

Some Issues on the Concept of Causality in Spatial Econometric Models^{*}

JEAN H.P. PAELINCK^a, JESÚS MUR^b

^a Emeritus Professor, Erasmus University Rotterdam, Burgemeester Oudlaan, 50, 3062 PA Rotterdam, Netherlands. E-mail: j.paelinck@upcmail.nl

^b Universidad de Zaragoza, Facultad de Economía y Empresa, Campus Rio Ebro, c/ María de Luna s/n, 50018 Zaragoza, España. E-mail: jmur@unizar.es

ABSTRACT

A specific feature of spatial econometric models is the simultaneity of the relations, which complicates the distinction between causes and effects. However, the notion of cause is of paramount importance in any specification. In fact, a model is a statement pointing that a variable, called endogenous, reacts to the variables that appear in the right hand side, the regressors. Our impression is that this problem has been scarcely treated in the Spatial Econometrics literature.

The content of the paper focuses on questions related to the specification procedure for spatial models. We examine what may be called the 'current traditional practice' and discuss the role that the concepts of identification and causality should play. Our purpose is to claim for the development of clear econometric guidelines to help the users to improve the theoretical foundations of their specifications. An application to the relation between per capita income and weight of the agricultural sector in the Spanish provinces illustrates our discussion.

Keywords: Causality, Identification, Spatial Econometric Models.

Algunos aspectos del concepto de causalidad en modelos econométricos espaciales

RESUMEN

La simultaneidad de las relaciones es un hecho distintivo de los modelos de econometría espacial, lo que dificulta la separación entre causas y efectos. Sin embargo, la noción de causa tiene una importancia fundamental en cualquier especificación de un modelo. De hecho, un modelo supone una afirmación señalando que una variable, denominada endógena, reacciona a las variables que aparecen en la parte derecha, los regresores. Nuestra impresión es que este problema se ha tratado insuficientemente por la literatura de Econometría espacial.

Este trabajo se centra en cuestiones relacionadas con la especificación de modelos espaciales. Examinamos lo que puede denominarse 'práctica tradicional actual' y discutimos el papel que deberían jugar los conceptos de identificación y causalidad. Nuestro propósito es reclamar el desarrollo de procedimientos econométricos claros que ayuden a los usuarios a mejorar sus especificaciones. Una aplicación al caso de la relación entre la renta per capita y peso del sector agrícola en las provincias españolas, en el año 2006, ilustra esta discusión.

Palabras clave: Causalidad, Identificación, Modelos econométricos espaciales.

JEL Classification: C21, C50, R15

^{*} A preliminary version of this paper was presented at the 3rd World Conference of the Spatial Econometrics Association held in Barcelona (Spain) in 2009. The authors thank the referees and editor for their helpful comments and suggestions. J. Mur has received financial support from the project ECO-2015-65758-P of the Spanish 'Ministerio de Economía y Competitividad'. Finally, this contribution is only possible thanks to the help of Claudette, Genevieve and Marc Paelinck.

Artículo recibido en septiembre 2017 y aceptado en octubre de 2017

Artículo disponible en versión electrónica en la página www.revista-eea.net, ref. 2-36101

1. INTRODUCTION

The notion of causality is basic for the specification of an econometric model. In general, it is supposed that the variables that appear on the right-hand side of the equation cause the variable that appears on the left-hand side. The treatment of the relation is simple if, furthermore, the variables that appear in the right-hand part are either exogenous or predetermined. This discussion is part of the routine in mainstream econometrics, although it has not been elaborated so much in a spatial context (Anselin, 1988).

One of the difficulties comes from the nature of the data because, on many occasions, we only have a single cross-section. Nevertheless, the literature on causality has evolved on the principle of temporal succession, under which the cause must precede the effect. If we do not have a temporal perspective, the discussion about causality is more complicated, but should not be discarded.

In the second Section we introduce some of the main problems that arise in a spatial framework. In the third Section we suggest several approaches to the question of causality. In the fourth Section we apply the proposed procedure to the case of the relation between per capita income and activity of the agricultural sector in the Spanish provinces. The paper finishes with a Section of conclusions and further developments.

2. A GENERAL SPATIAL MODEL. THE IDENTIFICATION PROBLEM

Let us assume that we have information for a set of R individuals over T periods. This panel allows us to specify a model, where the individuals interact between themselves, like the following:

$$y_{rt} = \alpha_{r1}y_{1t} + \dots + \alpha_{rR}y_{Rt} + x'_{1t}\beta_{r1} + \dots + x'_{Rt}\beta_{rR} + u_{rt} \quad (1)$$

$$r = 1, 2, \dots, R \quad t = 1, 2, \dots, T \quad \alpha_{rr} = 1$$

The r -th equation explains the behavior of individual r , y_{rt} , based on what happens in its neighborhood, and on a set of predetermined variables located at point r (which can be extended also to the neighbors); x_{mt} is a vector of order $(k \times 1)$ of observations taken at point m and β_{rm} the corresponding vector of parameters, also of order $(k \times 1)$; u_{rt} is an error term. Equation (1) has been normalized so that $\alpha_{rr}=1$.

Using a more compact notation:

$$BY = \Gamma X' + U \quad (2)$$

where:

$$\begin{aligned}
Y &= \begin{bmatrix} y_{11} & y_{12} & y_{13} & \cdots & y_{1T} \\ y_{21} & y_{22} & y_{23} & \cdots & y_{2T} \\ y_{31} & y_{32} & y_{33} & \cdots & y_{3T} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ y_{R1} & y_{R2} & y_{R3} & \cdots & y_{RT} \end{bmatrix}_{(RxT)} \\
B &= \begin{bmatrix} 1 & -\alpha_{12} & -\alpha_{13} & \cdots & -\alpha_{1R} \\ -\alpha_{22} & 1 & -\alpha_{23} & \cdots & -\alpha_{2R} \\ -\alpha_{31} & -\alpha_{32} & 1 & \cdots & -\alpha_{3R} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -\alpha_{R1} & -\alpha_{R2} & -\alpha_{R3} & \cdots & 1 \end{bmatrix}_{(RxR)} \\
U &= \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1T} \\ u_{21} & u_{22} & u_{23} & \cdots & u_{2T} \\ u_{31} & u_{32} & u_{33} & \cdots & u_{3T} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ u_{R1} & u_{R2} & u_{R3} & \cdots & u_{RT} \end{bmatrix}_{(RxT)} \\
X &= \begin{bmatrix} x_{11} & x_{21} & x_{31} & \cdots & x_{R1} \\ x_{12} & x_{22} & x_{32} & \cdots & x_{R2} \\ x_{13} & x_{23} & x_{33} & \cdots & x_{R3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ x_{1T} & x_{2T} & x_{3T} & \cdots & x_{RT} \end{bmatrix}_{(TxRk)} \\
\Gamma &= \begin{bmatrix} \beta_{11} & \beta_{21} & \beta_{31} & \cdots & \beta_{R1} \\ \beta_{12} & \beta_{22} & \beta_{32} & \cdots & \beta_{R2} \\ \beta_{13} & \beta_{23} & \beta_{33} & \cdots & \beta_{R3} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \beta_{1R} & \beta_{2R} & \beta_{3R} & \cdots & \beta_{RR} \end{bmatrix}_{(RxRk)} \\
&\quad \beta_{mr} = [\beta_{m1r} \ \beta_{m2r} \ \beta_{m3r} \ \cdots \ \beta_{mkr}]_{(k \times 1)}
\end{aligned}$$

The reduced form follows immediately:

$$\begin{aligned}
BY &= \Gamma X' + U \Rightarrow Y = \Pi X' + V \\
\Pi_{(RxRk)} &= B^{-1}\Gamma \\
V &= B^{-1}U
\end{aligned} \tag{3}$$

The following set of hypotheses completes the specification:

$$\begin{aligned}
z_t &= \begin{bmatrix} Y_t \\ X_t \end{bmatrix}_{(R(k+1) \times 1)} \sim N \left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}; \begin{pmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{pmatrix}_{R(k+1) \times R(k+1)} \right] \\
\mu_{t(R \times 1)} &= E[Y_t | X_t = x_t] = (\Sigma_{xx}^{-1} \Sigma_{xy}) x_t \\
\rightarrow v_{t(R \times 1)} &= Y_t - E[Y_t | X_t = x_t] = Y_t - \mu_t \\
E[v_t] &= E[E[v_t | X_t = x_t]] = 0 \\
E[v_t v'_s] &= E[E[v_t v'_s | X_t = x_t]] = \begin{cases} \Omega = \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}; & t = s \\ 0; & t \neq s \end{cases} \\
E[\mu_t v'_t] &= E[E[\mu_t v'_t | X_t = x_t]] E[\mu_t E[v'_t | X_t = x_t]] = 0; \forall t
\end{aligned} \tag{4}$$

There are $R(R-1)$ parameters in B and $R^2 k$ parameters in Γ , making a total of $R[R(k+1)-1]$ structural parameters of interest. Furthermore, in the reduced form of (3), there are $R^2 k$ parameters in the matrix Π . The covariance matrix Ω contains R^2 parameters. It is clear that the model is *underidentified*. To attain identification, we need at least $R(R-1)$ additional restrictions on the structural parameters.

These restrictions may come from different ways in a spatial context; for example, from matrix B in (3). If, as usual, we assume an exogenous, binary and known weighting matrix, which reflects the network of spatial dependencies, we can write:

$$B = I - \rho W$$

$$w_{rs} = \begin{cases} 1 & \text{if } s \in N_r \\ 0 & \text{if } s \notin N_r \end{cases} \quad (5)$$

where N_r is the set of neighbors which are related to point r and ρ a parameter of spatial autocorrelation. In (5) there are $[R(R-1)-1]$ restrictions and we need, at least, one more constraint to attain full identification, which may come from matrix Γ . For example, in many situations it is reasonable to limit the interaction between the endogenous, located at a given point, and the predetermined variables located elsewhere in space. This means that some of the β 's located outside the main diagonal of matrix Γ may be assumed to be zero. If, furthermore, we introduce the restriction of homogeneity of these vectors, we can write:

$$\begin{cases} \beta_{sr} = 0 & \text{if } s \neq r \\ \beta_{ss} = \beta_{rr} = \beta & \text{if } s=r \end{cases} \quad (6)$$

We obtain $(R^2-1)k$ restrictions on the matrix Γ , which, added to those already introduced in the composition of matrix B , give us a total of $[R^2(2k+1)-(R+k)]$ restrictions. The number of parameters of the reduced form continues to be R^2k while in the structural form only intervene $(k+1)$ parameters:

$$B^{-1} = [I_R - \rho W]^{-1} = \begin{bmatrix} b_{11} & b_{12} & b_{13} & \cdots & b_{1R} \\ b_{21} & b_{22} & b_{23} & \cdots & b_{2R} \\ b_{31} & b_{32} & b_{33} & \cdots & b_{3R} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{R1} & b_{R2} & b_{R3} & \cdots & b_{RR} \end{bmatrix} \quad \Gamma = \begin{bmatrix} \beta & 0 & 0 & \cdots & 0 \\ 0 & \beta & 0 & \cdots & 0 \\ 0 & 0 & \beta & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & \beta \end{bmatrix}$$

$$\Pi = B^{-1}\Gamma \Rightarrow \begin{bmatrix} \pi_{11} & \pi_{12} & \pi_{13} & \cdots & \pi_{1R} \\ \pi_{21} & \pi_{22} & \pi_{23} & \cdots & \pi_{2R} \\ \pi_{31} & \pi_{32} & \pi_{33} & \cdots & \pi_{3R} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \pi_{Rk1} & \pi_{Rk2} & \pi_{Rk3} & \cdots & \pi_{RkR} \end{bmatrix} = \begin{bmatrix} b_{11}\beta & b_{12}\beta & b_{13}\beta & \cdots & b_{1R}\beta \\ b_{21}\beta & b_{22}\beta & b_{23}\beta & \cdots & b_{2R}\beta \\ b_{31}\beta & b_{32}\beta & b_{33}\beta & \cdots & b_{3R}\beta \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{R1}\beta & b_{R2}\beta & b_{R3}\beta & \cdots & b_{RR}\beta \end{bmatrix} \quad (7)$$

$$\Pi = B^{-1} \otimes \beta \Leftrightarrow B\Pi = I_R \otimes \beta = \Gamma$$

Note that the b_{rs} terms of matrix B^{-1} depend on one unknown parameter, ρ , and on the weights in W . The model now is *overidentified*, so there are $R(R-1)$ restrictions of overidentification. The covariance matrices are an additional source of restrictions. For example, if we assume no correlation between the error terms of the cross-sections in (1):

$$u_t = B[Y_t - E(Y_t | X_t = x_t)]$$

$$E[u_t] = E[E[v_t | X_t = x_t]] = 0$$

$$E[u_t u'_s] = E[E[v_t v'_s | X_t = x_t]] = \begin{cases} \Lambda; & t = s \\ 0; & t \neq s \end{cases} \quad (8)$$

where Λ is a diagonal matrix; the error of the reduced form verify that $\Omega = B^{-1}\Lambda B'^{-1}$. Moreover, we have $R(R-1)$ new restrictions or (R^2-1) if we also assume homoskedasticity. In the first case (no correlation but skedastic variance), there is a total of $2R[R(k+1)-1]-k$ restrictions of *overidentification* and $2R^2(k-1)-(R+k+1)$ in the second (homoscedasticity). That is to say, as long as we have a sufficient number of temporal cross-sections ($T > Rk$), a spatially (restricted) simultaneous equations model like that of (1) can be estimated in the usual way.

The situation is more difficult if we have only one cross-section ($T=1$). Under this setting, the structural form of the model is:

$$y = \rho Wy + x\beta + u \Leftrightarrow \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ \dots \\ y_{Rt} \end{bmatrix} = \rho \begin{bmatrix} 0 & w_{21} & w_{31} & \dots & w_{R1} \\ w_{12} & 0 & w_{32} & \dots & w_{R2} \\ w_{13} & w_{23} & 0 & \dots & w_{R3} \\ \dots & \dots & \dots & \dots & \dots \\ w_{1R} & w_{2R} & w_{3R} & \dots & 0 \end{bmatrix} \begin{bmatrix} y_{1t} \\ y_{2t} \\ y_{3t} \\ \dots \\ y_{Rt} \end{bmatrix} + \begin{bmatrix} x_{11t} & x_{21t} & x_{31t} & \dots & x_{k1t} \\ x_{12t} & x_{22t} & x_{32t} & \dots & x_{k2t} \\ x_{13t} & x_{23t} & x_{33t} & \dots & x_{k3t} \\ \dots & \dots & \dots & \dots & \dots \\ x_{1Rt} & x_{2Rt} & x_{3Rt} & \dots & x_{kRt} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ \beta_3 \\ \dots \\ \beta_k \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \\ u_{3t} \\ \dots \\ u_{Rt} \end{bmatrix} \quad (9)$$

Equation (9) contains $(k+1)$ parameters, the same as those that appear in the reduced form:

$$y = (I - \rho W)^{-1} x\beta + (I - \rho W)^{-1} u \quad (10)$$

Assuming that the number of cross-sectional observations, R , is greater than the number of parameters, $k+1$, the model is *overidentified*. The reduced form of (10) is nonlinear in parameters although it can be 'linearized' by using the expansion of the inverse matrix:

$$\begin{aligned} (I - \rho W)^{-1} &\approx I + \rho W + \rho^2 W^2 + \rho^3 W^3 + \dots + \dots \\ \Rightarrow y &= x\beta + x^{(1)}\beta_1 + x^{(2)}\beta_2 + x^{(3)}\beta_3 + \dots + v \\ x^{(j)} &= W^j x \\ \beta_j &= \beta \rho^j \\ v &= (I - \rho W)^{-1} u \end{aligned} \quad (11)$$

being v the error term of the reduced form, spatially autocorrelated and heteroskedastic. As said, it is simpler to work with the structural form of (10) than with the reduced form of (11); see Paelinck and Nijkamp, 1978, and Paelinck and Klaassen, 1979, for more details

Another point to note in relation to the problem of identification is that the topology of the space does not specially matter; that is, the composition of the weighting matrix (which it is supposed to reflect the structure of the space) does

not have any incidence on that question. The minimum requirement is that, at least, two regions should be connected.

3. SPATIAL GRANGER CAUSALITY

The notion of Granger causality depends to a great extent on the principle of ‘temporal succession’ which implies that the cause must precede, in a temporal sense, the effect; another well-established principle refers to the ‘contiguity relation’ between cause and effect (Pearl, 2000). That is, both elements must coincide in a specific time and location. However, a cross-section contains observations that occur simultaneously over time and are not necessarily adjacent. In fact the ‘allotopy’ principle, stated by Ancot *et al.* (1990) as ‘very often, the factors that explain a given economic fact in a region of space are located in distinct places’, deactivates the requirement of physical contiguity.

Then, the question is whether the notion of spatial dynamics may replace the concept of temporal dynamics in the analysis of causality over Space.

Briefly stated, Granger causality test (Granger, 1969) develops around the idea of *predictability* in the sense that, if variable x causes variable y , the past of the first variable must help to improve the forecasts of the second: $\sigma^2(y_{t+1} | I_t) < \sigma^2(y_{t+1} | I_t - x_t^*)$, where $\sigma^2(-)$ denotes uncertainty (i.e., variance),

I_t is the informative content up to period t and x_t^* the information of variable x up to period t . The test is well-known (Holly, 1984). The null hypothesis is non-causality and lead us to a set of zero restrictions:

$$y_t = \alpha_0 + \sum_{k=1}^{k_1} \alpha_{1k} y_{t-k} + \sum_{k=1}^{k_2} \alpha_{2k} x_{t-k} + u_{1t}$$

$$\left. \begin{array}{l} H_0 : \alpha_{21} = \dots = \alpha_{2k_2} = 0 \\ H_A : \text{No } H_0 \end{array} \right\} \rightarrow F = \frac{SR_0 - SR_A}{SR_A} \frac{T - (k_1 + k_2)}{k_2} \sim F(T - (k_1 + k_2); k_2) \quad (12)$$

being SR_0 and SR_A the sum of squared residuals of the model of the null and alternative hypothesis, respectively. If we change the terms past/future by spatial proximity/remoteness, we obtain:

$$y = \alpha_0 + \sum_{k=1}^{k_1} \alpha_{1k} W_k y + \sum_{k=0}^{k_2} \alpha_{2k} W_k x + u$$

$$\left. \begin{array}{l} H_0 : \alpha_{20} = \alpha_{21} = \dots = \alpha_{2k_2} = 0 \\ H_A : \text{No } H_0 \end{array} \right\} \rightarrow LR = 2(l_{H_A} - l_{H_0}) \sim \chi^2(k_2 + 1) \quad (13)$$

where $\{W_1, W_2, \dots, W_{kj}\}$ is a sequence of weighting matrices of order 1, 2, etc., with $W_0 = I$; y and x are $(R \times 1)$ vectors of observations of the two variables, in period t , and u a vector of error terms. The LR ratio compares the estimated log-likelihood under the alternative, l_{H_A} , and under the null hypothesis, l_{H_0} .

The null hypothesis implies that the information about the spatial distribution of the x variable does not help to improve our knowledge about the y variable. Obviously, in continuation we have to invert the order of the variables to test the hypothesis that y does not cause x .

This two-step procedure can be combined in a direct formulation specifying a spatial VAR (Di Giacinto, 2003, 2006):

$$\left. \begin{aligned} y &= \alpha_0 + \sum_{k=1}^{k_{11}} \alpha_{1k} W_k y + \sum_{k=0}^{k_{12}} \alpha_{2k} W_k x + u_y \\ x &= \beta_0 + \sum_{k=1}^{k_{21}} \beta_{1k} W_k x + \sum_{k=0}^{k_{22}} \beta_{2k} W_k y + u_x \end{aligned} \right\}$$

$$\begin{aligned} \bullet \text{ } x \text{ does not cause } y &\rightarrow H_0 : \alpha_{20} = \alpha_{21} = \dots = \alpha_{2k_{12}} = 0 \\ \bullet \text{ } y \text{ does not cause } x &\rightarrow H_0 : \beta_{20} = \beta_{21} = \dots = \beta_{2k_{22}} = 0 \\ \bullet \text{ } x \text{ does not cause } y \text{ AND } y \text{ does not cause } x &\rightarrow H_0 : \alpha_{20} = \alpha_{21} = \dots = \alpha_{2k_{12}} = \beta_{20} = \beta_{21} = \dots = \beta_{2k_{22}} = 0 \end{aligned} \quad (14)$$

$$\rightarrow LR = 2(l_{H_A} - l_{H_0}) \sim \chi^2(k_J)$$

being k_J the number of restrictions corresponding to each case. The bivariate system of (14) can be estimated by maximum-likelihood methods.

As a kind of exploratory, preliminary analysis, it is interesting to combine the Simultaneous Dynamic Least Squares (SDLS from now on) estimators of Paelinck (1990) with a very simple and popular statistic in applied econometrics such as the partial correlation coefficient. Let us introduce the case assuming a spatial model, like the following:

$$y = \rho W y + x \beta_0 + W x \beta_1 + u \quad (15)$$

The SDLS estimators of (15) are equivalent to the estimators of the reduced form in (10), as shown in Griffith and Paelinck (2011). This implies that, if x is a matrix of strongly exogenous variables, the endogenously generated \hat{y} vector has the same property. Both elements can be combined in a matrix $z = [x, Wx, \hat{y}]$. Then the incremental contributions of all variables are computed according to Theil (1971). Recall that R_i^2 is defined as the multiple correlation coefficient from the multiple regression of y (observed) on z minus the i^{th} column; it is further known that:

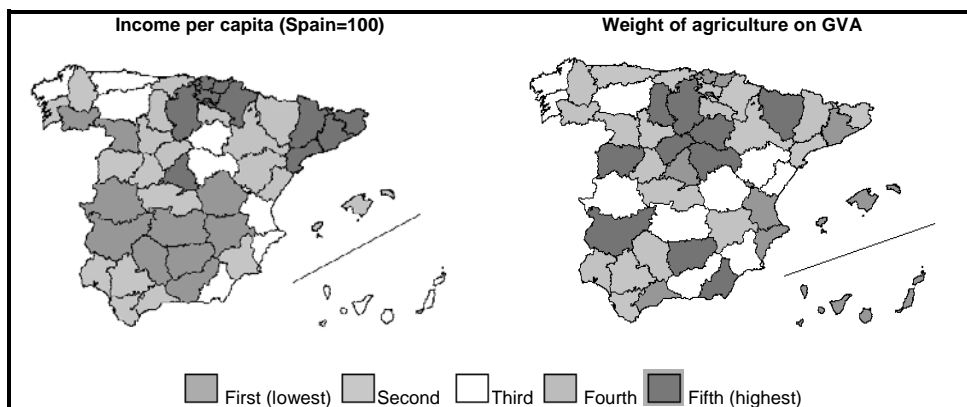
$$R^2 \approx \sum_{i=1}^k R_i^2 \quad (16)$$

being R^2 the global multiple correlation coefficient of the equation. Obviously, if the x variables are not relevant for explaining y over space, the marginal R^2 coefficient associated to these variables should be small in comparison with that attained from the lag structure of the y variable alone.

4. AN APPLICATION TO THE SPANISH CASE: PERSONAL INCOME VS AGRICULTURE

As an example, we present the case of the spatial distribution of the income per capita and the weight of the agricultural sector in the Gross Value Added of the Spanish provinces in the year 2006. The two variables appear in Figure 1. The first (*ipc*) is measured as an index, with value 100 for the national average, and the second (*ag*) as the percentage that the agricultural sector represents on the Gross Value Added of each province.

Figure 1
Spatial distribution of *ipc* and *ag*, in quantiles, in 2006



Source: Own elaboration.

The first question to note is that the spatial distribution of both variables is very different. The two are positively spatially correlated, but the structure of the second is more diffused (with a Moran's I of 0.15 and p value of 0.06) whereas the income per capita shows a strong Northeast-Southwest tendency (the Moran statistic is 0.66 with a p value of 0.00).

To begin with the discussion, let us introduce the equation connecting both variables:

$$y = \alpha_0 + \sum_{k=1}^2 \alpha_{1k} W_k y + \sum_{k=0}^2 \alpha_{2k} W_k x + u \quad (17)$$

y and x are our two variables of interest, ipc and ag . If y is identified with ipc and x with ag , we are testing for causality from agriculture (cause) to income (effect), and the contrary if we associate y with ag and ipc with x . In each case, we obtain the SDLS estimate of the corresponding y variable, \hat{y} , using a reduced version of (17), namely:

$$y = \alpha_0 + \alpha_{11}W_1y + \sum_{k=0}^2 \alpha_{2k}W_kx + u \quad (18)$$

This allows to complete the matrix $z = [W_1\hat{y}, W_2\hat{y}, x, W_1x, W_2x]$ and proceed as indicated in previous Section. Main results appear at the top of Table 1. Under the heading of ipc or ag , we show the cumulative percentage obtained from the corresponding marginal coefficients. For example, the SDLS regression for the ipc variable produces a multiple correlation coefficient of 0.77. The spatial structure of the variable ipc accounts for the 69.6% of this value, whereas the information of the ag variable only amounts to the 30.4%. In the case of the regression for the ag variable, the multiple coefficient is smaller, 0.53, and depends mainly on the information about the ipc variable, 75.1%, whereas the spatial distribution of the agricultural sector only accounts for the 24.9%. In short, it appears that income has a significant effect in the distribution of the agricultural sector but the reverse it is dubious.

Table 1
Income pc vs agriculture. Causality results

(1) SDLS+ (2) LS Estimation			
Explained	<i>ipc</i>		Explained <i>ag</i>
$W_1 ipc$	46.3%		$W_1 ag$ 19.4%
$W_2 ipc$	69.6%		$W_2 ag$ 24.9%
<i>ag</i>	15.3%		<i>ipc</i> 38.3%
$W_1 ag ag$	15.4%		$W_1 ipc$ 74.6%
$W_2 ag$	30.4%		$W_2 ipc$ 75.1%
Corr coef.	0.7667		Corr. coef. 0.5309
ML Estimation			
Explained	<i>ipc</i>		Explained <i>ag</i>
Log. unrest.	-202.193		Log. unrest. -135.068
Log. restri.	-195.785		Log. restri. -120.595
LR statistic	6.408		LR statistic 14.473
p-value	0.0934		p-value 0.0023

Source: Own elaboration.

At the bottom of Table 1, in the ML estimation panel, we present the results of the likelihood ratio of (14). The pvalue of the first relation (agriculture does not cause income) does not allow to reject the null hypothesis at the usual 5% significance level. In the second equation (income does not cause the distribution of agriculture) we obtain a very low pvalue for the LR statistics, which lead us to reject the null hypothesis.

5. CONCLUSIONS

This paper is a first approach to the analysis of causality in a spatial context. We are convinced that this is a very important topic that must be checked in order to assure the consistency of any econometric model in a spatial framework. There are obvious difficulties in solving the question such as the type of information used in this type of models. In this sense, the intention of our paper is to motivate the discussion.

We present some preliminary results in terms of an exploratory technique, based on the decomposition of the multiple correlation coefficient of a general regression, as well as a test which may be seen as an adjusted spatial version of the popular Granger causality test.

BIBLIOGRAPHY REFERENCES

- ANCOT, J.; KUIPER, J. and PAELINCK, J.H.P. (1990). "Five Principles of Spatial Econometric Illustrated". In CHATTERJY, M. and R. KUENNE: *Dynamics and Conflict in Regional Structural Change* (pp. 141-155). London: McMillan.
- ANSELIN L. (1988). *Spatial Econometrics: Methods and Models*. Dordrecht: Kluwer.
- Di GIACINTO, V. (2003). "Differential Regional Effects of Monetary Policy: A Geographical SVAR Approach". *International Regional Science Review*, 26, pp. 313-341.
- Di GIACINTO, V. (2006). "A Generalized Space-Time ARMA model with an Application to Regional Unemployment Analysis in Italy". *International Regional Science Review*, 29, pp. 159-198.
- GRANGER, C. (1969). "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods". *Econometrica*, 37, pp. 424-438.
- GRIFFITH, D. and J.H.P. PAELINCK (2011). *Non-standard spatial statistics and spatial econometrics*. Berlin: Springer.
- HOLLY, A. (1984). *Tests d'exogénéité*. I.M.E. Document de travail n° 71, Université de Dijon.
- PAELINCK, J.H.P. (1990). "Some new estimators in spatial econometrics". In D.A. GRIFFITH (ed.): *Spatial Statistics: Past, Present and Future* (pp. 163-181). Syracuse University, Institute of Mathematical Geography, Monograph n° 12.
- PAELINCK J.H.P. and P. NIJKAMP (1978). *Operational Theory and Method in Regional Economics*. Farnborough: Saxon House.

- PAELINCK J.H.P and L. KLAASSEN (1979). *Spatial Econometrics*. Farnborough: Saxon House.
- PEARL, J. (2000). *Causality: Models, Reasoning and Inference*. Cambridge: Cambridge University Press.
- THEIL, H. (1971). *Principles of Econometrics*. Amsterdam: North-Holland.

