

## Spatial Causality. An application to the Deforestation Process in Bolivia

Javier Aliaga \*, Marcos Herrera \*\*, Daniel Leguía \*, Jesús Mur \*\*,  
Manuel Ruiz \*\*\* Horacio Villegas \*

**ABSTRACT:** This paper analyses the causes of deforestation for a representative set of Bolivian municipalities. The literature on environmental economics insists on the importance of physical and social factors. We focus on the last group of variables. Our objective is to identify causal mechanisms between these factors of risk and the problem of deforestation. To this end, we present a testing strategy for spatial causality, based on a sequence of Lagrange Multipliers. The results that we obtain for the Bolivian case confirm only partially the traditional view of the problem of deforestation. Indeed, we only find unequivocal signs of causality in relation to the structure of property rights.

**JEL Classification:** C21, C50, R15.

**Keywords:** Risk of deforestation, Bolivia, municipalities, causality.

### Causalidad espacial. Una aplicación al proceso de deforestación en Bolivia

**RESUMEN:** Este trabajo analiza las causas de la deforestación para un conjunto representativo de municipios bolivianos. La literatura sobre economía ambiental insiste en la importancia de los factores físicos y sociales. Nos centramos en el último grupo de variables. Nuestro objetivo es identificar los mecanismos causales entre estos factores de riesgo y el problema de la deforestación. Con este fin, se presenta una estrategia de análisis para identificar mecanismos de causalidad espacial, basada en una secuencia de los multiplicadores de Lagrange. Los resultados que obtenemos para el caso de Bolivia confirman sólo parcialmente la visión tradicional del problema de la deforestación. De hecho, sólo encontramos

---

\* Institute of Socio-Economic Research (IISEC). Bolivian Catholic University.

\*\* Department of Economic Analysis. University of Zaragoza.

\*\*\* Department of Computing and Quantitative Methods. Technical University of Cartagena.

**Acknowledgements:** The authors would like to express their thanks to the Spanish International Cooperation Agency (*Agencia Española de Cooperación Internacional - AECI*) for its funding of project A/031125/10 as well as the support obtained through project ECO2009-10534 of the *Ministerio de Ciencia e Innovación del Reino de España*.

*Received: 13 april 2011 / Accepted: 28 november 2011.*

signos inequívocos de causalidad en relación con la estructura de los derechos de propiedad.

**Clasificación JEL:** C21, C50, R15.

**Palabras clave:** Riesgo de deforestación, Bolivia, municipios, causalidad.

## 1. Introduction

In the last few years, deforestation has become one of the hot topics on the research agenda of Economics and Environmental Economics. As it is generally acknowledged, damages to the environment have a great long run impact on the welfare conditions because of their effects on biodiversity reduction, natural resources depletion, climate change and soil degradation, among other factors (Kaimowitz and Angelsen, 1999). This is particularly relevant for Bolivia, where hundreds of thousands of hectares of rainforests and woodland are lost every year.

Different arguments have been used to explain this process. On the one hand, it is agreed that ploughing for agricultural purposes is in detriment of the woodland mass (Pacheco, 2004). At the same time, the improvement of transport infrastructures and demographic pressure increase the risk of deforestation. The lack of a well-defined property rights structure is another factor that facilitates the rainforest wasting. Some other physical or environmental factors also have a strong impact where deforestation takes place.

In this respect, Bolivia is a very interesting case as, approximately, 50% of the country is still grassland and rainforest. However, pressure for the transformation of the wilderness has increased significantly in the last few decades. The objective of our paper is to study the existence of causality mechanisms between the list of variables usually identified as factors of risk. The available deforestation indicators correspond for a set of 91 Bolivian municipalities that pertain to four departments, Beni, Pando and part of La Paz and Santa Cruz. These municipalities represent 60% of Bolivian territory and 40% of the population.

The peculiar aspect of our work is that we would like to go a bit further from the pure concept of dependence between risk factors and deforestation indices. In this sense, it must be remembered that a (spatial) econometric model relates a set of variables, trying to find their structure of dependence. However nothing is said in what respects to possible causality mechanisms between them. Causality is a central topic in mainstream econometrics that requires of a specific treatment but, surprisingly, this topic has had a very limited impact on the field of spatial econometrics using pure cross-sections (Weinhold and Nair, 2001, Hurlin and Venet, 2001, Hood *et al.*, 2008, or Tervo, 2009, for the case of spatial panel data). One of the purposes of this paper is to address the problem of spatial causality.

The paper is structured as follows: Section 2 provides a brief overview of the problem of deforestation and its consequences for the case of Bolivia. The Third section presents the problem of causality in a spatial cross-section and proposes some so-

lution. The Fourth section contains the results of the application of causality analysis to the data for the Bolivian municipalities. Main conclusions appear in Section 5.

## 2. Deforestation: the Bolivian case

In the last decades Bolivia has registered an exponential increase in deforestation. In the period of time between 1975 and 1993 a deforestation rate of 0.3% was produced, equivalent to the disappearance of 168,012 hectares of forests per year (Wachholtz, 2006). Between 1993 and 2000, the average increased to 270,000 hectares (Rojas *et al.*, 2003) and 280,000 hectares per year for the period 2004 to 2005.

In the previous data, we only consider the cases of deforestation that affected to a minimum of 5 hectares. Muñoz (2006) estimates that if the clearings of less than 5 hectares are taken into account, the number can easily reach half a million hectares per year. In per capita terms, a study lead by Andersen and Mamani (2009) found that the deforestation rates in Bolivia represent around 320 m<sup>2</sup>/person/year, which is 20 times greater than the world average (16 m<sup>2</sup>/person/year). This is one of the highest per capita deforestation rates in the world.

On the other part, according to the National Institute for Agrarian Reform (INRA, 2011), in 6 of the 8 Bolivian *ecoregions* (climatic systems with specific traits) more than 50% of its territory appears under the denomination of Communitarian Land of Origin. The areas where there are greater private ownership rights are the Integrated Central North, the Bosque Tucumano Boliviano, the Chaco and, in less scale, the Gran Chiquitania. The communitarian ownership rights are concentrated in the regions of the Amazon, the Bosque Tucumano Boliviano and the Chaco.

Finally, according to the United Nations Program for Development (PNUD, 2008), currently, the fringe that is suffering a greater pressure of deforestation is found between 142 masl<sup>1</sup> and 283 masl which explain the severe damages suffered in the departments of Pando, Beni and Santa Cruz. Deforestation has had, up to now, an smaller impact in the most elevated regions, as in the Humid Plateau of the Central Andes, in Yungas and in the Bosque Tucumano Boliviano. On the other hand, the pressure has been very intense in the regions of the Integrated Central North and the Chaco, due to their favorable conditions for agrarian and livestock production. The same process is beginning to occur now with the Amazon and the savannahs of Beni, although the typical seasonal floods of this zone slow the transformation of the forest into agrarian land (Lambin, 1997).

## 3. A procedure for testing spatial causality

Causality is one of the key issues in Economics to the extent that, for example, Heckman (2000) claims that «*the definition of causal parameters*» has been one of

---

<sup>1</sup> «Masl» means «meters above sea level».

the major contributions of Econometrics. There is a huge literature devoted to the topic where we can find different methods and approaches to the analysis of causality (Hoover, 2004). However, it is a bit surprising the little impact that this topic has had in a spatial context (we can only cite the work of Blommestein and Nijkamp, 1983). Recently, Herrera (2011) addressed the question of causality in a purely spatial context offering an updated perspective and new proposals. This section partly follows his suggestions.

Indeed, causality is not as simple as it seems, even in a time dimension: common causes, counterfactuals, non-experimentality, etc are problems that appear regularly in the literature. Difficulties increase in space where the first problem is to define the meaning of the term. For the sake of simplicity, let us think in the case of only two variables. We agree with the operational definition of Herrera (2011): *variable  $x$  causes variable  $y$ , in a spatial setting, if the first variable contains unique information in relation to the second variable, once we have taken into account all the information existing in the Space related to  $y$* . Consequently, we are going to use the term causality in information.

There are three points that need to be addressed when testing for causality between variables in pure spatial cross-sections, as it appears in Figure 1.

- i) The role of the Space: if the variables are spatially independent then it would be preferable to use a traditional approach to the problem (Heckman, 2000, or Pearl, 2009).
- ii) The relation between the variables: if the two variables were independent it would not make sense to talk about causality.
- iii) Assuming that Space is relevant and that the variables are related, causality in information implies that there is a one-way information flow between the two variables.

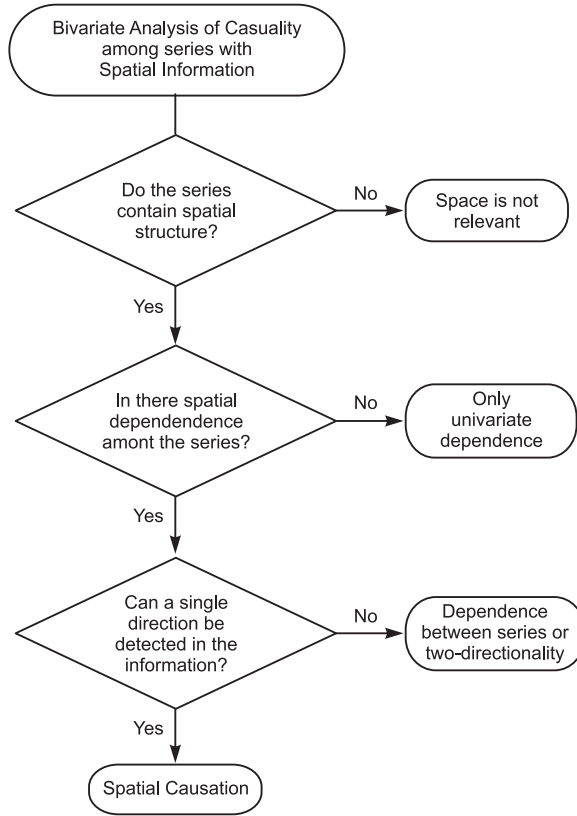
The first step means testing for the assumption of spatial independence of the data of each variable, for which a certain formalization of Space would be needed. In this sense, we follow the usual reasoning in terms using a finite sequence of weighting matrices, specified on a priori basis. Then, some of the well-known test of spatial dependence can be applied to each series (like the Moran's I, the Lagrange Multiplier, etc.). The results of this first step should be consistent: the same weighting matrix must intervene in the spatial structure of each series<sup>2</sup> and the hypothesis of spatial independence must be rejected for the two variables. In other words, Space should be relevant for the two variables and the spatial topology must coincide.

The second step, dependence between the variables, is a necessary condition to observe causality. In this case, we need a test of spatial dependence between the variables that takes into account the spatial structure of both series. The bivariate Moran's  $I_{yx}$  is a Mantel-type coefficient (Mantel, 1967), adapted by Wartenberg (1985) as an index to measure the spatial cross-correlation between two variables. Assuming that

---

<sup>2</sup> We mean that the same weighting matrix must be chosen as the *optimal spatial operator* (Herrera *et al.*, 2011) in order to account for the spatial dependence of each series.

**Figure 1.** Testing for causality between spatial series



the two variables are observed in  $R$  different locations, the expression of this statistic is as follows:

$$I_{xy} = \frac{\sum_{j=1}^R \sum_{i=1, i \neq j}^R y_i w_{ij} x_j}{S_0 \sqrt{\text{Var}(y) \text{Var}(x)}} \quad (1)$$

where  $w_{ij}$  is the  $(i, j)$ -th element of the weighting matrix  $W$  and  $S_0$  the sum of all the elements of  $W$ ;  $\text{Var}(y)$  and  $\text{Var}(x)$  refer to the (estimated) variance of the series  $y$  and  $x$ . The distribution function of the  $I_{yx}$  statistic is unknown.

Czaplewski and Reich (1993) obtain its moments,  $E(I_{yx})$  and  $V(I_{yx})$ , over all possible  $R!$  random permutations of the pairs  $\{y_s; x_s\}_{s \in S}$ , being  $S$  the set of locations whose cardinality is  $R$ . For moderate to large sample sizes (in any event,  $R > 40$ ), the authors

proposed the statistic:  $T_{xy} = \frac{I_{xy} - E(I_{yx})}{\sqrt{V(I_{yx})}}$ , which is distributed, approximately, like a

standard normal distribution. The null hypothesis of no correlation is rejected when  $|z_{xy}| > N_{\alpha/2}$ , where  $N_{\alpha/2}$  is the critical value corresponding to the standardized normal value that leaves a probability of  $\alpha/2$  on the right. Below we present another test for spatial independence, based on Lagrange Multipliers.

The purpose of the third step is to detect the direction of causality, if present, between the two variables. Following usual practice in time series analysis, we are going to specify an unrestricted spatial vector autoregressive model (SpVAR) to complete the testing strategy. Let us remind that the optimal weighting matrix,  $W$ , has been chosen before. For simplicity, we assume that the spatial dependence of both series is of the first order:

$$\begin{cases} [I_R - \rho_{yy}W]y + [\beta I_R + \rho_{yx}W]x + \eta_y = u_y \\ [\theta I_R + \rho_{xy}W]y + [I_R - \rho_{xx}W]x + \eta_x = u_x \end{cases} \quad (2)$$

where  $\{\rho_{yy}; \rho_{yx}; \rho_{xy}; \rho_{xx}\}$  are parameters of (crossed) spatial dependence,  $I_R$  is the identity of order  $R$ ,  $y$  and  $x$  ( $R \times 1$ ) vectors of observations of the variables of interest,  $\{\eta_y; \eta_x\}$  are two vectors of deterministic terms of order ( $R \times 1$ ) and  $\{u_y; u_x\}$  random vectors. More compact:

$$AY + \eta = u \quad (3)$$

where  $Y$  is a ( $2R \times 1$ ) vector such that  $Y' = [y'; x']$ . The  $\mu$  vector is also of order ( $2R \times 1$ ):  $\eta = [\eta_y; \eta_x]$ ; for simplicity, let us assume that the non-deterministic component of both series consist of only a constant, so  $\eta = m \otimes l$ , being  $l$  an ( $R \times 1$ ) vector and  $m$  a ( $2 \times 1$ ) vector of means  $[m_y; m_x]$ . The error vector is composed of two sub-vectors of order ( $R \times 1$ ):  $u' = [u_y'; u_x']$ , which is normally distributed  $u \sim N(0, \Xi)$ , where:

$$\Xi = \begin{bmatrix} \sigma_y^2 I_R & 0 \\ 0 & \sigma_x^2 I_R \end{bmatrix} = \begin{bmatrix} \sigma_y^2 & 0 \\ 0 & \sigma_x^2 \end{bmatrix} \otimes I_R = \Sigma \otimes I_R \quad (4)$$

Moreover,  $A$  is a ( $2R \times 2R$ ) matrix with the following structure:

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \rightarrow \begin{cases} A_{11} = I_R - \rho_{yy}W \\ A_{12} = \beta I_R + \rho_{yx}W \\ A_{21} = \theta I_R + \rho_{xy}W \\ A_{22} = I_R - \rho_{xx}W \end{cases} \quad (5)$$

Assuming normality, the log-likelihood function is:

$$L(Y; \Psi) = -\frac{R}{2} \ln(2\pi) - \frac{R}{2} \ln|\Sigma| + \ln|A| - \frac{[AY - \eta]' (\Sigma \otimes I_R)^{-1} [AY - \eta]}{2} \quad (6)$$

With  $\Psi' = [\rho_{yy}; \beta; \rho_{yx}; \eta_y; \sigma_y^2; \rho_{xx}; \theta; \rho_{xy}; \eta_x; \sigma_x^2]$ . The score vector is

$$l(\Psi) = \begin{bmatrix} \partial L / \partial \rho_{yy} \\ \partial L / \partial \beta \\ \partial L / \partial \rho_{yx} \\ \partial L / \partial \eta_y \\ \partial L / \partial \sigma_y^2 \\ \partial L / \partial \rho_{xx} \\ \partial L / \partial \theta \\ \partial L / \partial \rho_{xy} \\ \partial L / \partial \eta_x \\ \partial L / \partial \sigma_x^2 \end{bmatrix} = \begin{bmatrix} (1/\sigma_y^2)(y'W'u_y) - tra_{11} W \\ -(1/\sigma_y^2)(x'u_y) + tra_{21} \\ -(1/\sigma_y^2)(x'W'u_y) + tra_{21} W \\ -(1/\sigma_y^2)(l'u_y) \\ -(R/2\sigma_y^2) + (u'_y u'_y / 2\sigma_y^4) \\ (1/\sigma_x^2)(x'W'u_x) - tra_{22} W \\ -(1/\sigma_x^2)(y'u_x) + tra_{12} \\ -(1/\sigma_x^2)(y'W'u_x) + tra_{12} W \\ -(1/\sigma_x^2)(l'u_x) \\ -(R/2\sigma_x^2) + (u'_x u'_x / 2\sigma_x^4) \end{bmatrix} \quad (7)$$

where  $tr(-)$  is the trace operator,  $a_{11} = [A_{11} - A_{12}A_{22}^{-1}A_{21}]^{-1}$ ,  $a_{12} = -a_{11}A_{12}A_{22}^{-1}$ ,  $a_{21} = -A_{22}^{-1}A_{21}a_{11}$  and  $a_{22} = A_{22}^{-1} + A_{22}^{-1}A_{21}a_{11}A_{12}A_{22}^{-1}$ .

Using the framework of the SpVAR of (2), we can test: (1) independence between the series and (2) direction of causality (in information) between the series. Independence between the two series corresponds to the following null hypothesis:

$$\left. \begin{aligned} H_0: & A_{12} = A_{21} = 0 \\ H_1: & A_{12} \vee A_{21} \neq 0 \end{aligned} \right\} \quad (8)$$

The score vector (reordered according to the parameters in the null hypothesis) evaluated under the same null hypothesis of (8) becomes:

$$l(\Psi)_{|H_0} = \begin{bmatrix} \partial L / \partial \rho_{yx} \\ \partial L / \partial \beta \\ \partial L / \partial \rho_{xy} \\ \partial L / \partial \theta \\ \partial L / \partial \rho_{yy} \\ \partial L / \partial \eta_y \\ \partial L / \partial \sigma_y^2 \\ \partial L / \partial \rho_{xx} \\ \partial L / \partial \eta_x \\ \partial L / \partial \sigma_x^2 \end{bmatrix}_{|H_0} = - \begin{bmatrix} (1/\sigma_y^2)(x'W'u_y) \\ (1/\sigma_y^2)(x'u_y) \\ (1/\sigma_x^2)(y'W'u_x) \\ (1/\sigma_x^2)(y'u_x) \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \lambda_0 \\ \lambda_1 \end{bmatrix} \quad (9)$$

Being  $\lambda'_0 = -\left[ \frac{x'Wu_y}{\sigma_y^2} \frac{x'u_y}{\sigma_y^2} \frac{y'Wu_x}{\sigma_x^2} \frac{y'u_x}{\sigma_x^2} \right]'$  and  $\lambda'_1 = -[0 \ 0 \ 0 \ 0 \ 0 \ 0]'$ .

The Lagrange Multiplier is the quadratic form of the score vector on the inverse of the information matrix, both (vector and matrix) should be evaluated under the null hypothesis of (8). Combining these results, we obtain the expression of the Multiplier that enables us to test independence between the two series:

$$LM_I = \lambda'_0 I^{11} \lambda_0 \sim \chi^2(4) \tag{10}$$

where  $I^{11}$  is the inverse of the variance-covariance matrix of vector  $\lambda_0$ , whose expression can be found in equation (3.4.85) of Herrera (2011)<sup>3</sup>. Therefore, to test the assumption of no correlation:

$$H_0 : \{x_s\}_{s \in S} \text{ and } \{y_s\}_{s \in S} \text{ are uncorrelated processes}$$

The decision rule for the  $LM_I$  test with a confidence level of  $100(1-\alpha)\%$  is:

- If  $0 \leq LM_I \leq \chi^2_\alpha(4)$  the null hypothesis of (8) cannot be rejected.
- If  $LM_I > \chi^2_\alpha(4)$  reject the null hypothesis of (8).

Assuming that the null hypothesis of independence in the bivariate system of (2) has been rejected, the next step refers to the non-causality hypothesis. This is a double-lap exam: first we test that one variable, let us say  $x$ , does not cause in information the other,  $y$ ; then we change the order, testing that  $y$  does not cause, in information, to  $x$ . The null hypothesis of the first combination ( $x$  does not cause  $y$ ) is:

$$\left. \begin{aligned} H_0 : A_{12} &= 0 \\ H_1 : A_{12} &\neq 0 \end{aligned} \right\} \tag{11}$$

The score vector, evaluated under the null hypothesis of (11), is:

---

<sup>3</sup> Briefly  $I^{11}$  is a sub-matrix of the information matrix of the bivariate system of (2). As it is usual with the Lagrange Multipliers, the information matrix should be evaluated under the null hypothesis, in this case, of (8).



$$l(\Psi)_{H_0} = \begin{bmatrix} \frac{\partial L}{\partial \rho_{yx}} \\ \frac{\partial L}{\partial \beta} \\ \frac{\partial L}{\partial \rho_{xy}} \\ \frac{\partial L}{\partial \theta} \\ \frac{\partial L}{\partial \rho_{yy}} \\ \frac{\partial L}{\partial \eta_y} \\ \frac{\partial L}{\partial \sigma_y^2} \\ \frac{\partial L}{\partial \rho_{xx}} \\ \frac{\partial L}{\partial \eta_x} \\ \frac{\partial L}{\partial \sigma_x^2} \end{bmatrix}_{H_0} = - \begin{bmatrix} (1/\sigma_y^2)(x'Wu_y) + trA_{11}^{-1}A_{21}A_{22}^{-1}W \\ (1/\sigma_y^2)(x'u_y) + trA_{11}^{-1}A_{21}A_{22}^{-1} \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} \gamma_0 \\ \gamma_1 \end{bmatrix} \quad (12)$$

Where:

$$\gamma_0 = - \left[ \frac{x'W'u_y}{\sigma_y^2} + trA_{11}^{-1}A_{21}A_{22}^{-1}W \quad \frac{x'u_y}{\sigma_y^2} + trA_{11}^{-1}A_{21}A_{22}^{-1} \right]$$

$$\gamma_1 = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]'$$

In a more compact notation, the Lagrange Multiplier obtained for the null hypothesis of (11) is:

$$LM_{NC} = \lambda_0' I^{11} \lambda_0 \sim \chi^2(2) \quad (13)$$

Once again,  $I^{11}$  is the inverse of the variance-covariance matrix of the  $\gamma_0$  vector (expression (3.4.127) of Herrera, 2011). Consequently, to test the null hypothesis of:

$$H_0 : \{x_s\}_{s \in S} \text{ does not cause } \{y_s\}_{s \in S} \quad (14)$$

The decision rule for the  $LM_{NC}$  test with a confidence level of  $100(1-\alpha)\%$  is:

- If  $0 \leq LM_{NC} \leq \chi_\alpha^2(2)$  the null hypothesis of (11) cannot be rejected.
- If  $LM_{NC} > \chi_\alpha^2(2)$  reject the null hypothesis of (11).

#### 4. Deforestation in the Bolivian municipalities. A spatial approach

In this section, we apply the techniques developed previously to the information available on deforestation for a set of 91 Bolivian municipalities in the period 2004-2007. These municipalities belong to the departments of Bendi, 19 of them, Pando, 16, 23 come from the department of La Paz and 34 from Santa Cruz. They are

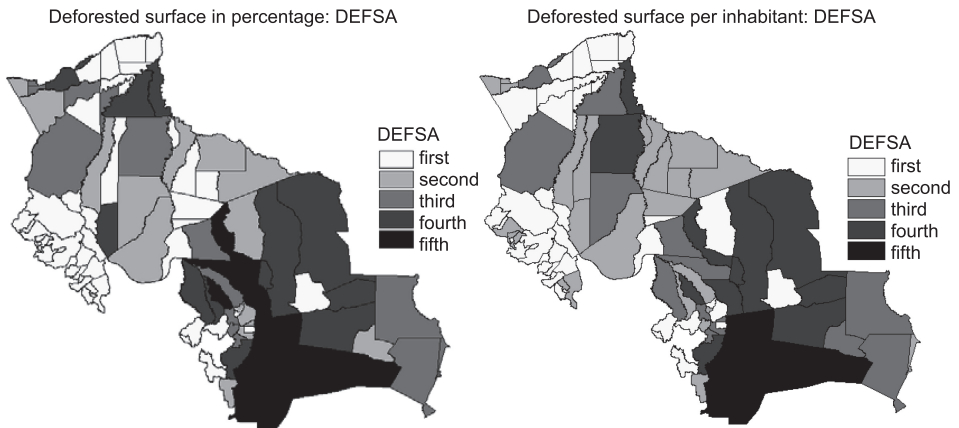
selected according to data availability (we could not find information for the other 222 Bolivian municipalities). Figure 2 shows the spatial layout of the municipalities included and not included in the study.

**Figure 2.** Bolivian municipalities in the deforestation study



Figure 3 depicts the indices of deforestation for these municipalities, using the quantiles of the distribution frequencies. The variables represented are the percentage of the land surface of each municipality classified as deforested in 2007 according

**Figure 3.** Deforestation indices in the Bolivian municipalities



to Rojas *et al* (2003), variable DEFSA, and the number of deforested hectares per inhabitant in the respective municipality, variable DEFPA. The spatial distribution of these data is what we are trying to explain.

There is an overall consensus in relation to the factors that are inducing the deforestation process (Kaimowitz and Angelsen, 1998). Some of them pertain to the block of physical environmental characteristics like rainfalls, temperature, climate instability, etc. However, we are interested in the impact of *human factors* in the sense that they reflect the consequences of social decisions in relation to economic growth, social organization, property rights, etc.

Due to statistical restrictions, we only have information for a limited number of risk factors: accessibility, measured in terms of density of principal and secondary roads per square kilometre (variable DECAT), population pressure, measured by the population density per Km<sup>2</sup> (variable DEPOB), urbanization, estimated by means of the percentage of population settled in rural areas, variable DEPOR, and property rights, as percentage of the municipality land surface privately owned, variable PROPI. The spatial distribution of the four variables, once again in quantiles, is shown in Figure 4.

**Figure 4.** Deforestation in the Bolivian municipalities. Risk factors

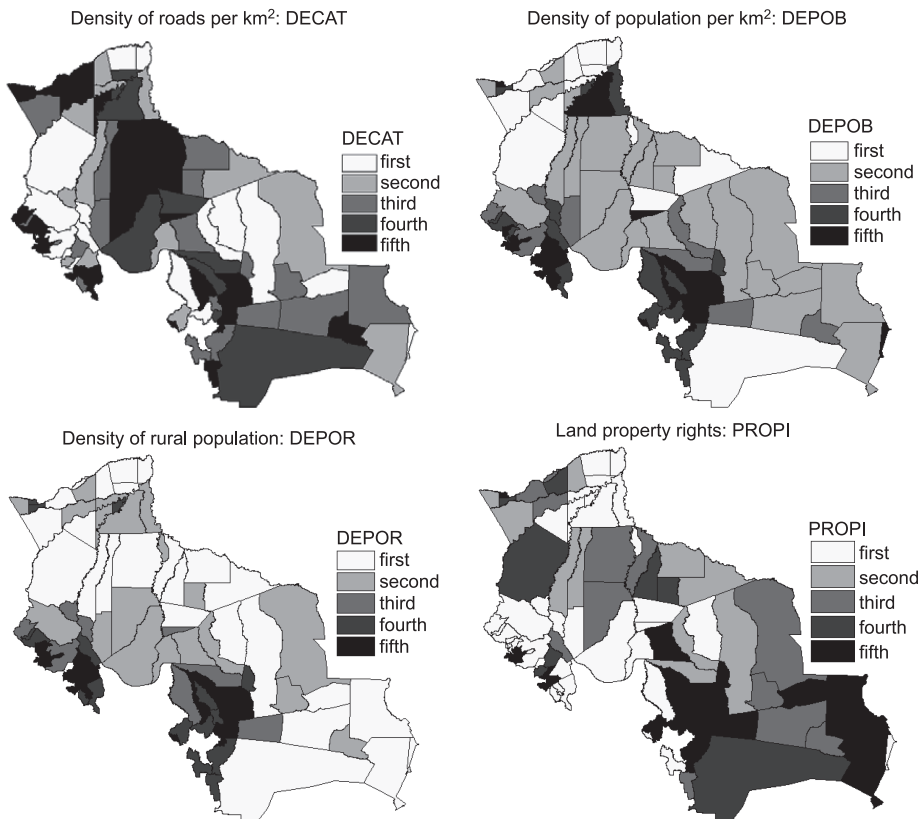


Table 1 presents some data for this group of variables. One important question to note is the great heterogeneity of the municipal records. For example, the average percentage of land classified as deforested is 2.03% for the 91 municipalities, but the figures range from 0.02% to 14.1%. The set of municipalities include cases with a very low density, 0.14 inhabitants per square kilometre, and others densely populated, with 1,175 inhabitants per square kilometre. The disparities in other indices such as road density or property rights structure are even greater. Moreover, all the variables are affected by severe non-normality problems. This is an important issue here because the strategy designed involves the use of maximum likelihood estimators, in which the assumption of normality plays a crucial role. Therefore in the following, we use the data in logarithms (transformed variables are identified with the symbol  $l$  before the respective code; the problems with the assumption of normality are corrected).

**Table 1.** Deforestation indices: Descriptive statistics

<i>Variables</i>	$\mu$	$\eta$	<i>Min.</i>	<i>Max.</i>	$\sigma$	$\alpha$	$\kappa$	<i>SW</i>	<i>I</i>
DEFSA	2.03	0.58	0.02	14.1	3.23	2.17	4.21	0.64 (0.00)	0.58 (0.00)
DEFFA	0.69	0.25	0.00	10.5	1.33	5.06	32.29	0.49 (0.00)	0.15 (0.01)
DEPOB	25.60	3.87	0.14	1,175.0	126.65	8.46	73.45	0.17 (0.00)	0.04 (0.04)
DEPOR	5.88	1.99	0.09	41.3	8.73	2.20	4.71	0.68 (0.00)	0.50 (0.00)
DECAT	98.75	82.66	0.00	330.7	59.81	-1.42	2.61	0.90 (0.00)	0.29 (0.00)
PROPI	30.44	6.18	0.00	100.0	39.75	-0.94	-0.87	0.71 (0.00)	0.10 (0.03)

$\mu$ : Mean;  $\eta$ : Median;  $\sigma$ : Standard deviation; *Min.*: Minimum value; *Max.*: Maximum value;  $\alpha$ : Skewness,  $\kappa$ : Kurtosis; *SW*: Shapiro-Wilks statistic; *MI*: Moran's I statistic. In parenthesis, *p*value.

Furthermore, as shown by the Moran index, there is a strong positive spatial dependence structure in the data of deforestation. This test of spatial independence is highly significant in all the cases. The weighting matrix employed to solve the test corresponds to the row-standardized version of the four nearest-neighbours (the conclusion of dependence is robust to the specification of the  $W$  matrix and, also, to the log transformation). The same matrix has been used in the causality analysis that follows.

As can be seen in Table 2, the linear correlation between the six indices of deforestation is medium to low. Except for two cases, DEFFA-DEFSA and LDEPOR-LDEPOB, the correlation coefficients are smaller than 0.5 in absolute value, although mostly of them are statistically significant (13 of the 15, at the usual significance level of 5%), with a predominance of positive scores (12 out of 15).

**Table 2.** Deforestation indices: Correlation matrix

	<i>DEFPA</i>	<i>LDEPOB</i>	<i>LDEPOR</i>	<i>LDECAT</i>	<i>LPROPI</i>
DEFSA	0.509	0.377	0.275	0.378	0.463
DEFPA		-0.093	-0.581	-0.071	0.051
LDEPOB			0.881	0.431	0.362
LDEPOR				0.421	0.376
LDECAT					0.314

95% confidence interval: (-0.21; 0.21).

All these data confirm, as expected, the relevance of the spatial dimension in the problem of deforestation. The role of the Space appears even more important when we consider bivariate spatial relationships. Table 3 shows the results of the bivariate Moran’s test,  $I_{yx}$ , and the Lagrange Multiplier,  $LM_I$ , for the assumption of spatial independence between the deforestation indicators and the risk variables. The results of the  $LM_I$  test in relation to the percentage of land surface deforested, DEFPA, are clearly anomalous. According to the simulations reported by Herrera (2011), the Lagrange Multiplier is more sensitive to the presence of outliers. The log-transformation is a smoothing transformation, useful for correcting non-normality problems, but probably not enough for the case of the Multiplier.

**Table 3.** Measures of bivariate spatial dependence

	$LM_I$		$I_{yx}$	
	<i>LDEFPA</i>	<i>LDEFSA</i>	<i>LDEFPA</i>	<i>LDEFSA</i>
LDEPOB	0.11 (0.99)	5.84 (0.21)	-3.15 (0.00)	2.98 (0.01)
LDEPOR	0.57 (0.97)	11.37 (0.00)	9.22 (0.00)	13.52 (0.00)
LDECAT	6.09 (0.19)	209.1 (0.00)	-7.24 (0.00)	2.02 (0.02)
LPROPI	14.81 (0.01)	22.93 (0.00)	16.96 (0.00)	3.46 (0.00)

*p*value in parenthesis.

Table 4 presents the results of the final step in our discussion of spatial causality. These results correspond to the Lagrange Multipliers of expression (13),  $LM_{NC}$ , whose null hypothesis is non-causality (in information). As indicated in Section 3, the results of this test are only relevant in the case that, previously, the assumption of spatial independence between series has been rejected. Furthermore, the identification of a certain direction of causality, in information, between the variables is subjected to the simultaneous fulfilment of two clauses: the null hypothesis of non-causality should be rejected in one direction but not rejected in the opposite direction.

Table 4 shows that, in relation to the indicator of per capita deforestation, LDEFPA, the test is non-conclusive in two cases. The population pressure, LDEPOB, and

**Table 4.** Causality results. Lagrange Multipliers,  $LM_{NC}$ 

	LDEFPA		LDEFSA	
	→	←	→	←
LDEPOB	0.87 (0.64)	0.90 (0.64)	2.99 (0.22)	33.83 (0.00)
LDEPOR	8.31 (0.02)	5.74 (0.05)	11.60 (0.00)	262.46 (0.00)
LDECAT	2.75 (0.25)	0.03 (0.99)	14.10 (0.00)	68.36 (0.00)
LPROPI	12.61 (0.00)	5.41 (0.07)	13.82 (0.00)	5.73 (0.06)
<i>p</i> value in parenthesis	→ : Causality is from the variable on the left to the variable on the right.		← : Causality is from the variable on the right to the variable on the left.	

the indicator of accessibility, LDECAT, do not cause deforestation, whereas urbanization, LDEPOR, and the structure of property rights, LPROPI, do cause this variable. The null of non-causality from LDEFPA to each of the four risks of deforestation cannot be rejected in any case, at a 5% level of significance (the conclusion is very tight with respect to LDEPOR and LPROPI).

Rejections of the null tend to predominate in the case of the percentage of land surface deforested in each municipality, LDEFSA. This is the case of LDEPOR and LDECAT, where the null of non-causality is rejected in both directions. According to the framework of Section 3, we cannot identify a unique direction for the information flow which prevents us of using the term causality (in information). On the contrary, the density of population, LDEPOB, appears to be caused by the deforestation process. Once again, property rights, LPROPI, emerge as a cause factor in the problem of deforestation

## 5. Final conclusions

Deforestation is an issue of great interest, particularly in regions that have preserved their environmental diversity; this is the case for most of South America in general, and Bolivia in particular. The literature on deforestation insists on the importance of physical variables related, for example, to climate and territory and other variables associated to social effects; human settlements, road infrastructure, and property rights are regularly identified as deforestation risk factors.

Our analysis has focused on the statistical part of the relationship, ignoring other aspects of the discussion. The problem that we consider is whether it is possible to detect causality relationships, in information, with a single cross-section of data and no time perspective. In this case, we wonder what occurs with the deforestation data available for a representative group of Bolivian municipalities.

The answer is positive to the first question: it is possible to develop a method for testing causality using purely spatial data. The strategy that we proposed is based on a sequence of Lagrange Multipliers obtained from a spatial VAR system. The applica-

tion of this strategy to the data available confirms only part of the traditional approach to the deforestation problem. Our conclusions can be summarised as follows:

- The variable with the greatest causal impact on the deforestation problem is the structure of property rights.
- Deforestation is found to be the cause of population distribution, as measured by population density.
- Other variables such as accessibility, measured through road density, or the importance of rural settlements do not appear to have a precise causal effect on deforestation.

It is important to note that the above results do not define the type of impact of the causal variables on the deforestation indices. The evidence available enables us to say, for instance, that an increase in private or communal land tenure will have a causal impact in the deforestation process. The same can be said of the relationship between deforestation and population density. The quantification of these relationships, in the sense of being able of forecasting tendencies, is in this project's future research agenda.

## References

- Andersen, L., and Mamani, R. (2009): «Cambio Climático en Bolivia hasta 2100: Síntesis de Costos y Oportunidades», *Estudio Regional de Economía del Cambio Climático en Sudamérica*, Comisión Económica para América Latina y el Caribe (CEPAL).
- Blommestein, H., and Nijkamp, P. (1983): «Causality Analysis in Soft Spatial Econometric Models», *Papers of the Regional Science Association*, 51, 65-77.
- Czaplewski, R., and Reich, R. (1993): «Expected value and variance of Moran's bivariate spatial autocorrelation statistic under permutation», *USDA For. Serv. Res. Paper RM-309*, Fort Collins, Colorado, pp. 1-13.
- Heckman, J. (2000): «Causal Parameters And Policy Analysis In Economics: A Twentieth Century Retrospective», *The Quarterly Journal of Economics*, 115 45-97.
- Herrera, M. (2011): «Causality. Contributions to Spatial Econometrics», *PhD. Dissertation*, University of Zaragoza.
- Herrera, M.; Mur, J., and Ruiz, M. (2011): «A Study on the Criteria for Selecting the Most Adequate Spatial Weighting Matrix», *Paper presented at the V World Congress of the Spatial Econometrics Association*, Toulouse, July.
- Hood III, M.; Kidd, Q., and Morris, I. (2008): «Two Sides of the Same Coin? Employing Granger Causality Tests in a Time Series Cross-Section Framework», *Political Analysis*, 16: 324-344.
- Hoover, K. (2004): «Lost causes», *Journal of the History of Economic Thought*, 26: 149-164.
- Hurlin, C., and Venet, B. (2001): *Granger Causality Tests in Panel Data Models with Fixed Coefficients*, University of Paris Dauphine: Working Paper Eurisco, 2001-2009.
- INRA (2011): *Plan Trienal de Ajuste al Nuevo Estado Plurinacional 2010-2012*, La Paz: Ministerio de Desarrollo Rural y Tierras.
- Kaimowitz, D., and Angelsen, A. (1998): *Economic Models of Tropical Deforestation: A review*, Center for Internacional Forestry Research (CIFOR).
- Lambin, E. (1997): «Modelling and monitoring land-cover change processes in tropical regions», *Progress in Physical Geography*, 21: 375-393.

- Mantel, N. (1967): «The Detection of Disease Clustering and a Generalized Regression Approach», *Cancer Research*, 27: 209-220.
- Muñoz, A. (2006): *Apuntes de la Audiencia Técnica-Avances de la Deforestación en Bolivia*, Santa Cruz, Superintendencia Forestal, Proyecto Bolfor II.
- Pacheco, P. (2004): *Las fronteras agrícolas en el trópico boliviano: entre las situaciones heredadas y los desafíos del presente*, Bogor Barat: Center for International Forestry Research (CIFOR).
- Pearl, J. (2009): «Causal Inference in Statistics: An Overview», *Statistics Surveys*, 3: 96-146.
- PNUD (2008): *Más allá del impase entre desarrollo y conservación. En la otra frontera: usos alternativos de recursos naturales en Bolivia*, Programa de las Naciones Unidas para el Desarrollo, PNUD, La Paz.
- Rojas, D.; Martínez, I.; Cordero, W., and Contreras, F. (2003): *Tasa de deforestación de Bolivia: 1993-2000*, Santa Cruz, Superintendencia Forestal, Proyecto Bolfor I.
- Tervo, H. (2009): «Centres and Peripheries in Finland: Granger Causality Tests Using Panel Data», *Spatial Economic Analysis*, 4: 377-390.
- Wachholtz, R. (2006): *Avance de la deforestación mecanizada en Bolivia: Tasa anual de deforestación mecanizada en los años 2004 y 2005*, Santa Cruz, Superintendencia Forestal, Proyecto Bolfor II.
- Wartenberg, D. (1985): «Multivariate spatial correlation: a method for exploratory geographical analysis», *Geographical Analysis*, 17: 263-283.
- Weinhold, D., and Nair, U. (2001): «Causality Tests for Cross Country Panels: A New Look at FDI and Economic Growth in Less Developed Countries», *Oxford Bulletin of Economics and Statistics*, 63: 153-172.