SPACE-TIME LAGS: SPECIFICATION STRATEGY IN SPATIAL REGRESSION MODELS

by

F.A. López and C. Chasco

REAL 04-T-17 November, 2004

SPACE-TIME LAGS: SPECIFICATION STRATEGY IN SPATIAL REGRESSION MODELS*

LÓPEZ, F.A. (*) and C. CHASCO (**)
(*) *Department of Quantitative and Computer Methods. Universidad Politécnica de Cartagena*. e-mail: Fernando.Lopez@upct.es
(**) *Department of Applied Economics. Universidad Autónoma de Madrid*. e-mail: coro.chasco@uam.es;

ABSTRACT: The purpose of this article is to analyze the dynamic trend of spatial dependence, which is not only contemporary but time-lagged in many socio-economic phenomena. Firstly, we show some of the commonly used exploratory spatial data analysis (ESDA) techniques and we propose other new ones, the exploratory space-time data analysis (ESTDA) that evaluates the instantaneity of spatial dependence. We also propose the space-time correlogram as an instrument for a better specification of spatial lag models, which should include both kind of spatial dependence. Some applications with economic data for Spanish provinces shed some light upon these issues.

Key words: Spatial dependence, spatial diffusion, ESDA, correlogram, Spanish provinces

JEL Classification: **C21, C33, C51, C53, D14, O18**

1. INTRODUCTION

The purpose of this article is to analyze the dynamic trend of spatial dependence, making a differentiation between two types of spatial dependence: contemporary or instant and noncontemporary or lagged. The first typr is the consequence of a very quick diffusion of the process over the neighboring locations, while the second one implies that a shock in a certain location needs several periods of time to diffuse over its neighborhood. It is not easy to separate both types of spatial dependence but they must both be present very frequently when specifying a spatial dependence model.

In effect, spatial dependence has usually been defined as a spatial effect, which is related to the spatial interaction existing between geographic locations and *takes place in a particular moment of time*. In other words, spatial dependence is considered as the contemporary coincidence of value similarity with locational similarity and is formally expressed as a spatial autoregressive model, in which a variable y is a function of its spatial lag Wy (a weighted average of the value

* Earlier versions of this paper were presented at the XVIII Asepelt-Spain Annual Meeting 2004 at University of León (Spain) in June 2004, the Regional Economic Applications Laboratory seminars at the University of Illinois at Urbana-Champaign in September 2004 and the First Seminar in Spatial Econometrics at Zaragoza (Spain) in October 2004. We appreciate comments from participants at these meetings.

of y in the neighboring locations, or spatial lag)[1] for a same moment of time. However, in most socio-economic phenomena this "coincidence in values-locations" (Anselin, 2001a) is not only an instant coincidence but also a final effect of some cause that happened in the past, one which has spread through geographic space during a certain period of time.

In effect, when spatial dependence is produced by the existence of spatial interaction, spatial spillovers or spatial hierarchies in the endogenous variable of a spatial regression model, the spatial-lag or "simultaneous model of spatial dependence" (Anselin, 2001b) has been frequently mentioned in the literature as the solution. Nevertheless, there are some authors that have considered this instantaneity of spatial dependence as problematic (Upton and Fingleton 1985, pp. 369), suggesting the introduction of a time-lag in the spatial dependence term, i.e. $y_t=\rho W y_{t-k}$; $k=1,2,\ldots$ In socioeconomic contexts, there is no doubt that a shock produced in a certain location (e.g. an income growth) will be probably diffused over its neighboring locations over a time interval. Recently, Elhorst (2001) considered this topic and presents several single equation models that include a wide range of non-contemporary spatial dependence lags, not only in the endogenous but also in the exogenous variables. Anselin (2001a) has presented a brief taxonomy for panel data models with different kind of spatial dependence structure for the endogenous variable (space, time and space-time), referring to them as pure space-recursive, time-space recursive, time-space simultaneous and time-space dynamic models. Space-time dependence has also been specified in either theoretical frameworks (Baltagi *et al.*, 2003; Pace *et al.*, 2000) or panel data applications (Case, 1991; Elhorst, 2001; Yilmaz *et al.*, 2002; Baltagi and Li, 2003; Mobley, 2003).

In this article, we analyze the spatial dependence structure in regression models allowing for not only horizontal (static) but also space-time interaction (dynamic). We propose the use of the space-time dependence regression model, that better expresses the effects due to spatial interaction as spatial diffusion phenomena, which is not only "horizontal" – simultaneous – but also time-wise. Further, we try to answer two questions that can only be quantified by these kinds of models: when a shock is produced in a certain location, 1) "what proportion of this shock will be tranferred to the surroundings?" and 2) "how long does it take until the diffusion process is complete?"

[1] W matrix is the so-called spatial weight matrix, which has been profusely defined in the literature (e.g., Cliff and Ord 1975, 1981; Anselin 1988).

The paper proceeds as follows: in the next section, we show some of the commonly used exploratory spatial data analysis (ESDA) techniques and we propose some new ones, the exploratory space-time data analysis (ESTDA) that evaluates the instantaneity of spatial dependence. In section 3, a space-time correlogram is used as an instrument for the identification of space-time dependence models. It is illustrated with some examples for economic series of Spanish provinces is the period 1986-2002. Some summary conclusions complete the paper.

2. EXPLORATORY SPACE-TIME DATA ANALYSIS (ESTDA)

Before analyzing the space-time confirmatory process, some ESDA tools are shown to illustrate space-time processes. First, we briefly present (without going into further details) the bivariate Moran spatial autocorrelation statistic to specifically apply it to analyze the same variable at two different moments of time. In this section, we present new indicators and exploratory tools for the analysis of space-time processes, which have been defined as "exploratory space-time data analysis" or ESTDA.

Our goal is twofold: first, we contribute towards obtaining appropriate indicators to evaluate the dynamic diffusion of spatial dependence and secondly, we develop some statistics for the detection of contemporary – simultaneous – and non-contemporary – lagged – spatial dependence present in a wide-range of socio-economic phenomena.

2.1. Bivariate Moran spatial autocorrelation statistic

The bivariate Moran spatial autocorrelation statistic quantifies the spatial dependence degree existent between two variables Y_k, Y_l in a same location i (Anselin *et al*. 2002). This yields a multivariate counterpart of a Moran-like spatial autocorrelation statistic as follows:

(1) $$I_{kl} = \frac{z_k' W z_l}{z_k' z_k}$$

(2) or: $$I_{kl} = z_k' W z_l / n$$

where: $z_k = [Y_k - \mu_k]/\sigma_k$ and $z_l = [Y_l - \mu_l]/\sigma_l$ have been standardized such that the mean is zero and standard deviation equals one; W is the familiar spatial weight matrix that defines the neighborhood interactions existent in a spatial sample (Anselin, 1988); in this context, the usual row-standardized form of the spatial weights matrix can be used, yielding an interpretation of the spatial lag as an "average"[2] of neighboring values; and n is the number of observations. Since the z variables are standardized, the sum of squares used in the denominator of (1) is constant and equal to n, no matter whether z_k or z_l are used.

Our focus is on the linear association between a variable z_k at a location i (z_{ik}), and the corresponding "spatial lag" for the other variable, $[Wz_l]_i$. This concept was derived from multivariate spatial correlation (Wartenberg, 1985) and thus centers on the extent to which values for one variable z_k observed at a given location show a systematic (more than likely under spatial randomness) association with another variable z_l observed at the neighboring locations. The significance of this multivariate spatial correlation can be assessed in the usual fashion by means of a randomization (or permutation) approach. In this case, the observed values for one of the variables are randomly reallocated to locations and the statistic is recomputed for each such random pattern.

There is also a bivariate generalization of Moran scatterplot that corresponds with a scatterplot with the z_l spatial lag, Wz_l, on the vertical axis and the variable z_k on the horizontal axis (using the variables in standardized form). The slope of the regression line in this scatterplot is equal to the value of the expression (2). In addition, it is also possible to examine each individual location in terms of its placement within the four quadrants of the scatterplot, which define the four types of multivariate spatial association.

2.2. Space-time autocorrelation

When considering both space-time dimensions, some ESTDA tools can be defined to analyze and visualize the space-time structure of a distribution. These are the cases of the Moran space-time autocorrelation statistic, space-time Moran scatterplot, Moran space-time autocorrelation function (MSTAF) and Moran's I line graph.

2.2.1. Moran space-time autocorrelation statistic

Instead of being completely different, variables z_l and z_k could be the same variable observed in two instants of time, t and t-k, with the only limitation that future values cannot explain past

[2] It corresponds with an average but it is not a mean in a strict sense.

ones. In this case, the bivariate Moran's *I* computes the relationship between the spatial lag, Wz_t, at time t and the original variable, z, at time t-k (k order time lag). Therefore, this statistic quantifies the influence that a change in a spatial variable z, that operated in the past (t-k) in an individual location i (z_{t-k}) exerts over its neighborhood at the present time (Wz_t). Hence, it is possible to define the Moran space-time autocorrelation statistic as follows:

$$(3) \quad I_{t-k,t} = \frac{z'_{t-k} W z_t}{z'_{t-k} z_{t-k}}$$

where, as in the last case, the denominator can be substituted by n as this variable z is also standardized. The value adopted by this index, as in (1), corresponds with the slope in the regression line of Wz_t on z_{t-k}. Note that for $k=0$, this statistic (3) coincides with the familiar univariate Moran's I that from now on, we denote as I_t.

Since the Moran's space-time autocorrelation coefficient equals to the slope of the regression of Wz_{t-k} on z_t, it is possible to connect this statistic with the standard Pearson correlation coefficient between these two variables:

$$(4) \quad Corr(z_{t-k}, Wz_t) = r_{z_{t-k}, Wz_t} = \frac{Cov(z_{t-k}, Wz_t)}{\sqrt{Var(z_{t-k})}\sqrt{Var(Wz_t)}} = \frac{\frac{1}{n} z'_{t-k} W z_t}{\sqrt{Var(Wz_t)}}$$

$$(5) \quad \text{or:} \qquad r_{z_{t-k}, Wz_t} = \frac{I_{t-k,t}}{\sqrt{Var(Wz_t)}}$$

The Moran space-time autocorrelation statistic can also be expressed as:

$$(6) \quad I_{t-k,t} = r_{z_{t-k}, Wz_t} \sqrt{Var(Wz_t)}$$

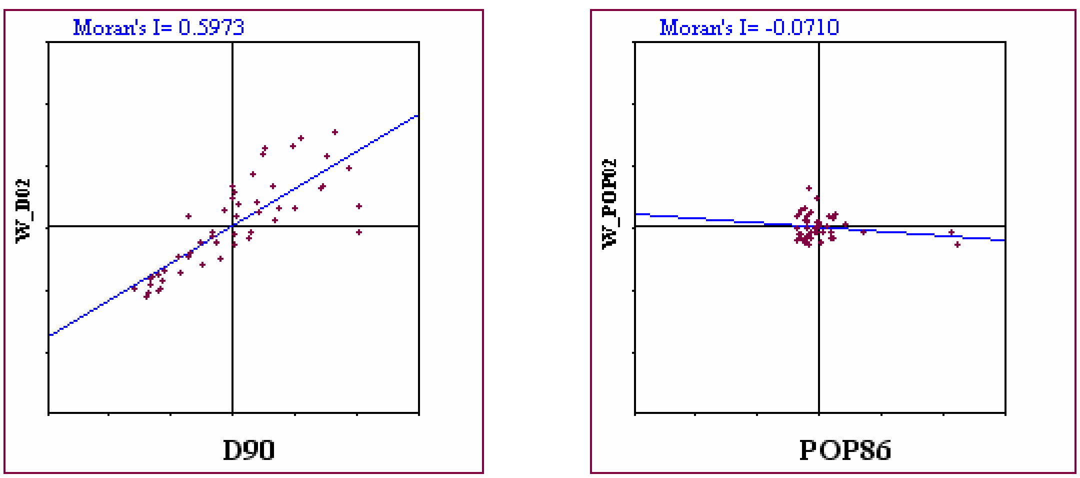
2.2.2. Space-time Moran scatterplot

In parallel with the bivariate Moran scatterplot (Anselin *et al*. 2002), the space-time Moran scatterplot corresponds with a scatterplot with the z_t spatial lag, Wz_t, on the vertical axis and the variable z_{t-k} on the horizontal axis (using the variables in standardized form). The slope of the regression is equal to the value of the expression (3). In this case, it is also possible to examine each individual location as associated with the four quadrants of the scatterplot, which are the four types of space-time spatial association.

In figure 1, two different cases are shown. On the left hand side, the space-time Moran scatterplot shows on the horizontal axis the bank deposits rate[3] of the 50 Spanish provinces in 1990 (*D90*) and on the vertical axis, the correspondent spatial lag in 2002 (*W_D02*), considering *W* as a row-standardized contiguity matrix (two provinces are neighbors if they share a common border). As can be seen, there is a high connection between the bank deposits rate variable in 1990 and its spatial lag 12 years later, as it is shown by the Moran space-time autocorrelation statistic ($I_{90,02}=0.5973$; *p-value=0.001*). In the right graph, we have represented a different situation of non-space-time autocorrelation between population for Spanish provinces in 1986 and the corresponding spatial lag in 2002 ($I_{86,02}=-0.071$; *p-value=0.427*) .

Figure 1: **Space-time Moran scatterplot**

Bank deposits rate **Population**



Source: Estimates by authors with GeoDa (Anselin, 2003)

2.2.3. Moran space-time autocorrelation function (MSTAF)

The MSTAF is the result of plotting the values of the Moran space-time autocorrelation statistic (3), adopted by a variable in a certain period of time. It is a coordinate graph in which the Moran coefficient values are plotted on the vertical axis and the time lags on the horizontal one. The first value corresponds to the contemporary case, $k=0$, which is the univariate Moran's I (I_t), whereas the other ones are proper Moran space-time autocorrelation coefficients ($I_{t-k,t}$). This graph visualizes the influence that a change in a spatial variable *z*, that operated in the past (from

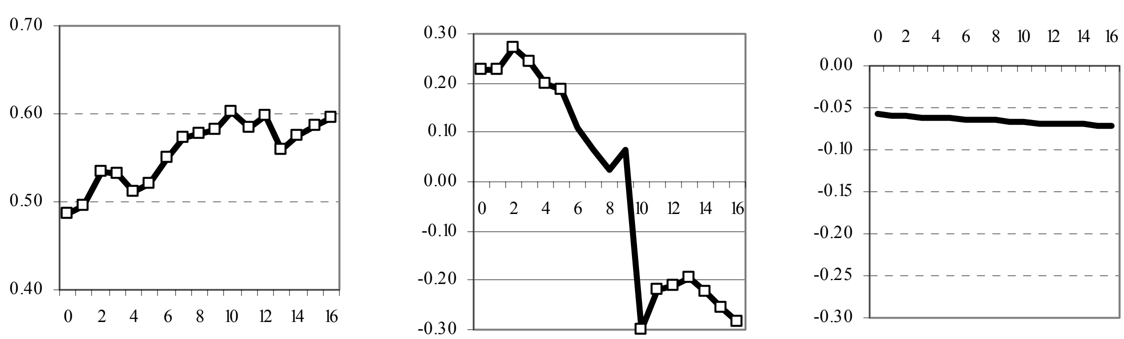
[3] Definitions for all the variables used in this paper are shown in Table 1.

t to *t-k*) in an individual location *i* (z_{t-k}) exert over its neighborhood in the present time (Wz_t). Inference is necessary to evaluate the significance of $I_{t-k,t}$ values and, as a result, the existence or absence of spatial autocorrelation, either a contemporary or non-contemporary one.

In figure 2, we have represented this MSTAF function for three variables: bank deposits rate, price index and population of the 50 Spanish provinces in year 2002. In the horizontal axis, we have represented 16 time lags (period 1986-2002) and the initial moment, lag 0 (year 2002). As can be appreciated, there is clear evidence of non-contemporary spatial dependence in the first function, as all the values exhibit significant high time-lagged values[4] and it shows an increasing trend (the influence of past values of this variable in a certain location over its neighborhood in the present increases with time). The price index variable also has some high time-lagged values in absolute terms (positive in the beginning and negative the end of the period), providing evidence of non-contemporary spatial dependence. Nevertheless it shows a decreasing trend, i.e. the influence of past values of this variable in a certain location over its neighborhood in the present decreases with time in this period.

Figure 2: Moran space-time autocorrelation function (MSTAF)

Bank deposits rate, 2002	**Price index, 2002**	**Population, 2002**



Note: □ is **5%** significance level, permutation approach (999 permutations). *Source: Self-elaboration.*

Regarding the population MSTAF function, it shows very low, almost constant, negative values that are not significant throughout all the period. We can conclude that there is no spatial

[4] All inference computations were done with the permutation approach and 999 permutations.

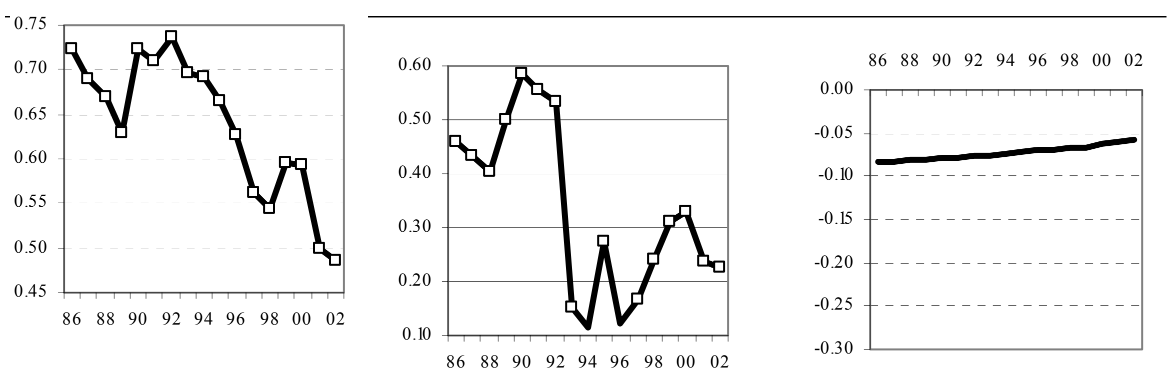
autocorrelation, either contemporary or lagged ones, in the population function of Spanish provinces during 1986-2002.

2.2.4. Moran's *I* line graph.

The Moran's *I* line graph allows showing the evolution of spatial dependence in a period of time. Other authors have already used this plot to explore the dynamics of contemporary spatial dependence in a period of time (Rey and Montouri, 1999). In figure 3, we have represented this coefficient trend for the former provincial variables from 1986 to 2002: bank deposits rate, index price and population. As can be seen, contemporary spatial dependence is very significant and constantly high in all bank deposits rate distributions (especially in 1990 and 1992) and almost all the price index ones (with the exception of 1994 and 1996). On the contrary, all the population variables have non-significant low Moran's *I* values during the whole period. The economic interpretation of these results points out a different behavior of these distributions throughout the Spanish provinces in 1986-2002: while provinces with relatively high (low) bank deposits rate/price index tend to be located, in a same moment of time, nearby other provinces with high (low) bank deposits/prices more often than would be expected due to random chance, population distributions are not clustered at all.

Figure 3: Figure 3: **Moran's I line graph**

Bank deposits rate, 2002	**Price index, 2002**	**Population, 2002**



Note: □ is **5%** significance level, permutation approach (999 permutations). *Source: Self-elaboration.*

2.3. Moran space-time partial autocorrelation coefficients

There is no doubt that the spatial dependence measures that have been presented include different sources of dependence that are difficult to separate. Formally:

$$(7) \quad Cov(z_{it}, z_{js}) \neq 0$$

where sub-indexes i, j are different spatial locations and t, s are different instants of time. Therefore, we consider the following types of dependencies:

(a) there is dependence in expression (7) that is the result of time evolution:

$$(8) \quad Cov(z_{it}, z_{js}) \neq 0 \quad ; \quad \forall i = j$$

This expression affirms that (for s=t-k) the value of the z variable in period t is more or less related to t-k. This assertion is more correct for lower values of k.

(b) there is a dependence in expression (7) that is the result of spatial interactions:

$$(9) \quad Cov(z_{it}, z_{js}) \neq 0\,; \quad \forall t = s$$

This second type of dependence –spatial dependence- can be produced by two sources:

(b1) **Simultaneous or contemporary dependence** constitutes the usual definition of spatial dependence in the literature and it is the consequence of an instant, very rapid, spatial diffusion of a phenomenon in geographic space. It can be connected to or the consequence of a lack of concordance between a spatial observation and the region in which the phenomenon is analyzed.

(b2) **Lagged or non-contemporary dependence** is the result of a slower diffusion of a phenomenon towards the surrounding space. This kind of dependence is due to the usual interchange flows existing between neighbor areas, which requires of a certain period of time to be tested.

Although it is very difficult to divide spatial dependence into its two dimensions, it is worth trying to compute them separately in order to correctly specify a spatial process that exhibits spatial dependence. One of the aims of this article is to show a new range of ESTDA tools that allow justifying the inclusion of both kind of spatial lags, contemporary (Wy_t) and time-lagged (Wy_{t-k}) ones, to explain y_t in a spatial regression. Some coefficients can be defined to evaluate the inclusion of a space-time lag term in a spatial regression. The basic underlying idea consists of eliminating the influence of one of the dimensions in order to compute separately contemporary and non-contemporary dependence. For this purpose, we substitute in (6) the space-time correlation coefficient by a partial correlation one.

A first expression computes the correlation between variable *z* in period *t-k* and its spatial lag *Wz* in period *t* removing the influence of *z* in period *t*. It can be defined as **Moran space-time partial autocorrelation statistic**:

(10) $I_{t-k,t}^{P} = Corr(z_{t-k}, Wz_t \mid z_t)\sqrt{Var(Wz_t)} \quad ; \quad k = 1,2,...,t-1$

where $Corr(z_{t-k}, Wz_t \mid z_t)$ is the partial correlation coefficient of variables z_{t-k} and Wz_t after eliminating the correlation from z_t: $Corr(z_{t-k}, Wz_t \mid z_t) = \dfrac{r_{z_{t-k},Wz_t} - r_{z_{t-k},z_t} \cdot r_{Wz_t,z_t}}{\sqrt{1-r^2_{z_{t-k},z_t}} \cdot \sqrt{1-r^2_{Wz_t,z_t}}}$, where *r* is the standard Pearson correlation coefficient. Therefore, it is possible to express this statistic as a function of both Moran's *I* and space-time Moran's *I*: $I_{t-k}^{P} = \dfrac{I_{t-k,t} - r_{z_{t-k},z_t} \cdot I_t}{\sqrt{1-r^2_{z_{t-k},z_t}} \cdot \sqrt{1-r^2_{Wz_t,z_t}}}$.

This indicator removes contemporary spatial dependence from the relationship between variables z_{t-k} and Wz_t. If the pattern of spatial dependence is one that can be captured by contemporary spatial autoregression, then the partial autocorrelation will be close to zero. On the contrary, if the process is one that can be captured by non-contemporary spatial dependence, then $I_{t-k,t}^{P}$ will be significantly different from zero.

The complementary expression consists of computing contemporary, or instant, spatial dependence after removing lagged spatial dependence by the means of an index that can be defined as the **partial Moran's *I***:

(11) $I_t^{P_k} = Corr(z_t, Wz_t \mid z_{t-k})\sqrt{Var(Wz_t)} \quad ; \quad k = 1,2,...,t-1$

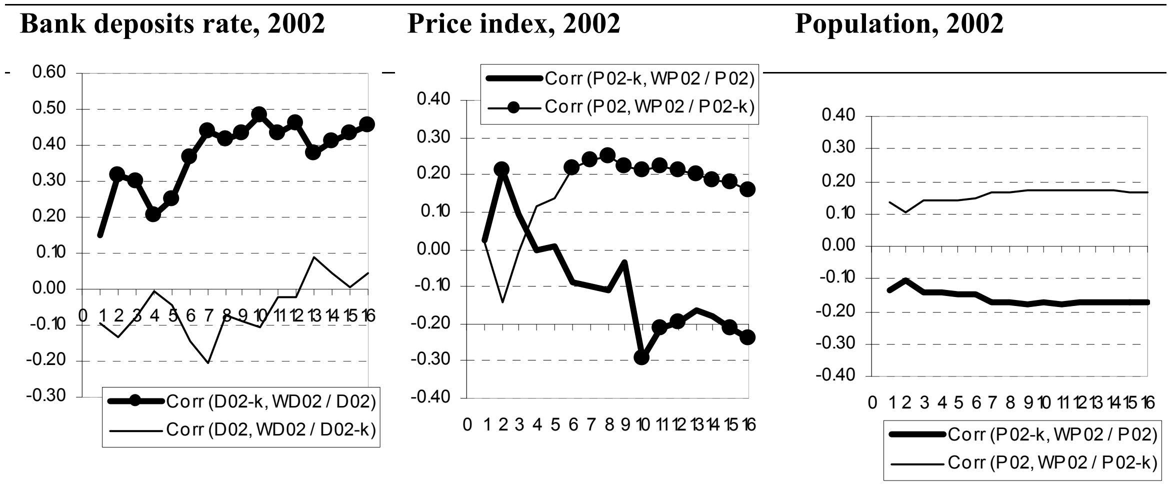
where $Corr(z_t, Wz_t \mid z_{t-k})$ is the partial correlation coefficient of variables z_t and Wz_t after eliminating the correlation from Wz_{t-k}: $Corr(z_t, Wz_t \mid z_{t-k}) = \dfrac{r_{z_t,Wz_t} - r_{z_t,z_{t-k}} \cdot r_{Wz_t,z_{t-k}}}{\sqrt{1-r^2_{z_t,z_{t-k}}} \cdot \sqrt{1-r^2_{Wz_t,z_{t-k}}}}$. Therefore, it is possible to express this statistic as a function of both Moran's *I* and space-time Moran's *I*:

$$I_t^{P_k} = \frac{I_t - r_{z_{t-k},z_t} \cdot I_{t-k,t}}{\sqrt{1-r^2_{z_{t-k},z_t}} \cdot \sqrt{1-r^2_{Wz_t,z_{t-k}}}}.$$

This indicator removes lagged spatial dependence from the contemporary relationship between variables z_t and Wz_t. If the pattern of spatial dependence is one that can be captured by lagged spatial autoregression, then the partial Moran's *I* will be close to zero. On the contrary, if the process is one that can be captured by contemporary spatial dependence, then $I_t^{P_k}$ will be significantly different from zero. The inference of the partial correlation coefficients can be applied to both space-time autocorrelation statistics, as they are the result of multiplying the former by a constant.

The results obtained for these indexes for provincial bank deposits rate, price index and population distributions in period 2002 are in figure 4. The bold line corresponds to $I_{t-k,t}^{P}$ (Moran's space-time partial correlation coefficient) whereas the thin line is the one of $I_t^{P_k}$ (partial Moran's I). The interpretation of the results is as follows. In the case of the bank deposit rate, the $I_{t-k,t}^{P}$ function has higher significant values than the $I_t^{P_k}$ during the whole period, while in the case of price index, $I_{t-k,t}^{P}$ function has higher values than $I_t^{P_k}$ only in the first three years and there is an inflexion in lag 4 (1998), from which the partial Moran's *I* values are higher. In this case, the $I_t^{P_k}$ has significant values from lag 6 to the end of the period and $I_{t-k,t}^{P}$ function has significant values for lags 2, 10, 11, 12, 15, 16, though the last 5 ones have negative values, which makes no economic sense. So we can conclude on the one hand, that there is only a non-contemporary spatial dependence in the explanation of bank deposits rate in 2002, which is increases to the end of the period. On the other hand, it can also be stated that non-contemporary spatial dependence is also present as a factor in the explanation of the price index in 2002, but only in lag 2 (2000), does it have both economic and statistical significance. Therefore, a slower spatial diffusion speed is expected in bank deposits rate, but the price index has a much quicker diffusion speed.

Figure 4: Moran space-time partial autocorrelation functions



Note: ● is **5%** significance level. *Source: Self-elaboration.*

In the case of population, there is no evidence of either instant or lagged spatial effect, as both $I^{P}_{t-k,t}$ and $I^{P_k}_{t}$ values are constant and very close to zero.

3. IDENTIFICATION OF SPACE-TIME REGRESSION MODELS

The joint representation of the Moran's space-time autocorrelation function (MSTAF) in combination with the Moran space-time partial autocorrelation functions (MSTPF) leads to the **space-time correlogram**, which is useful to identify space-time autocorrelation processes. Therefore, this correlogram is a two-graph representation of three functions: a total autocorrelation function ($I_{t-k,t}$) and two partial ones ($I^{P}_{t-k,t}$, $I^{P_k}_{t}$).

The identification process should be conducted in two steps as follows:

- First, the Moran space-time autocorrelation function (MSTAF) indicates the existence (or non-existence) of dynamic spatial dependence. If the MSTAF values are significant (using the regular inference process) we can conclude that there is spatial and temporal dependence in the corresponding distribution, and vice versa. The MSTAF trend is an indicator of the diffusion speed of the spatial distribution:

A decreasing trend in this function points out a **quicker space-time diffusion** of the spatial distribution, as correlations between z_{t-k} and Wz_t are lower for higher values of k.

An increasing trend in MSTAF is a symptom of a **slower space-time diffusion** of the spatial distribution, as past values z_{t-k} will have more influence on the present spatial lagged ones (Wz_t).

A lack of trend, with low values, in MSTAF indicates that there **is no space-time diffusion** in the spatial distribution, so the analysis should finish in this step.

- Secondly, the Moran space-time partial autocorrelation functions (MSTPF) are the instrument to determine whether the existent spatial dependence can be divided into instant and lagged or if it is only instant or only lagged spatial dependence.

Contemporary or instant spatial dependence is present in a variable if only the partial Moran's I has significant values. In this case, only the present values of variable y (y_t) can explain its present spatial lag (Wy_t). In a spatial regression, if an endogenous variable y_t exhibits significant MSTAF and partial Moran's I values, the spatial autocorrelation detected by this correlogram can be completely captured by a contemporary spatial lag of y_t (Wy_t) as an explicative variable in the model.

$$(12)\quad y_t = \alpha + \rho W y_t + \varepsilon_t$$

ρ is the spatial parameter to estimate and ε the error term. This the well-know spatial-lag model. In this case, the use of ordinary least squares (OLS) in the presence of non-spherical errors would yield inconsistent estimators due to the presence of a stochastic regressor Wy_t. Therefore, this model must be estimated by ML or instrumental variables (IV) method (for a more extensive review, see Anselin, 1988).

Non-contemporary or lagged spatial dependence is present in a variable if only the Moran space-time partial autocorrelation function has significant values. In effect, past values of variable y (y_{t-k}) completely explain its present spatial lag (Wy_t), so y_t and Wy_t are not correlated. In this case, the existence of spatial dependence in an endogenous variable y_t, detected by the correlogram, can be completely captured by a space-time lag of y (Wy_{t-k}) as an explicative, exogenous, variable in the model.

(13) $y_t = \alpha + \rho W y_{t-k} + \varepsilon_t$

This model can be estimated by OLS, as the spatial-lag is not correlated with the errors. The most explicative time lagged variable (y_{t-k}) should be the one with highest value in the MSTAF. In practice, the most explicative time lagged variable (y_{t-k}) is not so easy to find. Several alternative models should be estimated to determine the best space-time lag from the group of most relevant values in MSTAF.

If these significant values of the Moran space-time autocorrelation function are closer to the present time (lag 0), it can be said that the corresponding variables reveal a quick spatial diffusion process. On the other hand, if the significant values are concentrated farther from the initial time lags, the variable will have a slower spatial diffusion process.

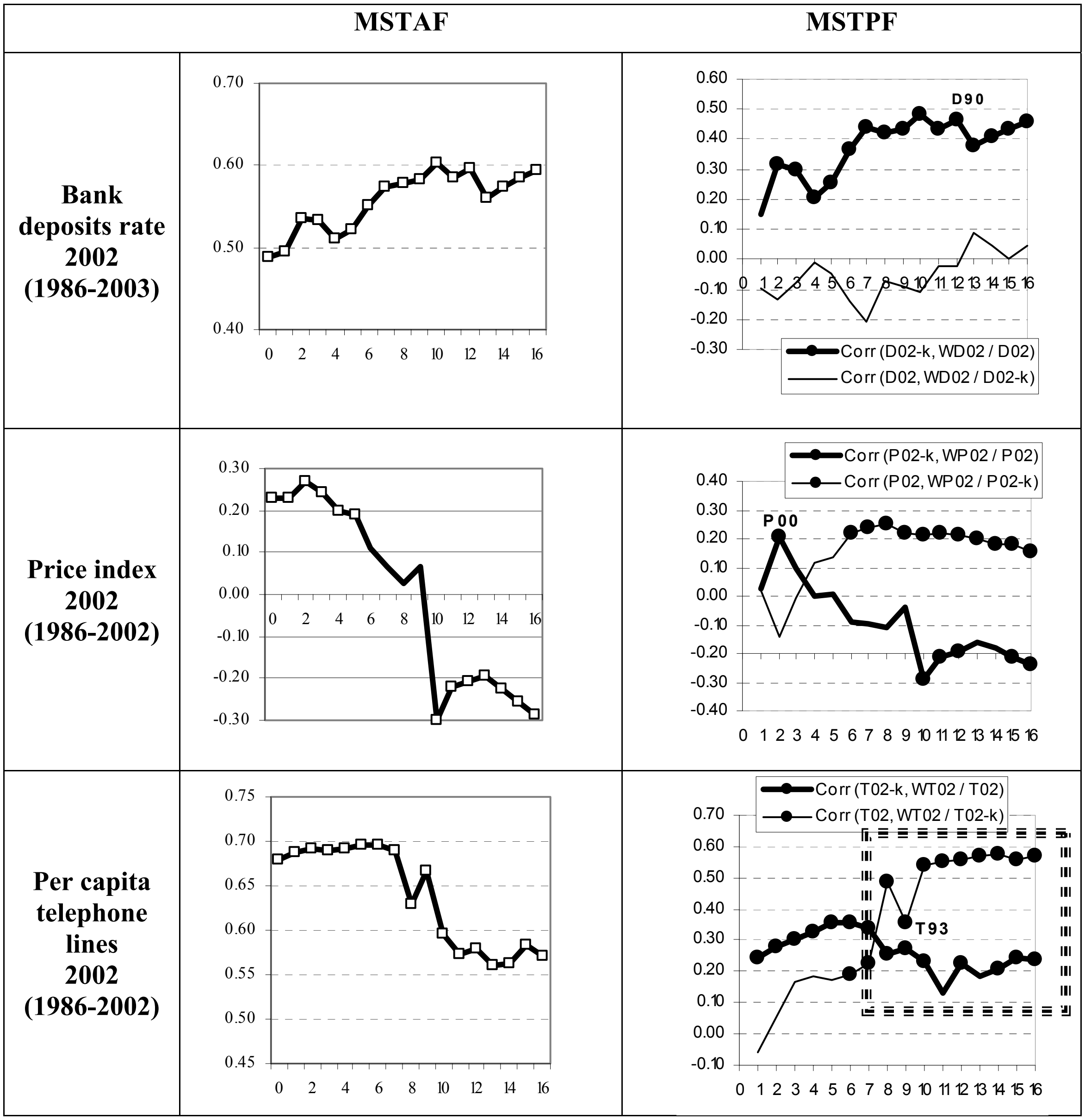
Mixed contemporary and non-contemporary spatial dependence is present in a variable if both partial functions have high significant values for the same periods. In this case, not only present but also past values of variable *z* can completely explain its present spatial lag. Therefore, the existence of spatial dependence in an endogenous variable y_t, detected by the correlogram, can be captured by both a spatial lag and a space-time lag of y (Wy_t, Wy_{t-k}) as explicative (exogenous) variables in the model.

(14) $y_t = \alpha + \rho_1 W y_t + \rho_2 W y_{t-k} + \varepsilon_t$

ρ_1, ρ_2 are spatial parameters to estimate. This is also a space-lag model so it can be only estimated by ML or IV. In this case, the most explicative time lagged variable (y_{t-k}) should also be determined from the group of most significant values in both partial functions.

In figure 5, we show the space-time correlograms of 3 variables: banks deposits rate, price index and per capita telephone lines (population is excluded as no space-time autocorrelation has been detected).

Figure 5: Space-time correlograms



Note: □ is **5%** Moran's I significance level (MSTAF). ● is **5%** significance level (MSTPF). *Source: Self-elaboration.*

The MSTAF of bank deposits rate in 2002 has highly significant increasing values throughout the period 1986-2002; so strong space-time dependence with a slower diffusion is expected in this variable. On the contrary, the MSTAF of the second variable, index price is 2002, has a decreasing trend and only the first 6 time lags have significant positive values, hence weaker space-time dependence with a quick diffusion is expected here. The MSTAF of the third variable, per capita telephone lines, has a peculiar shape, with high and almost constant values from lags 1-9 (2001-1993), which is indicative of non-contemporary spatial dependence that decreases from lag 9 to the end of the period, pointing out the predominance of instant dependence in this period. Therefore, two possible solutions could be expected in this variable: either non-contemporary spatial dependence during 1993-2001 or both contemporary and non-contemporary spatial dependence afterwards.

Regarding to the MSTPF plots, in the bank deposits rate variable 2002 the Moran space-time partial autocorrelation function is always higher than the partial Moran's *I* and all its values have less than 5% significance level (except lag 1). So this is a case of pure non-contemporary or lagged spatial dependence (expression 13). That is why the existence of spatial dependence in bank deposits rate in 2002 is captured by a space-time lag variable, that must be found from the group of those Moran's space-time partial autocorrelation with most significant (highest) values (e.g. lags 10, 11, 12). The selection of the best specification must be determined after the corresponding estimation of these space-time models (for instance, the model with best measure of fit –e.g. the AIC[5]). As it shown in table 2, in the case of bank deposits rate, all the models points out the clear supremacy of lagged spatial dependence over instant one, which is never significant. The best specification corresponds with model (13), a non-contemporary spatial model, in which the only exogenous variable of bank deposits rate in 2002 is its corresponding space-time lag in 1990 ($k=12$). Thus, it can also be concluded that the spatial diffusion process of this variable in 2002, at the provincial level in Spain, is strong (with a high coefficient) and slow, as it lasts 12 years.

(15) $\hat{D}02 = 0.008 + 0.85 \cdot WD90$

In the price index variable, partial Moran's *I* function is always higher than the Moran space-time partial autocorrelation, with the exception of lags 1 to 3, though only lag 2 have less than 5% significance level (as we have excluded significant negative values for not having economic sense). Therefore, this is also a non-contemporary spatial dependence case, where lagged spatial dependence is stronger than instant one. So spatial dependence in the price index variable in

[5] The Akaike Information Criterion (AIC) is a ML-based statistic that, as well as the log likelihood (LIK) measure, is appropriate to compare models estimated by different methods (e.g. OLS and ML). But AIC corrects the LIK for overfitting, which is very important when also comparing models with different number of regressors (Anselin, 1992). In order to preserve the normality assumption, which is convenient when using the AIC, all the models have been estimated in deviations to the mean. It also allows making comparisons between coefficients.

2002 is captured by a space-time lag variable in lag 2 (P00). After the corresponding estimation of these space-time models, shown in table 2, in the price index variable only a few models have significant estimates and most of them have a lagged spatial dependence variable. The best specification corresponds with model (13), a non-contemporary spatial model, in which the only exogenous variable of price index in 2002 is its corresponding space-time lag in 2000 ($k=2$). Thus, it can also be concluded that the spatial diffusion process of this variable in 2002, at the provincial level in Spain, is quick -it lasts only 2 years- but weaker than in the previous example (lower coefficient).

(16) $\hat{P}02 = 0.02 + 0.59 \cdot WP00$

In per capita telephone lines, the Moran space-time partial autocorrelation function is higher than the partial Moran's *I* from the beginning to lag 7 and the contrary from lag 8 to the end. There are a lot of significant values (with less than 5% significance level) in both functions, with the exception of lags 11 and 13, in the first one, and lags 1 to 6, in the second. There are some periods in which both functions have significant values: lags 7-10 (1992-1995), 12 (1990) and 15-16 (1986-1987). This is a clear indication that spatial dependence can be decomposed in two components: instant and lagged. Spatial dependence in the per capita telephone lines variable in 2002 is captured by both contemporary and non-contemporary spatial dependence specification. The lagged spatial lag term must be selected from the periods in which both partial functions have significant values. After the corresponding estimation of these space-time models, shown in table 2, the best specification corresponds to a mixed contemporary and non-contemporary spatial model, in which the explicative variables of per capita telephone lines in 2002 are its corresponding space-time lags in 2002 and 1993 ($k=9$).

(17) $\hat{T}02 = 0.38 + 0.38 \cdot WT02 + 0.69 \cdot WT93$

Thus, it can also be concluded that this variable in 2002, at the provincial level in Spain, has two spatial diffusion processes: an instant (very quick) process and a slower (9 years) one, though the slower-lagged process is almost half stronger than the quicker-instant one.

The other two variables also have been analyzed and their correspondent space-time correlograms are shown in the figure in the Appendix. We have obtained their corresponding space-time correlogram and highlighted the significant values and we have also estimated all the models (the results are in the Appendix). Note that the employment rate can has a strong spatial dependence that can be decomposed in two components, lagged and instant. So spatial diffusion has also two speeds for this variable: a quicker one and a slower spatial diffusion, which is stronger and lasts 10 years. And finally, per capita registered cars variable is another case of pure non-contemporary spatial dependence model with a strong slower spatial diffusion effect (12 years).

6. CONCLUSIONS

The main aim of this paper was the analysis of the dynamics of spatial dependence making a differentiation between two types of spatial dependence: instant or contemporaneous and lagged or non-contemporaneous. The first one is the consequence of a very quick diffusion of the process over the neighboring locations, while the second one implies that a shock in a certain location needs of several periods of time to take place and be tested over its neighborhood. Hence, we propose the use of the space-time dependence regression model, which better expresses the effects due to spatial interaction as spatial diffusion phenomena, which is not only "horizontal" or simultaneous but also time-wise.

For the fulfillment of this aim, we propose new exploratory space-time data analysis (ESTDA) tools that evaluate the instantaneity of spatial dependence and a space-time correlogram is used as a valid instrument for the identification of space-time dependence models. In the second part of this paper, we illustrated the process for the identification of different types of spatial dependence in some variables, with the help of the space-time correlogram. We have shown that spatial dependence, when present in a variable, can be decomposed in two components -lagged and instant- or in case of weaker instant spatial dependence, only a space-lag spatial dependence should be specified.

All the shown instruments allow us to answer the two questions we proposed in the introduction, namely when a shock is produced in a certain location, 1) "which proportion of this shock will be translated to the surroundings?" and 2) "how long does it take until the diffusion process completely end up?"

Referring to the first question, in all the analyzed variables, a greater proportion of a shock in a location was translated to its surroundings. In the case of per capita telephone lines and employment rate, both kind of spatial dependence are present, though the proportion of a shock occurred in the past is bigger in both cases than the impact of a shock happened in the present moment. Analysis of the second question reveals that there is more diversity between variables. For example, index price shows a quick –but weaker- speed of spatial diffusion (2 years), whereas per capita registered cars and bank deposits present a strong but lower speed (12 years both).

6. REFERENCES

Anselin, L. (1988) *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.

Anselin, L. (1992), *"Space Stat tutorial. A workbook for using SpaceStat in the analysis of spatial data".* Technical Report S-92-1, National Center for Geographic Information and Analysis, University of California. Santa Barbara, CA.

Anselin, L. (2001a) "Spatial econometrics". In *A companion to theoretical econometrics*, ed. Baltagi, Oxford: Basil Blackwell; pp. 310-330.

Anselin, L. (2001b), "Spatial effects in econometric practice in environmental and resource economics". *American Journal of Agricultural Economics*, 83 (3), pp. 705-710.

Anselin, L., I. Syabri and O. Smirnov (2002). "Visualizing multivariate spatial correlation with dynamically linked windows". In L. Anselin and S. Rey (Eds.), *New Tools in Spatial Data Analysis, Proceedings of a Workshop*. Center for Spatially Integrated Social Science, University of California, Santa Barbara (CD-ROM).

Anselin, L. (2003), *"GeoDa^{TM} 0.9 User's Guide"*. Spatial Analysis Laboratory (SAL), University of Illinois at Urbana-Champaign. http://sal.agecon.uiuc.edu

Baltagi, B.H. and D. Li (2003) "Prediction in the panel data model with spatial correlation". In: Anselin, L., R. Florax and S. Rey (Eds.), *New Advances in Spatial Econometrics*, Springer-Verlag; pp. 283-296.

Baltagi, B.H., S.H. Song and W. Koh (2003) "Testing panel data regression models with spatial error correlation". *Journal of Econometrics*, 117 (1), Nov. 2003; pp. 123-150

Case, A. (1991) "Spatial patterns in household demand". *Econometrica* 59, 953–965.

Cliff, A. and J. Ord (1973), *Spatial autocorrelation.* London: Pion.

Cliff, A. and J. Ord (1981), *Spatial processes, models and applications.* London: Pion.

Elhorst, J.P. (2001) "Dynamic models in space and time". *Geographical Analysis* 33; 119-140.

Mobley, L.R. (2003) "Estimating hospital market pricing: An equilibrium approach using spatial econometrics". *Regional Science and Urban Economics* 33; pp. 489–516.

Pace R.K., R. Barry, O.W. Gilley y C.F. Sirmans (2000) "A method for spatial-temporal forecasting with an application to real estate prices International". *Journal of Forecasting*, 16; pp. 229-246.

Rey, S. and B. Montouri (1999), "US regional income convergence: A spatial econometric perspective". *Regional Studies*, vol. 33.2; pp. 143-156.

Upton, G. and B. Fingleton (1985) *Spatial data analysis by example*. Volume 1. Wiley, Chichester.

Wartenberg D. 1985. "Multivariate spatial correlation: A method for exploratory geographical analysis", *Geographical Analysis*, 17: 263–283.

Yilmaz, S., K.E. Haynes and M. Dinc (2002), "Geographic and network neighbors: Spillonver effects of telecommunications infrastructure". Journal of Regional Science, 42 (2); pp. 339-360.

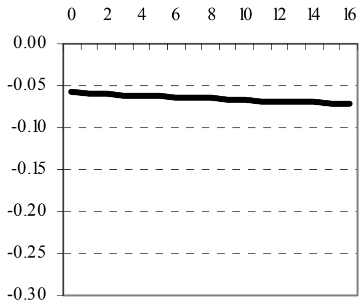
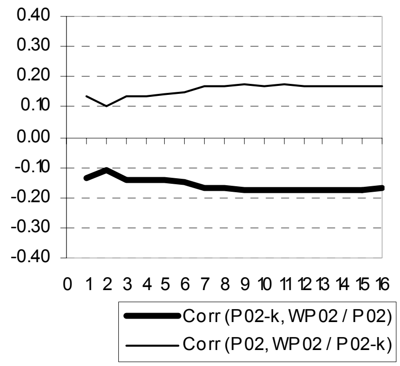
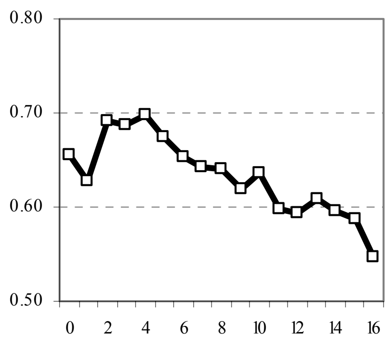
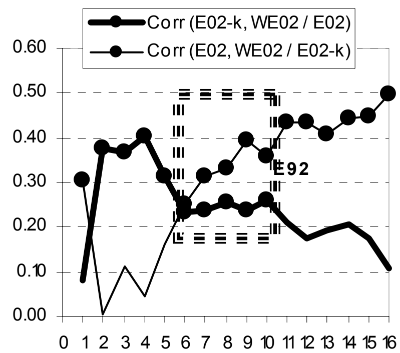
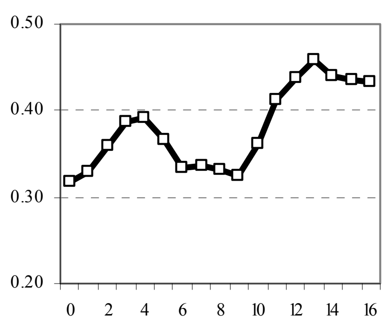
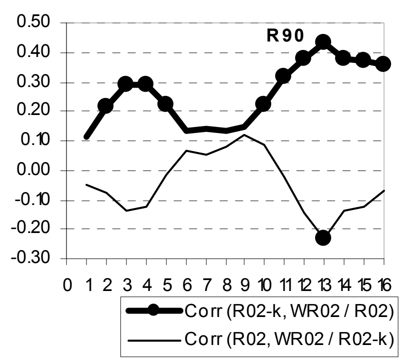
Appendix

Table 1

Definition of the variables used in the models

Variable		**Source**
D	Bank deposits rate (per capita), period 1986-2002, Spanish provinces (euro)	Banco de España
POP	Population, period 1986-2002, Spanish provinces	INE
P	Price index, period 1986-2002, Spanish provinces (1992=100)	INE
T	Per capita telephone lines, period 1986-2002, Spanish provinces	Telefónica, S.A.
E	Employment rate, period 1986-2002, Spanish provinces	INE
R	Per capita registered cars, period 1986-2002, Spanish provinces	DGT

Annex: Space-time correlograms of other variables

	MSTAF	MSTPF
Population (1986-2002)		
Employment rate (1986-2002)		
pc Registered cars (1986-2002)		

Note: □ is **5%** Moran's I significance level (MSTAF). ● is **5%** significance level (MSTPF). *Source: Self-elaboration.*

Table 2

Time	Spatial	Bank deposits rate				Price index				pc Telephone lines				Employment rate				pc Registered cars			
		(13)		(14)		(13)		(14)		(13)		(14)		(13)		(14)		(13)		(14)	
lag	lag	Coef.	AIC	Coef.	AIC	Coef.	AIC	Coef.	AIC	Coef.	AIC	Coef.	AIC	Coef.	AIC	Coef.	AIC	Coef.	AIC	Coef.	AIC
1	2001 2002	**0.91**	116.35	**0.93** -0.02	118.35	**0.55**	139.21	**0.54** 0.01	141.20	**1.08**	76.27	**1.00** 0.08	78.19	**0.94**	97.71	**0.71** 0.24	98.83	**0.93**	127.40	**0.91** 0.03	129.39
2	2000 2002	**0.87**	115.01	**0.88** -0.01	117.01	**0.59**	137.20	**0.67** -0.09	139.10	**1.06**	77.07	**0.91** 0.15	78.71	**0.95**	89.77	**0.90** 0.06	91.72	**0.76**	130.07	**0.61** 0.19	131.50
3	1999 2002	**0.85**	116.16	**0.80** 0.06	118.12	**0.55**	138.47	**0.53** 0.03	140.46	**1.05**	78.70	**0.82** 0.25	79.61	**0.95**	89.46	**0.90** 0.06	91.39	**0.74**	129.26	**0.63** 0.15	130.90
4	1998 2002	**0.87**	116.49	**0.83** 0.05	118.46	**0.50**	140.87	0.35 0.17	142.34	**1.04**	79.27	**0.79** 0.27	79.82	**0.96**	87.94	**0.99** -0.03	89.93	**0.73**	129.26	**0.61** 0.17	130.82
5	1997 2002	**0.87**	116.44	**0.82** 0.05	118.41	**0.57**	141.17	0.38 0.19	142.46	**1.04**	78.90	**0.79** 0.27	79.51	**0.96**	91.71	**0.84** 0.13	93.38	**0.69**	130.98	**0.53** 0.23	132.05
6	1996 2002	**0.84**	115.63	**0.81** 0.03	117.62	0.33	144.52	0.03 0.33	143.77	**1.04**	79.09	**0.78** 0.28	79.54	**0.97**	95.08	**0.70** 0.28	96.72	**0.70**	132.16	**0.51** 0.27	132.85
7	1995 2002	**0.85**	113.42	**0.89** -0.05	115.39	0.13	145.55	-0.03 **0.35**	143.76	**1.03**	80.21	**0.76** 0.30	80.39	**0.97**	96.47	**0.70** 0.28	96.72	**0.67**	132.79	**0.47** 0.28	133.29
8	1994 2002	**0.83**	114.24	**0.79** 0.05	116.19	0.03	145.88	-0.12 **0.37**	143.52	**1.01**	94.08	**0.51** **0.55**	86.43	**1.01**	95.33	**0.72** 0.30	94.99	**0.66**	133.58	**0.45** 0.31	133.62
9	1993 2002	**0.84**	113.44	**0.82** 0.02	115.43	0.13	145.58	-0.01 **0.34**	143.78	**1.04**	82.87	**0.69** **0.38**	81.52	**0.96**	101.46	**0.58** **0.42**	97.95	**0.66**	134.28	**0.43** **0.34**	133.87
10	1992 2002	**0.84**	112.36	**0.83** 0.01	114.36	**-0.50**	138.57	**-0.41** 0.23	138.88	**1.00**	97.99	**0.47** **0.60**	87.15	**0.95**	100.40	**0.59** **0.40**	97.64	**0.77**	130.62	**0.59** 0.25	131.46
11	1991 2002	**0.84**	113.56	**0.78** 0.06	115.48	-0.37	142.11	-0.30 0.29	141.07	**1.03**	97.79	**0.48** **0.59**	87.63	**0.90**	107.40	**0.47** **0.50**	100.78	**0.79**	127.03	**0.70** 0.13	128.76
12	1990 2002	**0.85**	111.90	**0.83** 0.02	113.89	-0.34	142.53	-0.27 0.29	141.60	**1.00**	99.84	**0.46** **0.62**	87.38	**0.92**	105.64	**0.50** **0.48**	100.35	**0.81**	125.86	**0.76** 0.05	127.82
13	1989 2002	**0.86**	115.50	**0.73** 0.15	117.04	-0.35	142.78	-0.27 0.29	141.60	**1.01**	101.25	**0.45** **0.63**	87.95	**0.96**	101.19	**0.58** **0.41**	97.97	**0.75**	126.08	**0.71** 0.07	128.01
14	1988 2002	**0.85**	114.75	**0.74** 0.12	116.44	**-0.47**	140.81	-0.36 0.25	140.75	**1.02**	100.27	**0.46** **0.62**	87.58	**0.95**	103.01	**0.55** **0.45**	98.46	**0.69**	128.58	**0.58** 0.17	130.08
15	1987 2002	**0.84**	114.43	**0.75** 0.11	116.21	**-0.51**	139.24	**-0.41** 0.23	139.60	**1.01**	98.69	**0.47** **0.60**	87.48	**0.93**	104.67	**0.51** **0.47**	99.64	**0.68**	129.06	**0.56** 0.18	130.47
16	1986 2002	**0.83**	114.69	**0.72** 0.14	116.30 0.39	**-0.56**	137.45	**-0.47** 0.18	138.40	**1.01**	99.51	**0.46** **0.61**	87.47	**0.88**	111.31	**0.41** **0.55**	102.59	**0.64**	131.21	**0.48** 0.26	131.84

Note: In black, 5% significance level; black and underlining, 1% significance level; grey, final solution. AIC: Akaike Information Criterion.