Determinants of spatial diffusion and adoption of European agri-environmental supports related to extensive grazing in France

Gilles ALLAIRE

INRA-ODR
INRA Toulouse, BP 52627, 31326 Auzeville, France
allaire@toulouse.inra.fr

Eric CAHUZAC

INRA-ODR
INRA Toulouse, BP 52627, 31326 Auzeville, France
eric.cahuzac@toulouse.inra.fr

and

Michel SIMIONI

INRA, UMR 1110 MOISA
Inra-SupAgro, 2, place Pierre Viala - Bât. 26, 34 060 Montpellier Cedex 2
michel.simioni@supagro.inra.fr

Determinants of spatial diffusion and adoption of European agri-environmental support for extensive grazing in France

Abstract

This paper analyzes farmers' participation in agri-environmental programs by integrating the spatial diversity of both extrinsic and intrinsic factors and their interactions. Special attention is devoted to the assessment of the spatial diffusion of agri-environmental programs related to extensive grassland management programs in two French regions with very different breeding practices. The empirical analysis involves the application of spatial Durbin probit models to data on the behavior of farmers in these two regions. Prediction tests show that the estimated models succeed in replicating the spatial distribution of observed participation rates for most of the considered geographical aggregation scales. Results for the marginal effects of the variables describing local farming conditions illustrate the importance of considering the spatial dimension in an analysis of farmers' participation in environmental schemes.

Keywords: Agri-environmental measures, Rural Development Program, Common Agricultural Policy, farmer's participation, spatial Durbin model, Probit, geographical diffusion.

JEL classification : C11, C21, Q01, Q18.

1 Introduction

Under the European rural development regulation policy, each EU member state is required to set up agri-environmental measures (AEMs) in its Rural Development Plan (RDP). Farmers can choose to engage in designed environmental contracts with commitments that exceed the relevant mandatory standards. Farmers who choose this option sign a contract with the administration and receive a payment for the cost of implementing environmental commitments and for any loss of income the commitments entail. AEMs may be designed by national or regional bodies, and requirements and payments can be adapted at the local level (e.g., natural resources preservation areas) such that the AEMs can be adapted to particular farming systems and local environmental conditions (e.g., mountains and less favored areas). Participating farmers provide environmental services at a given price in a given place. These services are intended to provide a public good. This public good is spatially structured in its design, delivery and benefits.

Knowledge of the factors that have led farmers to participate in agri-environmental programs provides crucial information for policy evaluation and for further policy design. In recent years, research on these determinants has received increased attention due to the growing EU share of expenditures on AEMs and the resulting need for feedback data. Defrancesco *et al.* (2008) emphasize that there is a current consensus in the literature that participation in voluntary programs, such as agri-environmental programs, depends on farmers' attitudes and behavioral responses, defined as extrinsic factors (see Vanslembrouk *et al.*, 2002) and how well the AEMs are suited to farming systems and economic contexts, defined as the intrinsic characteristics of the measures. Various papers using surveys to elicit farmers' behavioral attributes show that the level or type of professional degrees or training of farmers and their civic (non-economic) preferences, which some authors call environmental sensibility, act as determinant extrinsic factors in farmers' participation in AEMs (see, among others, Barreiro-Hurlé *et al.*, 2010). Intrinsic characteristics of AEMs include zoning and targeting practices. Local public authorities and lobbies participate in the design of AEMs at several geographical levels. Several authors emphasize the influence of the institutional characteristics of territories on the local management of environmental programs (references can be found in Peerlings and Polman, 2009).These characteristics impact both intrinsic and extrinsic factors. Considering the intensity of the adoption of AEMs at the EU regional level during the period from 2001-2004 from a political economy approach, Bertoni and Olper (2008) find a determinant role of "variable proxies related to farmer political weight, political institutions and the demand for positive externalities" in local implementation of AEMs. This indicates that social networks that farmers are a part of influence their decision to participate. These social networks are geographically structured, and, for a given location, they provide a collective capacity for farmers' groups to participate in a given program, which supports individual decisions. Generally, this collective capacity corresponds to the political weight of professional agricultural groups on local governments and territory planning initiatives, and we can then expect that the interactions between intrinsic and extrinsic factors related to the diffusion of AEMs are spatially differentiated. Institutional factors contributing to individual decisions are broadly defined as social capital in Bertoni *et al.* (2008).

This paper analyzes farmers' participation in agri-environmental programs by integrating the spatial diversity of both extrinsic and intrinsic factors and other interactions. We focus on one category of AEMs, those implemented in France that support sustainable extensive grazing through several schemes, or agri-environmental schemes (AESs). We focus specifically on farmers' decisions on whether to participate in one of the AESs without distinguishing between the AESs. In this sense, our paper differs from most other works on this subject because most papers focus on one AES and study the determinants of participating in this measure. Here, we are interested in the implementation of one of the objectives of the AEM (named "f") from the 1999 RDP that sought to prevent the suppression of permanent or semi-permanent pastures resulting from the intensification of cattle breeding practices. This objective was implemented by different, complementary programs. The main contribution of this paper is to assess the success of this objective, given the spatial context in which farmers made their decisions. It is difficult to determine the exact terms of the various contracts for each farmer, as the terms, in most cases, have been adapted to the geographical context of the farm. The modeling of farmers' decisions must then take into account that we cannot observe the characteristics of these spatially heterogeneous contracts. One solution would be to introduce fixed effects for these sub-regions in a standard probit model to capture this spatial heterogeneity when modeling choices.[1] We choose another approach based on recent work in spatial econometrics that allows us to consider omitted variables that are spatially correlated, such as contract characteristics, in modeling farmers' decisions. More specifically, our modeling approach of a farmer's AES adoption decision relies on the use of a spatial version of the probit model, the spatial Durbin probit model (SDPM). The advantages of this modeling choice are the following. First, this model allows us to implicitly consider the spatial dependence on the adoption decisions of farmers in the same neighborhood (spatial spillovers) and to assess the strength of spatial dependence on farmers' decision making. This assessment is of peculiar interest when contextual factors are added to farm and farmer characteristics as adoption determinants. The relevance of contextual factors in explaining the spatial dimension involved in farmers' decision making (e.g., importance of local farming systems and of local institutions) can then be assessed. Second, when spatial dependence is present, the effect of each adoption determinant can be decomposed into a direct effect of the decision of the farmer and an indirect effect due to the responses of neighboring farms through spillover effects. Third, the inability to observe all potential determinants of an adoption decision in our database can result in an endogeneity bias due to the possible correlations between observed factors and omitted (non-observable) factors that play a role in the farmer's decision (for example, the farmer's preference for nature conservation). In line with the findings of Pace and Lesage (2010) on the case of the linear model, this omitted variable bias can be mitigated by the use of SDPM. Lastly, the predictive ability of the estimated SPDM at different geographic aggregation scales. The objective is to determine whether this model capture the spatial diffusion of the studied AESs well. This assessment is based on the implementation of prediction tests (Ben-Akiva and Lerman, 1985) that compare predicted and observed participation rates on the same geographic scale.

Like most of the AESs in France, their environmental objectives and part of their funding are defined at the regional level. Conditions of eligibility and payments vary depending on decisions made regionally and according to various measures taken at the time the Rural Development Plan (RDP) was developed. In addition, farmers' unions or lobbies organized by the type of agricultural production (here, cattle breeding) have significant influence on the implementation of agricultural policies at the department (NUTS3) level.[2] Finally, the economy of cattle breeding is influenced by regional and local characteristics for both markets and production systems. As emphasized by Bertoni and Olper (2008), a national focus may mask several key details that could be important for the study of the adoption and spatial diffusion of these AEMs. Thereafter, the analysis is conducted at the regional level, focusing on two different French regions (NUTS2): Basse-Normandie and Auvergne. The first region is known for milk production, mainly from pastures (particularly wetlands), and the second region is known for milk production from mountainous pastures, with intensive agriculture in the valleys. These two regions have benefitted differently from the diversity of the French AESs supporting sustainable extensive grazing, which were successively offered to farmers during the 2000-2006 RDP period.

The remainder of the paper proceeds as follows: Section 2 presents the considered AESs and discusses them with respect to their spatial targeting. Section 3 presents the spatial Durbin probit model. Section 4 describes the data. In Section 5, the estimation results are presented and discussed. Section 6 provides the conclusion.

2 Background

Many EU Member States have basic AESs designed to attract large numbers of farmers and to address regional issues. In addition, these AES include more focused programs aimed at fewer farmers to engage in large restructuring (e.g., conversion to organic farming) or intended address localized issues, such as the Natura 2000 sites.[3] All these AESs are supported by an environmental baseline. Because payments can only be made above this baseline, the level at which each member state sets the baseline has an effect on how much farmers can be paid for their actions. In this paper, we focus on AESs related to extensive grazing. These programs share the same ecological and strategic targets: to preserve grass land areas from intensification to protect biodiversity and to support specific local production systems valorizing pastures. The objective of all AESs is to prevent the decline of permanent or semi-permanent pastures resulting from the intensification of cattle production (milk or meat) by paying farmers for the cost of preventing the intensification of the pastures. In France, the first implemented scheme, the PMSEE,[4] was set up according to the framework of the EU regulation 2098. This regulation, which was activated in 1994, sought to subsidize extensive grazing and to counteract incentives to develop more intensive agricultural practices that arose due to the premium paid for maize production. The PMSEE is designed to have a wide uptake among farmers, as it involves modest per-hectare payments to farmers and makes correspondingly modest demands on farmers. The essential condition of entry into the program was that extensive grazing areas represented of the agricultural area of the farm. When setting-up the 2000-2006 RDP, the European Commission decided that the PMSEE did not meet the criteria for an agri-environmental measure. The last contracts under the PMSEE were signed in 1999 and were paid for until 2003, i.e., for a period of four years. New programs were then instituted for replacing or succeeding the PMSEE. During the first three years of the implementation of the RDP, several programs were successively implemented due to the recommendations of the European Commission and the changes of government in France. These programs were more demanding than the PMSEE. In the case of the CTE, which was active from 2000 to 2002, and the CAD, which was active from 2003 to 2006, support for extensive grazing is part of a multi-purpose contract, which, in the case of the CTE, asks a minimum capital expenditure of the farmer. The last AES devoted to extensive grazing, or the PHAE, which was active from 2003 to 2006, and definitively replaced the PMSEE, leaves at the NUTS3 level the definition of the desirable level of extensification, and adds additional good maintaining practices to the contractual specifications.

In France, extensive grazing programs and the ICHN constitute the most important basic measures of the RDP between 2000 and 2006. These measures affect the majority of farmers in the mountains and the less-favored areas of the nation except the southeast Mediterranean. During the 2004 campaign, the agri-environmental programs were at the peak of their diffusion, and they can be subscribed into three different schemes: PHAE, CTE and CAD. Extensive grazing payments do not sustain the conversion from intensive to extensive grazing. Others measures can help with this conversion. Extensive grazing programs seek to avoid the conversion from extensive agriculture to intensive agriculture or to avoid land abandonment by compensating farmers for the opportunity cost of no change in areas that currently have extensive agricultural practices but where intensifying is attractive. The cost for no change and, consequently, the attractiveness of the contract depend not only on the characteristics of the farms but also on the spatial environment.

Potentially, these schemes related to extensive grazing offered contracting within the entire French metropolitan territory, but restrictive conditions related to farm specialization and livestock density varied according to the region and the department, which was part of the role of the public authorities managing these programs. The programs are more or less demanding in terms of environmental requirements, and thus, the payments vary. Some programs include subsidies for complementary investments (e.g., building fences). For a given area, in general, a farmer had the choice between two types of contracts: those that did or did not include additional environmental options (reducing fertilizers) or complementary supports for investment. Assuming that participating farmers are rational and choose the best contract considering the available information and the local circumstances, we consider only two situations: the choice whether to agree to an AES sustaining extensive grazing contract and the program for which the contract is signed. Two arguments support our models, which do not directly address the individual choice between alternative contract settings. First, establishing the terms of the alternative contract for each farmer is difficult, particularly because the possible combinations of options was large and the option to modify the five-year contract if the administration modified the rules to the programs was not systematically offered. Second, we do not consider differences in the characteristics of the program a farmer chooses because we focus on the common objective of all the considered programs, which is to preserve grassland areas by avoiding intensification. Success for this objective depends on the spatial context. Indeed, we assume that the spatial diffusion of the extensive grazing support measures is permitted by the variety of the programs offered to farmers; accordingly, we expect to identify local patterns of program selection in analyzing the spatial diffusion of such programs. In addition, we assume that there are interactions between the specific, intrinsic characteristics of the programs locally available and the characteristics of the territory in term of the farmers' political influence and the local demand for environmental measures .

3 Spatial Durbin Probit Model

Our dependent variable is a dichotomous choice variable, we denote by y_i, which takes a value of 1 when farmer i participates in an agri-environmental scheme, and 0 if he does not. The farmer's choice to participate depends on the difference in utilities, $U_{1i} - U_{0i}$, associated with participating (U_{1i}) and not participating (U_{0i}). We denote this difference by y_i^*. We do not observe y_i^*; we only observe the choice made. Based on utility maximization, when choosing to participate or not, the relationship between the observed choice y_i and the latent variable y_i^* is

$$\begin{cases} y_i = 1, & \text{if } y_i^* \geq 0 \\ \quad \text{or} & \\ y_i = 0, & \text{if } y_i^* < 0 \end{cases} \tag{1}$$

Thus, the probability of participating can be expressed as

$$\Pr(y_i = 1) = \Pr(U_{1i} - U_{0i} \geq 0) = \Pr(y_i^* \geq 0)$$

The standard probit model, i.e., the model without any spatial effects, assumes the following non-spatial regression relationship

$$y_i^* = x_i \beta + \varepsilon_i \tag{2}$$

where x_i is an $1 \times k$ vector of the observations for farm i on the independent (explanatory) variables related to the the farm or to the farmer. The error term, ε_i, is assumed to be normally distributed, i.e., $\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2)$.

Spatial dependencies in Probit model can be incorporated in several ways (see Chapter 10 in Lesage and Pace, 2009). We choose to estimate them using the spatial Durbin probit model, or SPDM, where equation (2), expressed in matrix form, becomes

$$y^* = \rho W y^* + X\beta + WX\delta + \varepsilon \tag{3}$$

where y^* is an $n \times 1$ vector of realizations of the latent dependent variable, X is an $n \times k$ matrix of observations on the explanatory variables, W is a matrix of spatial weights, and ε is an $n \times 1$ vector of i.i.d. random error terms. Some neighboring criteria are chosen to determine the structure of the spatial weights, which are routinely based on contiguity (farmers belonging to the same county, for example) or a distance criterion. The weights in W are usually row-standardized such that the elements of each row sum to one. Thus, the SDPM includes spatial lags of the dependent variable (Wy^*) and of the independent variables (WX), in addition to the independent variables (X). The implication of (3) in the present study is that the difference in utilities that determines the choice of a given farmer to participate or not is a function of the weighted average of the differences in utilities that determine the choices of the farmer's neighbors, of independent variables related to the farm or to the farmer, as well as the weighted average of these independent variables in a given neighborhood.

The main reason for choosing a SDPM is related to the possible endogeneity of some independent variables. By definition, our data set includes only those variables observed at the farm level and at some spatially aggregated locations associated with each individual observation. However, there are a host of latent, unobservable and frequently non measurable influences that likely impact the decision to participate by a farmer. For example, unlike studies using stated choice data such as Barreiro-Hurlé *et al.* (2010), we are unable to observe factors that may exert an influence on the difference in utilities in equation (5), such as a farmer's preference for nature conservation. It is sometimes possible to incorporate observable variables that capture the effect of such latent factors. For instance, Bertoni *et al.* (2008) use observable variables that describe the political and institutional environment, including membership in a specific farmers' organization or the ideological orientation of the district, as proxies for the transaction costs of the political bargaining process that might influence the probability of individual participation in agri-environmental programs. However, it is unlikely that observable variables can capture all the possible latent explanatory variables. Because the omitted and included explanatory variables are both likely to exhibit spatial dependence based on the same spatial connectivity structure, it seems likely that omitted and included variables will exhibit non-zero covariance. Therefore, we may experience endogeneity biases due to omitted explanatory variables.

Endogeneity in a spatial context has received attention only in linear regression models (see, among others, Fingleton and Le Gallo, 2010). However, to our knowledge, no tools currently eliminate endogeneity biases in spatial probit models. We are only aware of the spatial sampling technique employed by Carrión-Flores and Irwin (2004) and Albers *et al.* (2008) to eliminate bias from spatial autocorrelation using non-spatial probit techniques (more precisely, an instrumental variable probit analysis) to explore the potential consequences of endogeneity in empirical applications. Nevertheless, in the context of linear regression models, estimates based on a spatial Durbin specification shrink the endogeneity bias relative to OLS (Pace and LeSage, 2010). Indeed, spatially dependent omitted variables result in a model with a design matrix containing spatially lagged versions of the dependent and the explanatory variables or, in other words, a spatial Durbin model. The set of equations resulting in a spatial Durbin model can be expressed as

$$y = X\beta + \eta,\ \eta = \rho W\eta + \varepsilon,\ \text{and}\ \varepsilon = X\gamma + u$$

where y is an $n \times 1$ vector of observations on the continuous dependent variable, X is an $n \times k$ matrix of observations on the explanatory variables, η is an $n \times 1$ vector of a spatially correlated omitted variable following a spatial autoregressive process with autoregressive coefficient ρ, and u is an $n \times 1$ vector of i.i.d. random error terms. The last equation shows that the omitted variable is correlated with the explanatory variables when $\gamma \neq 0$. Using this set of equations and the cofactor restrictions, Lesage and Pace (2009) show that the resulting model is a spatial Durbin model (see also, Brown *et al.*, 2009).

Additionally, the magnitude of the omitted variable bias in the spatial Durbin model does not exhibit the magnification of an OLS, and it no longer depends on the magnitude of spatial dependence in the disturbances or the dependent or independent variables. As emphasized by Pace and Lesage (2010), these desirable properties of the spatial Durbin model provide a strong motivation to use this model in applied practice. Pace and Lesage (2010) do not generalize their results to the case of spatial probit models when using a spatial Durbin specification, but we can show that a similar reasoning to that of Lesage and Pace (2009) may be performed to derive equation (5) from a more general model with a latent dependent variable and omitted spatially dependent variables. We can speculate that the omitted variable bias is also mitigated in the case of spatial probit models when using a spatial Durbin specification that matches the implied data-generating process for this set of circumstances.

In sum, the implications of selecting the latent model to describe contracting decisions, i.e. equation (3), are that the difference in utilities of a given farmer is a function of the weighted average of its neighbors' difference in utilities (Wy^*), a set of possible determinants of the contracting decision (X) and the weighted average of these variables in a given neighborhood (WX). When contextual factors are introduced in the analysis as explanatory variables, equation (3) becomes

$$y^* = \rho W y^* + X\beta + WX\delta + Z\gamma + \varepsilon \tag{4}$$

where Z denotes the vector of these contextual factors such that $WZ = Z$. Indeed, the contextual factors are defined at a geographical level that includes the contiguity criterion used for defining the spatial weights matrix W.

The spatial Durbin probit model estimates are obtained using a Monte Carlo Markov Chain (MCMC) procedure, i.e., Gibbs sampling with a Metropolis-Hastings algorithm, as proposed by Lesage (2000). This procedure entails specifying a complete conditional distribution for the model parameters and iteratively sampling from this conditional distribution. The conditioning variables for each set of model parameters are the most recent draws of other model parameters. The sequence of the resulting parameter draws converges to the joint posterior distribution of parameters after a sufficient number of draws. To construct the values of the latent dependent variable, Lesage (2000) adds an additional conditional distribution for the posterior distribution of this latent variable that is conditional on all other parameters and takes the form of a truncated normal distribution, as proposed by Albert and Chib (1993) in their pioneering work on the Bayesian analysis of binary and polychotomous response models.[5]

4 Data

The empirical analysis uses data assembled in the framework of the "Observatoire des Programmes Communautaires de Développement Rural" or ODR. The ODR was created in December 2006 to meet the statistical needs for the evaluation of the French RDP. The data we use include all farms with more than three hectares for which a request for aid from the first or second CAP pillars was filed to benefit from direct payments linked to animal breeding or from agri-environmental support related to extensive grazing. The data include 259,316 observations for the year 2004 from both individual farms and partnerships, with 18,311 observations in Basse-Normandie and 20,109 in Auvergne. The definitions of all the variables we use and the corresponding acronyms are summarized in Table 1.

Table 1 – **Variable Definitions**

Variable name	Definition	Unit
Farm and farmer characteristics		
AES	Participation in an AES	1=yes, 0=no
AUC	Area under cultivation	ha
AGE	Farmer's age or age of the younger partner in the case of partnership	year
PHAE	Eligible to participate in PHAE	1=yes, 0=no
STATUS	Farm status (1=individual, 2=partnership)	indicator
ICHN	Participation in ICHN program	1=yes, 0=no
Contextual factors		
LFA	Less favored area (1=mountain, 2=less-favored, 3=plain)	indicator
PBEEF	Average direct payment for beef	10^3 euro
PEXTEN	Average compensation payment for extensive grazing	10^3 euro
PGOAT	Average direct payment for goat meat	10^3 euro
GRASS	Share of grassland (per area under cultivation)	> 0 and < 1
LIVE	Share of livestock farms	> 0 and < 1
CTE	Share of participants in CTE 19-20	> 0 and < 1
PMSEE	Share of participants in PMSEE in 2002	> 0 and < 1
CLD	Cattle livestock density (per total area)	number/ha
MILK	Share of milk livestock (in cattle livestock)	> 0 and < 1

4.1 Dependent variable

The dependent variable, denoted by AES, is a dichotomous variable indicating whether a farm participates (1) or does not participate (0) in one AES. Table 2 reports the average values of all variables for the entire sample, i.e., at the national level, and for the two subsamples of the two regions under investigation.[6] The AES participation rate in Auvergne is above the national rate, while the AES participation rate in Basse-Normandie is lower than the national rate. The spatial dispersion of the observed participation rates at the C27 level is shown on the maps on the left sides of Figures 1 and 2. In Auvergne, the large mountainous areas show a participation rate above 80% . The Limagne plain, located in the center of the region and home to the river Allier, has participation rates below 30% . This geographical area is not devoted to milk production, as in the mountainous regions, or meat production; this area focuses on cereal crop production. The 30% participation rate corresponds to the lower sextile in Auvergne and the higher sextile in Basse-Normandie. In this region, the areas with high participation rates are notably wetlands areas (principally in the west of the region), and the area that produces the cheese "Camembert de Normandie" which benefits from a PDO (Protected Designation of Origin) labeling. In this area, contracts are signed by farmers to preserve a traditional, extensive milk system of production

Table 2 – **Some descriptive statistics**

Variables means	France	Basse-Normandie	Auvergne
AES	0.300	0.166	0.669
AUC	71.264	59.858	65.949
AGE	48.017	50.487	46.437
PHAE	0.500	0.477	0.869
STATUS1	0.720	0.753	0.787
STATUS2	0.280	0.247	0.213
ICHN	0.367	0.061	0.830
LFA1	0.234	0.000	0.744
LFA2	0.300	0.193	0.21
LFA3	0.466	0.807	0.046
PBEEF	140.424	104.918	156.372
PEXTEN	20.288	10.736	35.744
PGOAT	14.181	2.855	13.296
GRASS	0.400	0.512	0.671
LIVE	0.758	0.959	0.905
CTE	0.331	0.423	0.235
PMSEE	0.284	0.145	0.635
CLD	0.442	0.767	0.477
MILK	0.338	0.410	0.350
N	265316	18311	20109

4.2 Independent variables

Variables used in the model are measured at the individual level (farm and farmer characteristics) and at the municipality level (contextual factors). Tables 1 and 2 give a complete description of the variables defined in the following two subsections. The variables we used as determinants of a farmer's contracting decision capture different dimensions such as (1) the main orientation of farm production, (2) the human capital on the farm, (3) the social capital (network and learning effects), and (4) the institutional quality of the area surrounding the farm.

4.2.1 Farm and farmer characteristics

The literature on farmers' willingness to participate in AESs indicates that variables related to farm and farmer characteristics are critical factors for explaining participation in AESs (Defrancesco and al. 2006, Barreiro-Hurlé and al. 2008, Vanslembrouck and al. 2002). Only one variable is available at the farm level to summarize the farm's main production: the share of grasslands in the total farm area under cultivation. We transform this variable into a dummy variable describing eligibility for participation in PHAE. This new variable, denoted by $PHAE$, is equal to one if the share of grasslands is higher than the share defined at the department (NUTS3) level as the eligibility threshold and is zero otherwise. The eligibility threshold varies by department from 20% to 80% in our sample; the most common threshold is 75%. We do not have information regarding the farmer's level of education, and therefore, the human capital on the farm is assessed indirectly. We use the farmer's age or the age of the younger partner if the farm is a partnership, AGE, and partnership status, $STATUS$, as proxies. Another critical farm characteristic is farm size. Here, we use the area under cultivation, AUC. The social capital is represented by the variable participation in the ICHN program, $ICHN$. The ICHN program is a basic agri-environmental program designed to attract a large number of farmers from less-favored areas. Participation in this program can be an indicator of the social integration of the farmer.

4.2.2 Contextual factors

Contextual factors are added to farm and farmer characteristics to assess the farm environment. To describe the main agricultural production in the area surrounding a farm, we add variables from the municipality level such as the share of grasslands per area under cultivation, $GRASS$, the share of livestock farms, $LIVE$, the cattle livestock density, CLD, the share of milk livestock in total cattle livestock, $MILK$, and the total amount of direct payments for beef, $PBEEF$, and goat breeding, $PGOAT$, as well as compensation for extensive agriculture from the CAP, $PEXTEN$. These variables allow us to identify local breeding systems, particularly in terms of the use of grass, which is an issue associated with the studied AESs.

The institutional quality of the surroundings of the farm is captured by the share of participants in a previous agri-environmental scheme, PMSEE, and by the share of the most exigent contract, CTE 19-20, at the municipality level, $PMSEE$ and CTE , respectively. Finally, we characterize the level of incentives of farms in plains, less-favored areas or mountains, as some payments are adjusted by location, LFA. Unfortunately, participation in the ICHN program is closely linked to this latter variable that we include in the contextual factors. Thus, participation in the ICHN program, which is available at the farm level, will be excluded from the analysis when considering models that involve contextual factors as explanatory variables.

5 Empirical Results

5.1 Estimation issue - The choice of W

The estimation of the SDPM requires the specification of a spatial weight matrix W. As stated previously, each row i of this matrix contains elements that are $1/m_i$ or 0, where m_i is the number of neighbors of farm i. If farm j represents one of the m_i farms neighboring farm i, then the (i,j)th element of W contains the value $1/m_i$. All elements in the ith row the matrix W not associated with neighboring observations take values of 0. By construction, the elements of W are non-negative, and each row sums to 1. There are different possible definitions of neighborhood. One definition is the physical distance between two farms in terms of longitude and latitude, an integral part of the available data. Then, m_i can be defined as some number of the nearest neighbors of farm i. Our data do not contain such specific information, and therefore, other criteria must be used. The only information we have about the location of a farm is its location in a given municipality. We then choose to build the matrix W by considering those farms located in the same municipality as neighbors. One limitation to this choice is that two distant farms in the same municipality can be considered as neighbors, while two others farms very close to each other but in two different municipalities would not. We could increase the area defined to be a neighborhood, but the problem would remain and the concept of the neighborhood would become too large. Moreover, municipality has been shown as the relevant geographical area when defining the neighborhood of a farm in studies dealing with the issue on how professional standards are built and disseminated in groups of French farmers and in groups of breeders (see, among others, Darré, 1984).

5.2 Diagnostics for the Presence of Spatial Effects

The last two columns of Table 3 provide the results of the estimates from the SDPM with contextual factors for the two considered regions, Basse Normandie and Auvergne. The models fit well for cross-sectional data, with pseudo-R^2 values ranging from $0.62 - 0.76$.

Table 3 – **Parameter Estimates from Spatial Durbin probit models**

	Models without contextual factors		**Model with contextual factors**	
	Basse-Normandie	Auvergne	Basse-Normandie	Auvergne
Intercept	-4.202***	-5.138***	-5.345***	-6.240***
AUC	0.005***	0.007***	0.006***	0.009***
AGE	0.139***	0.161***	0.130***	0.195***
AGE2	-0.002***	-0.002***	-0.002***	-0.002***
PHAE	0.964***	1.685***	0.952***	1.792***
STATUS2	0.061**	-0.119***	0.046*	-0.161***
ICHN	0.479***	1.200***	.	.
$W\times$ AUC	0.001	-0.001	-0.001	-0.001*
$W\times$ AGE	0.005	0.006	0.013*	-0.001
$W\times$ AGE2	0.000**	0.000	0.000	0.000
$W\times$ PHAE	0.391***	-0.161**	-0.323***	-0.608***
$W\times$ STATUS2	0.006	-0.121*	0.076	0.324***
$W\times$ ICHN	-0.218***	-0.864***	.	.
LFA1	.	.	.	0.194***
LFA3	.	.	0.108***	-0.035
PBEEF	.	.	-0.001***	-0.004***
PEXTEN	.	.	0.010***	0.013***
PGOAT	.	.	0.000	-0.002***
GRASS	.	.	0.915***	0.601***
LIVE	.	.	0.022***	0.286***
CTE	.	.	0.535***	0.242***
PMSEE	.	.	1.493***	0.950***
CLD	.	.	-0.136***	0.137***
MILK	.	.	0.381***	-0.206**
ρ	0.201***	0.395***	-0.008	0.137***
Pseudo R^2	0.62	0.74	0.71	0.76
AUROC	0.791	0.845	0.836	0.852
N	18311	20109	18311	20109

Note: *,**, and *** denote that the estimated parameter is significantly different from 0 at the usual 10%, 5%, and 1% levels respectively.

Moreover, the models perform well in classifying the farms as participating farms or not, as shown by the high values of the areas under the ROC curve, or $AUROC$ (see Lahiri and Yang, 2012). The magnitude of the estimated value of the autoregressive parameter ρ is small for the two regions, although significantly different from zero for Auvergne. This implies that a farmer's decision to contract in Auvergne has a positive effect on the decisions of his neighbors and vice versa.

To assess the impact of the introduction of contextual factors on spatial dependence when estimating SDPM, in Table 3, we report the results from the estimation of a SDPM that does not include contextual factors as explanatory variables for each region. In this model, we include the $ICHN$ variable as an explanatory variable. This variable was excluded from the SDPM with contextual variables due to its collinearity with the LFA variable. The performance of the models with contextual factors in terms of overall quality of fit or of the accuracy of the classification of the farms is comparable to the performance of models without contextual variables, except for Basse Normandie, where significant improvements in pseudo-R^2 and $AUROC$ are due to the introduction of the contextual factors. Furthermore, their introduction leads to a significant reduction in the estimated value of the autoregressive parameter for either region. The impact of omitted variables becomes less important in Auvergne and even disappears in Basse Normandie when contextual factors are included in the spatial Durbin probit models.

5.3 Assessment of the predictive performance of the SDPM at various geographical scales

We now consider the ability of the SDPM to replicate the spatial distribution of observed contracting shares in the two regions. This ability can be assessed at different geographical scales, such as the NUTS5 or the C27 local unit, depending on the objectives of the assessment. The tests and measures that have been proposed to compare the predictive ability of discrete choice models (see Haener *et al.*, 2001, for an overview) can then be applied. Many of these tests operate at the aggregate level, comparing observed and predicted shares. In line with Ben-Akiva and Lerman (1985), shares for a given geographical unit c can be computed as $(1/n_c)\sum_{i=1}^{n} d_{ic}\, y_i$ and $(1/n_c)\sum_{i=1}^{n} d_{ic}\, \widehat{P}_i$, respectively, where n_c denotes the number of farmers located in the geographical unit, d_{ic} is an indicator of whether or not farm i is located in this geographical unit, and $\widehat{P}_i$ denotes the predicted probability that a farmer participates in the program. We implement the regression tests for a slope of one and an intercept of zero in the regression of observed aggregate shares on predicted shares (Horowitz, 1988).

Table 4 – **Empirical levels of significance for geographical prediction tests**

	Basse-Normandie	**Auvergne**
Geographical unit:		
NUTS5	$< 10^{-3}$	0.269
C27 local unit	0.209	0.542
NUTS4	0.028	0.474
PRA local unit	0.813	0.408
NUTS3	0.541	0.652

Results of the prediction tests at the different geographical scales are reported in Table 4. The reported empirical levels of significance clearly indicate that the null hypothesis is not rejected, which has a slope of one and an intercept of zero for most geographical scales. This result is particularly encouraging because the tests are not corrected for the predicted shares being estimated in a first step. Simulations have demonstrated that, without this correction, these regression-based tests reject the null hypothesis too often (see Horowitz and Louviere, 1993). Clearly, the estimated SDPM performs well in the replication of the spatial distribution of observed shares for Auvergne, whatever the chosen level of geographic aggregation. A different result is obtained for Basse-Normandie; the model failed to replicate the observed shares for two levels of geographical aggregation, not surprisingly, NUTS5 and NUTS4. For this second scale of geographic aggregation, the rejection of the null hypothesis is the result of a few influential NUTS4 areas where the observed shares are very high compared to the average share.[7] In addition, the difference between the results of prediction at the geographical aggregation scale ofC27 local units and NUTS4 for Basse-Normandie illustrates a common problem in geographical aggregation of spatial data known as the Modifiable Areal Unit Problem, or MAUP (Openshaw, 1984). This problem concerns not only the well known problem of the aggregation of smaller units into larger ones, or scaling bias, but additionally the less-often-addressed problem of alternative allocations of zonal boundaries, or gerrymandering. The second aspect of MAUP matters here. NUTS4 boundaries are defined at the administrative level, and therefore, some NUTS4 may be very heterogeneous regarding some intrinsic aspects of AESs. For example, only a part of a NUTS4 is classified as a less-favored area. On the contrary, C27 local units were built with the aim of evaluating rural development policies. Their boundaries thus define a homogeneous geographical area with respect to agri-environmental measures. Overall, the estimated SDPM allows for correctly recovering the spatial distribution of the observed rates in the two regions at the C27 local unit aggregation scale, i.e., from the level of geographic aggregation, considered relevant when evaluating rural development policies.

The quality assessment of the estimated SDPM to explain the spatial distribution of participation rates in agri-environmental programs can be enhanced by the inspection of maps. The spatial distributions of observed and predicted shares evaluated at the C27 local unit scale are represented in Figure 1 for Basse-Normandie and in Figure 2 for Auvergne. 276 (resp. 268) C27 local units are present in Basse-Normandie (resp. Auvergne). Beneath each map, we report estimates of the densities of these shares as histogram estimators. Overall, we see the same spatial trends in the distributions of observed and predicted shares when comparing the maps for each region. Nevertheless, the SDPM tends to overpredict zero participation rates in both regions and high participation rates for Basse-Normandie. This leads to a higher number of C27 local units with high predicted shares or with predicted shares equal or close to zero. For instance, in Basse-Normandie, 14 areas have a predicted share greater than 70% in contrast to no participation among the observed shares. Similarly, 49 C27 local units have a predicted share smaller than 1% instead of the observed share of 16%. Extreme shares are poorly predicted because some C27 local units have very small numbers of breeders, as expected.

5.4 Evaluation of Marginal Effects

As shown above, the autoregressive parameter ρ interacts with the other explanatory variables, and their impact on the probability of participation can be assessed. As a result, the estimated values of the coefficients cannot be compared directly between the different models. Marginal effects must be evaluated. As emphasized by Lesage and Pace (2009), one advantage of the MCMC estimation technique of the spatial probit model is that the sample draws from the estimates can be used to produce separate marginal effects for every observation at each iteration. Averaging over these results at each iteration, the global nature of spatial spillovers that cumulate across interrelated observations in the sample can be recognized, and this reflects the joint posterior distribution of the marginal effects. Table 5 reports the mean values of the direct, indirect and total marginal effects for all the estimated spatial Durbin probit models with contextual factors (see Appendix 2 for their computation). The significance of these effects is assessed by testing if value of zero belongs in the 95% confidence interval as deduced from the empirical distribution of the estimated marginal effects.

Consider, first, the estimated values of the average total marginal effects of the individual characteristics. They are significant except for the effect of the farm being a partnership in Basse-Normandie, i.e., $STATUS2$. Most of these effects have the same positive sign in both regions. Other studies have shown positive effects of these individual variables with the notable exception of AGE (see Defrancesco *et al.*, 2008, for a survey). Thus, a positive effect of this variable can be expected when the AES does not involve any specific investments, such as the measure studied here, and a negative effect if not. Older farmers are expected to

Table 5 – **Direct, Indirect and Total Marginal Effects Estimates**

	Basse-Normandie			Auvergne		
	DME	IME	TME	DME	IME	TME
AUC	0.001*	0.000	0.001*	0.002*	0.001*	0.002*
AGE	0.026*	-0.001	0.026*	0.046*	0.015*	0.061*
AGE2	0.000*	0.000	0.000*	-0.001*	0.000*	-0.001*
PHAE	0.117*	-0.004	0.113*	0.250*	0.080*	0.330*
STATUS2	0.025	-0.001	0.024	0.036*	0.011*	0.047*
LFA1	.	.	.	.	.	0.061*
LFA3	.	.	0.021	.	.	-0.004
PBEEF	.	.	0.000*	.	.	-0.001*
PEXTEN	.	.	0.002*	.	.	0.004*
PGOAT	.	.	0.000	.	.	-0.001
GRASS	.	.	0.167*	.	.	0.165*
LIVE	.	.	0.004*	.	.	0.081*
CTE	.	.	0.098*	.	.	0.071*
PMSEE	.	.	0.277*	.	.	0.260*
CLD	.	.	-0.024*	.	.	0.035
MILK	.	.	0.068*	.	.	-0.056*

Note: – DME, IME, and TME denote average direct, indirect, and total marginal effects respectively.
– ∗ indicates a significant marginal effect.

show a higher propensity for measures involving only small changes to their usual farming practices, while young farmers are more willing to take risks and are therefore more inclined to enter more demanding AESs. Barreiro *et al.* (2010) emphasize that when the AES is focused on extensification, older farmers are more prone to enroll in this type of AES, as it requires less labor and does not require new investments in either labor and/or knowledge, the main reasons that deter older farmers from participating.

Consider, first, the estimated values of the average total marginal effects of the individual characteristics. They are significant except for the effect of the farm being a partnership in Basse-Normandie, i.e., . Most of these effects have the same positive sign in both regions. Other studies have shown positive effects of these individual variables with the notable exception of (see Defrancesco *et al.*, 2008, for a survey). Thus, a positive effect of this variable can be expected when the AES does not involve any specific investments, such as the measure studied here, and a negative effect if not. Older farmers are expected to show a higher propensity for measures involving only small changes to their usual farming practices, while young farmers are more willing to take risks and are therefore more inclined to enter more demanding AESs. Barreiro *et al.* (2010) emphasize that when the AES is focused on extensification, older farmers are more prone to enrol in this type of AES, as it requires less labor and does not require new investments in either labor and/or knowledge, the main reasons that deter older farmers from participating.

Average total marginal effects are all higher for Auvergne than for Basse-Normandie because both the average direct and indirect marginal effects are higher for the first region than for the second. Moreover, no average indirect marginal effects are significant for Basse-Normandie, whereas they are all significant and positive for Auvergne. The interpretation of this result is that, on average, the effect of increasing one of the individual characteristics of farm is positive, regardless of region, but the spillover to the neighboring farms in Auvergne is positive, and therefore, the probability of participation increases even more when one individual characteristic is increased. Finally, the estimated average total marginal effect of being eligible for the PHAE program is very high for the two regions. This is linked with the high level of contracting in Auvergne. Extensive grazing contracts are similar to entry-level programs, and local spillover effects intervene in the diffusion. This effect is more pronounced for Auvergne than for Basse-Normandie. To interpret this result, it is important to remember that the AESs under study cover a large number of potential farmers and seek to avoid or limit the intensification of agriculture rather than promote extensification. The proportion of eligible farmers in Basse-Normandie, although high (48%), is relatively small when compared to the proportion in Auvergne (87%). Moreover, the participation rate in the first region, 17%, is significantly smaller than for the second region, 67%. The adoption of the measure can then be interpreted as a highly personal decision in Basse-Normandie, whereas there appears to be a ripple effect in Auvergne. This may explain why the marginal total effect of being eligible for the measure ($PHAE$ variable) is more pronounced for Basse-Normandie than for Auvergne, but there is a significant indirect marginal effect in the Auvergne.

Consider now the estimated values of the average total marginal effects of the contextual factors. As shown above, direct and total marginal effects are confounded because the value of contextual factor is the same for a given farm and its neighbors by definition. Most effects are significant except for $LFA3$ and $PGOAT$ in both regions and for CLD in Auvergne. The marginal effects of direct compensation or first pillar payments are relatively small, but significant. As expected, an increase in the share of grassland in the municipality where a farm is located significantly increases the farm's probability of participating. This increase is relatively high but identical across the two regions. Similarly, an increase in the share of livestock farms in the farm's municipality increases its probability of participating, but the effect is smaller in both regions. The marginal effect of the shares of participants in the CTE and PMSEE programs in the municipality are significant and positive, and their magnitudes are comparable across the two regions. The marginal effect of the share of participants in the second program is very strong in relation to the basic nature of this program. Cattle livestock density in the municipality has a significant effect only for Basse-Normandie, and this effect is negative. We find the reverse for the marginal effect of the share of milk livestock in the municipality, the effect being significant for the two regions, but in different directions.

In sum, two results for the effects of contextual factors are notable. First, past participation in an AES such as PMSEE significantly increases the probability of participating again, indicating continuity in the programs. The effect is high because the objective of the contracts is to maintain grassland areas and farms that are specialized in cattle production and that use extensive grazing are generally located in specific geographic areas, often mountainous ones. Under these conditions, farmers who do not intensify their production have benefited from the PMSEE for some time. The PMSEE was a measure targeting extensive grazing areas, while CTE was a measure covering a broader category of farmers in less-specialized areas. The smaller effect of the share of CTE participants in the municipality is thus an indicator of the spatial diffusion of rural development programs in comparison to the former PMSEE program. The effect is higher in Basse-Normandie because the share of farmers who benefited from the CTE is higher in this region. Second, the reverse effects of the share of milk and livestock farms in the municipality in the two regions may be explained by differences in dairy systems in these regions. The two regions are covered in large part by geographical areas devoted to the production of PDO (Protected Designation of Origin) cheeses or butters. Introducing a dummy variable indicating whether the municipality where the farm is located belongs to a geographical area devoted to PDO cheese or butter production does not change our main results. Indeed, the coefficients of this variable were not significantly different from zero.[8] Characteristics of milk production in the two regions are well captured by the chosen variables. Milk production in Auvergne is concentrated in the less mountainous areas (valleys and plateaus) and is generally more intensive than the meat production that occurs in the mountainous areas that include extensive grazing. Instead, milk production in Basse-Normandie is notably valorizing grasslands and wetlands, some of which belong to natural parks where agriculture practices cannot be intensified (see Figure 1).[9] Conversely, intensive practices characterize the meat production in this region, which is mainly devoted to milk production. Dairy farmers in Basse-Normandie are more likely to adopt the AESs, unlike dairy farmers in Auvergne. To conclude, different regional production systems lead to different participation profiles.

6 Conclusion

This paper contributes to the literature on the determinants of farmers' participation in agri-environmental programs, dealing explicitly with the spatial dimension and analyzing observed real world behaviors instead of stated intentional or contingent ones. More specifically, this paper integrates the spatial diversity of both extrinsic and intrinsic factors and their spatial interactions in the analysis of farmers' participation. Special attention is paid to the assessment of the spatial diffusion of agri-environmental programs related to extensive grassland management. The analysis is performed on two samples of farmers located in two French regions, Auvergne and Basse-Normandie, with very different breeding practices, to investigate the link between these practices and farmers' decisions to participate in these programs. Our modeling approach of farmer's participation decisions relies on the use of spatial Durbin probit models. These models explicitly consider possible spatial spillovers on individual decisions and allow for an explicit treatment of the omitted variable problem frequently encountered when an analysis is based on observed and not stated data.

Determinants of a farmer's participation in the different models included factors describing the background of the farmer's decision, or contextual factors, and farm characteristics. These contextual factors are available at some aggregated geographical scale and describe the main agricultural production system in the area where the farm is located. Our results show that adding these contextual factors to the individual factors in the set of potential factors significantly reduces the estimated value of the correlation between farmers' decisions. This correlation becomes less important but remains significantly different from zero in Auvergne, while it becomes insignificant in Basse-Normandie. The impact of omitted variables becomes less important in Auvergne, and even disappears in Basse-Normandie, when contextual factors are considered in the spatial Durbin probit models. Our results for the marginal effects of farm characteristics confirm those in the literature. A farmer's eligibility for agri-environmental programs has a non-negligible impact on participation probability. This indicates correct targeting by these programs. As for the contextual variables, belonging to a geographical area where many farmers participated in past agri-environmental programs has a positive effect on the probability of participating in current programs, showing continuity and an expansion in the French rural development policy. Finally, the reverse effects of the contextual factors describing differences in the milk production systems between the two regions show that different regional production systems can lead to very different contracting profiles.

The assessment of the performances of the estimated spatial Durbin probit models at different geographical scales through prediction tests show that these models perform well in capturing the spatial diffusion of the considered agri-environmental programs. Overall, the estimated models allow for correctly recovering the spatial distribution of observed participation rates in the two regions from the C27 local unit aggregation level, i.e., the level of geographic aggregation considered relevant when evaluating rural development policies.

7 References

Albers, H.J., Ando, A.W., and X. Chen (2008). Spatial-econometric analysis of attraction and repulsion of private conservation by public reserves. Journal of Environmental Economics and Management 56: 33-49.

Albert, J.H., and S. Chib (1993). Bayesian analysis of binary and polychotomous response data. Journal of the American Statistical Association 88: 669-679.

Anselin, L., and D. Arribas-Bel (2013). Spatial fixed effects and spatial dependence. Papers in Regional Science 92: 3-17.

Barreiro-Hurlé, J., Espinosa-Goded, M., and P. Dupraz (2010). Does intensity of change matter? Factors affecting adoption in two agri-environmental schemes. Journal of Environmental Planning 53: 891-905.

Baylis, K., Peplow, S., Rausser, G., and L. Simon (2008). Agri-environmental policies in the EU and United States: A comparison. Ecological Economics 65: 753-764.

Ben-Akiva, M., and S.R. Lerman (1985). Discrete Choice Analysis: Theory and Application to Travel Demand. Cambridge, MIT University Press.

Bertoni, D., and A. Olper (2008). The political economy of EU agri-environmental measures: An empirical assessment et the regional level. Contributed paper, 12th EAAE Congress, Ghent, Belgium.

Bertoni, D., Cavicchioli, D., Pretolani, R., and A. Olper (2008). Agri-environmental measures adoption: New evidence from Lombardy region. Paper prepared for the 109th EAAE Seminar '*The CAP after the Fischler Reform: National Implementations, Impact Assessment and the Agenda for Future Reforms*. Viterbo, Italy.

Brown, J., Florax, R., and K. McNamara (2009). Determinants of investment flows in U.S. manufacturing. Review of Regional Science 39: 77-97.

Carrión-Flores, C., and E.G. Irwin (2004). "Determinants of residential land-use conversion and sprawl at the rural-urban fringe. American Journal of Agricultural Economics 86: 889-904.

Darré, J.-P. (1984). La production des normes au sein d'un réseau professionnel: l'exemple d'un groupe d'éleveurs. Sociologie du Travail 26: 141-156.

Defrancesco, E., Gatto, P., Runge, F., and S. Trestini (2008). Factors affecting farmers' participation in agri-environmental measures; A Northern Italian perspective. Journal of Agricultural Economics 59: 114-131.

Eurostat (2007). Regions in the European Union - Nomenclature of territorial units for statistics - NUTS 2006 / EU-27. Luxembourg, Office of the Official Publications of the EU.

Fingleton, B., and J. Le Gallo (2010). Endogeneity in a spatial context: Properties of estimators. In: A. Paez, J. Le Gallo, R. Buliung, and S. Dall'Arba, eds. Progress in Spatial Econometrics: Theory and Methods, and Thematic Applications. Berlin: Springer, pp. 59-73.

Fleming, M. (2004). Techniques for estimating spatially dependent discrete choice models. In: L. Anselin, L., Florax, R.J.G.M. Florax, and S.J. Rey, eds. Advances in Spatial Econometrics. Berlin: Springer, pp. 145-168.
Haener, M.K., Boxall, P.C., and W.L. Adamowicz (2001). Modeling recreation site choice: Do hypothetical choices reflect actual behavior? American Journal of Agricultural Economics 83: 629-642.
Holloway, G., Shankar, B., and S. Rahman (2002). Bayesian spatial probit estimation: A primer and an application to HYV rice adoption. Agricultural Economics 27: 383-402.
Horowitz, J.L. (1988). Specification tests for probabilistic discrete choice models of consumer behavior. In: R.G. Golledge, and H. Timmermans, eds. Behavioral Modeling in Geography and Planning. London: Croom-Helm, pp. 124-137.
Horowitz, J.L., and S. Louviere (1993). "Testing predicted choices against observations in probabilistic discrete-choice models. Marketing Science 12: 270-279.
Lahiri, K., and L. Yang (2012). Forecasting Binary Outcomes. In: G. Elliott and A. Timmermann, eds. Handbook of Economic Forecasting. Vol. 2: in press.
Lesage, J.P., Pace, R.K., Lam, N., Campanella, R., and X. Liu (2011). New Orleans Business Recovery in the Aftermath of Hurricane Katrina. Journal of the Royal Statistical Society, Series A 174: 1007-1027.
Lesage, J.P., and R.K. Pace (2009). Introduction to Spatial Econometrics. CRC Press, Taylor and Francis Group.
Lesage, J.P. (2000). Bayesian estimation of limited dependent variable spatial autoregressive models. Geographical Analysis 32: 19-35.
Openshaw, S. (1984). The modifiable areal unit problem. Concepts and Techniques in Modern Geography 38.
Pace, R.K., and J.P. LeSage (2010). Omitted variable biases of OLS and spatial lag models. In: A. Paez, J. Le Gallo, R. Buliung, and S. Dall'Arba, eds. Progress in Spatial Econometrics: Theory and Methods, and Thematic Applications. Berlin: Springer, pp. 17-28.
Peerlings, J., and N. Polman. (2009). Farm choice between agri-environmental contracts in the European Union. Journal of Environmental Planning and Management 52: 593-612.
Train, K.E. (2009). Discrete Choice Methods with Simulation, 2nd edition. Cambridge: Cambridge University Press.
Vanslembrouck, I., van Huylenbroeck, G., and W. Verbeke (2002). Determinants of the willingness of Belgian farmers to participate in agri-environmental measures. Journal of Agricultural Economics 53, 489-511.

Notes

[1]Due to the definition of neighbors we have chosen in spatial modeling: The fact that farms belong to the same municipality, it is possible to show that these two approaches are equivalent (Anselin and Arribas-Bel, 2013). But the spatial econometric model is more parsimonious in terms of the number of parameters to be estimated and we have not enough observations in many municipalities in order to get reliable estimates of the fixed effects.

[2]The Nomenclature of Units for Territorial Statistics, or NUTS (see Eurostat, 2007). We additionally consider two other types of geographical units called the "C27 local unit" and the "PRA ('Petite Région Agricole') local unit". A C27 local unit groups contiguous municipalities with the same economic characteristics. A PRA local unit groups contiguous NUTS4 units that have the same main agricultural products.

[3]We refer the reader to Baylis *et al.* (2008) for a comparison of agri-environmental policies in the EU and United States.

[4]Acronyms related to the considered AESs are defined in Appendix 1.

[5]It is beyond the scope of this paper to include a complete description of the assumptions underlying the Lesage (2000) procedure and the algorithms employed to implement it. We direct the interested reader to Train (2009) for a clear exposition of the MCMC approach and Holloway *et al.* (2002) for an accessible primer and sample application. Other possible estimation techniques are documented in Fleming (2004).

[6] A table with a complete statistical description of the variables is available upon request from the authors.

[7]Thus, the significance level of 0.006 when the NUTS4 unit with the highest shares is removed changes to 0.03 when the usual rule to detect outliers in a sample, i.e., to have a share greater than the 75% quartile plus 1.5 times the difference between this quartile and the 25% quartile, or 47%, is applied (4 NUTS4 units are removed) and to 0.09 when the six NUTS4 units with observed shares greater than 45% are removed. These units are mostly located in the same national park named "Parc du Cotentin".

[8]These results are available from the authors upon request.

[9]The dummy variable indicating that the farm is located in the natural park called "Parc du Cotentin" in Basse-Normandie (see the north-western part of Figure 1) had a significant impact on the probability of participating in an AES, reflecting the local park authority's policy, but its effect was negligible when compared to other covariates.

Appendix 1 - Definitions of acronyms

- PHAE, or "Prime Herbagère Agri-Environnementale" : Agri-environmental Payment for Extensive Grazing.
- ICHN, or "Indemnité Compensatoire d'Handicap Naturel" : Payments to farmers in mountain or others handicapped areas (including wetlands areas).
- PMSEE, or "Prime au Maintien des Systèmes d'Elevages Extensifs".
- CTE, or "Contrat Territorial d'Exploitation" : French territorial contract combining various AESs.
- CAD, or "Contrat d'Agriculture Durable" : Contracts following CTE.

Appendix 2 - Computation of Marginal Effects

In models containing spatial lags of the explanatory variables or dependent variables, interpretation of parameters becomes richer but also more complicated. This interpretation becomes even more complicated in case of discrete choice models for calculating the marginal effect of change in an explanatory variable on the conditional probability that $y = 1$. In the case of spatial Durbin probit models, these marginal effects can be decomposed into direct and indirect effects. Lesage and Pace (2009) and, more recently, Lesage *et al.* (2011) demonstrate how these effects can be computed in the case of spatial lag Probit models. To illustrate this point, consider the spatial Durbin model defined in (3), which can be rewritten as

$$y^* = (I_n - \rho W)^{-1}(X\beta + WX\delta) + (I_n - \rho W)^{-1}\varepsilon \tag{5}$$

or equivalently as

$$y^* = \sum_{r=1}^{k} S_r(W)x_r + V(w)\varepsilon \tag{6}$$

where x_r denotes the n by 1 vector of the observations for the rth explanatory variable, $S_r(W) = V(W)\ (I_n\beta_r + W\delta_r)$ and $V(W) = (I_n - \rho W)^{-1}$. For farmer i, equation (6) becomes

$$y_i^* = \sum_{r=1}^{k} S_{r,i}(W)x_r + V(w)_i\ \varepsilon \tag{7}$$

where $S_{r,i}(W)$ and $V(W)_i$ denote the i-th rows of the $S_r(W)$ and $V(W)$ matrices, respectively. The probability that farmer i contracts is

$$\text{Prob}\,(y_i = 1) = \Phi\left(\sum_{r=1}^{k} S_{r,i}(W)x_r\right) \tag{8}$$

Consider the impact on this probability due to a change in the r-th explanatory variable for farm j. This marginal effect can be measured as

$$\frac{\partial \text{Prob}\,(y_i = 1)}{\partial x_{r,j}} = \phi\left(\sum_{r=1}^{k} S_{r,i}(W)x_r\right)\ S_{r,ij}(W) \tag{9}$$

where ϕ denotes the probability density function of the normal distribution and $S_{r,ij}(W)$ denotes the j-th element of the i-th row of matrix $S_r(W)$. Note that this latter reduces to β_r if $j = i$, and to 0 if $j \neq i$ when $S_r(W) = I_n\beta_r$, or, in other terms, when no spatial effect are modeled, as in the standard Probit model. In this model, changes in the values of explanatory variables of neighboring farm j have no impact on the decision of farm i, while such spillovers are explicitly modeled in the spatial Durbin model.

Consider a matrix version of the problem of marginal effects in spatial Durbin models. Let $D(\phi(Q))$ denote the n by n diagonal matrix of which the (i,i)-th element is

$$\phi(Q_i) \equiv \phi\left(\sum_{r=1}^{k} S_{r,i}(W)x_r\right). \tag{10}$$

Then, the matrix of the marginal effects of the r-th explanatory variable, denoted by dP/dx_r, for which (i,j)-th element is $\partial\text{Prob}\,(y_i = 1)\,/\partial x_{r,j}$, can be expressed as

$$\frac{dP}{dx_r} = D\left(\phi(Q)\right)\ \left(I_n - \rho W\right)^{-1}\ \left(I_n\beta_r + W\delta_r\right) \tag{11}$$

The marginal effect due to a change in an explanatory variable can then be summarized in three ways, as proposed by Lesage and Pace (2009). The first is the average total effect on an observation. The row sums of the elements of the matrix dP/dx_r represent the total effect on each observation due to a unit change in the r-th explanatory variable across all the observations. It can easily be shown that the vector of these row sums, or cumulative effects, equals

$$d\left(\phi(Q)\right)\left(1-\rho\right)^{-1}\left(\beta_r + \delta_r\right)$$

where $d(\phi(Q))$ denotes the n by 1 vector of which the i-th element is $\phi(Q_i)$. Taking the average of the vector of cumulative effects yields the *average total effect*. The second impact is referred to as the direct effect, which is the effect of changes in the i-th observation on the decision probability. The *average direct effect* is measured by summing the trace elements of the matrix dP/dx_r in (11) and dividing the sum by the number of observations. The third effect is the *average indirect effect*, which constitutes feedback effects through neighbors. This last effect is the difference between the average total effect and the average direct effect.

When contextual variables are introduced in a spatial Durbin model, as in (4), their marginal effects can be expressed as

$$\frac{dP}{dz_r} = D\left(\phi(Q)\right)(1-\rho)^{-1}\gamma_r \tag{12}$$

It can be easily shown that $\left(I_n - \rho W\right)^{-1} Z\gamma = (1-\rho)^{-1}Z\gamma$. The average total effects and average direct effects of contextual variables are thus equivalent.

Appendix 3 - Figures

Figure 1 – **Spatial Distributions of Observed and Predicted Participation Rates in Basse-Normandie**

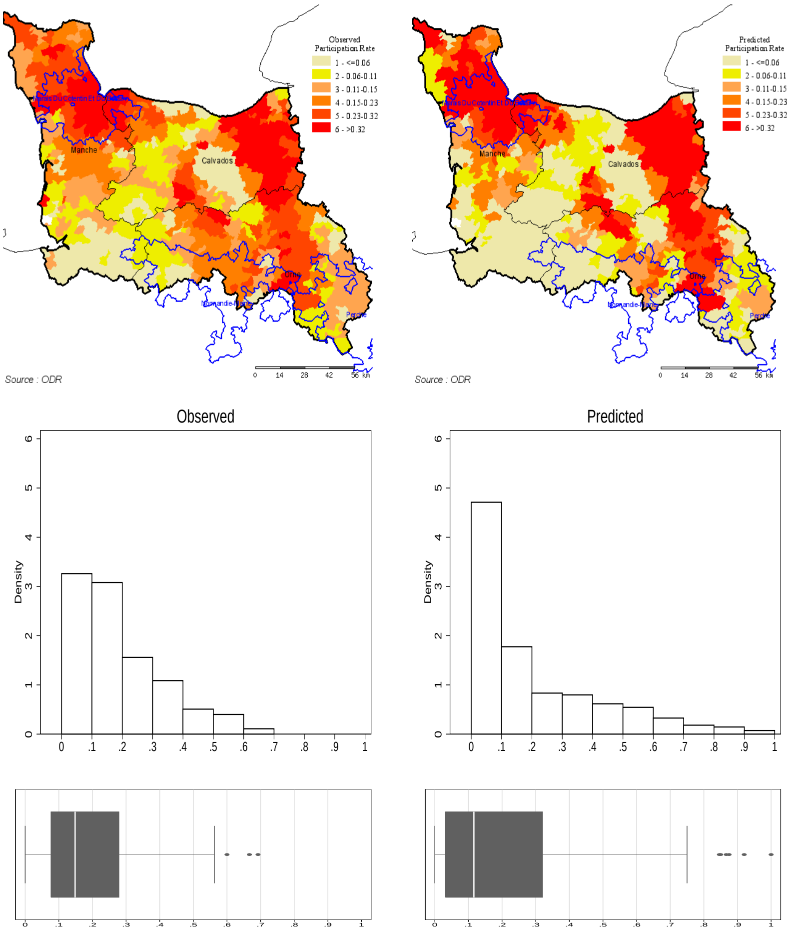

Figure 2 – **Spatial Distributions of Observed and Predicted Participation Rates in Auvergne**