 km	10 km	20 km	30 km	10 km	20 km	30 km	10 km	20 km	30 km
Urbanisation economies	0.422***	0.302***	0.192***	0.456***	0.323***	0.203***	0.483***	0.355***	0.230***	0.504***	0.381***	0.252***
	(30.873)	(42.620)	(39.998)	(33.363)	(45.613)	(42.294)	(35.341)	(50.075)	(47.921)	(36.898)	(53.714)	(52.375)
Population density	0.443***	0.308***	0.189***	0.454***	0.322***	0.202***	0.471***	0.342***	0.220***	0.493***	0.367***	0.241***
	(32.468)	(43.550)	(39.428)	(33.239)	(45.522)	(42.056)	(34.482)	(48.231)	(45.675)	(36.023)	(51.856)	(50.188)
Industrial diversity	0.122***	0.080***	0.066***	0.454***	0.322***	0.202***	0.261***	0.186***	0.157***	0.241***	0.186***	0.154***
	(9.016)	(11.411)	(13.797)	(33.239)	(45.522)	(42.026)	(19.165)	(26.271)	(32.718)	(17.648)	(26.322)	(32.016)
Human capital (university-level)	0.113***	0.072***	0.066***	0.124***	0.064***	0.062***	0.194***	0.116***	0.112***	0.216***	0.133***	0.117***
	(8.345)	(10.300)	(13.942)	(9.142)	(9.214)	(13.057)	(14.234)	(16.483)	(23.332)	(15.855)	(18.839)	(24.506)
Manufacturing jobs (%)	0.474***	0.371***	0.328***	0.547***	0.434***	0.376***	0.486***	0.389***	0.338***	0.550***	0.435***	0.366***
	(34.719)	(52.321)	(68.150)	(40.002)	(61.150)	(78.047)	(35.542)	(54.849)	(70.150)	(40.223)	(61.397)	(75.962)
Service jobs (%)	0.211***	0.174***	0.152***	0.298***	0.226***	0.187***	0.359***	0.281***	0.236***	0.407***	0.316***	0.250***
	(15.492)	(24.605)	(31.709)	(21.812)	(31.901)	(38.957)	(26.323)	(39.631)	(49.016)	(29.814)	(44.543)	(51.924)
Human capital (intermediate-level)	0.246***	0.192***	0.144***	0.204***	0.148***	0.116***	0.267***	0.194***	0.154***	0.316***	0.232***	0.172***
	(18.078)	(27.172)	(29.960)	(14.961)	(20.967)	(24.245)	(19.570)	(27.379)	(32.085)	(23.166)	(32.799)	(35.758)
Firm entries	0.192***	0.133***	0.090***	0.219***	0.173***	0.127***	0.291***	0.234***	0.163***	0.197***	0.176***	0.142***
	(14.107)	(18.827)	(18.903)	(16.088)	(24.518)	(26.515)	(21.346)	(33.046)	(34.033)	(14.450)	(24.914)	(29.560)

(***) Significance at 1%, (**) significance at 5% and (*) significance at 10%. z-statistic in brackets.