ICAE
29th | Milan Italy 2015
UNIVERSITÀ DEGLI STUDI DI MILANO AUGUST 8 - 14
AGRICULTURE IN AN INTERCONNECTED WORLD

Spatial Heterogeneity in Production Functions Models

Billé AG.[1], *Salvioni C.*[2], *and Benedetti R.*[2]

[1] *Department of Economics and Finance, Tor Vergata University, Rome, Italy*

[2] *Department of Economics, University of Chieti-Pescara, Pescara, Italy*

Abstract. *Controlling for unobserved heterogeneity is a fundamental challenge in empirical research, as failing to do so can introduce omitted variables biases and preclude causal inference. In this paper we develop an innovative method – the Iterative Geographically Weighted Regression (IGWR) method – to identify clusters of farms that follow a similar local production econometric model, taking explicitly unobserved spatial heterogeneity into account. The proposed method is the perfect combination of the GWR approach and the adaptive weights smoothing (AWS) procedure. This method is applied to regional samples of olive growing farms in Italy. The main finding is that the conditional global IGWR model fits the data best, proving that explicitly accounting for unobserved spatial heterogeneity is of crucial importance when modeling the production function of firms particularly for those operating in land based industries*

IAAE

1. Introduction

Although theory supports firms do not operate on a common production function (homogeneous technology), in most empirical literature a global production function is proposed that assumes production technology is invariant over space and across firms. As a matter of fact, especially in land based industries such as agriculture, the analysis of production often need to be more nuanced and account for the spatial variations in production technology arising from local-specific solutions that satisfy, for example, the environmental constraint of a geographical area. Assuming a common production function encompassing every sample observation, i.e. failing to recognize the geographical variations in technology, leads to biased estimates.

When it is possible to classify the sample observations into certain categories defined on the basis of a priori sample separation, for example the location in different geographical or administrative regions, then a production function can be estimated for each group of firms. These estimates can then be used to build a meta-production, which represents the envelope of the production points of the most efficient groups (Hayami and Ruttan, 1971; Mundlak and Hellinghausen, 1982; Lau and Yotopoulos ,1989; Battese and Rao, 2002). There are cases, though, in which the groups of firms sharing a common technology and the boundaries of the regions in which they operate cannot be defined on the basis of information known a priori. In this case cross firm heterogeneity can be modeled in the form of continuous parameter variation, such as in state-dependent production functions (Munlak, 2012). As a matter of fact the data referring to the state of nature are usually highly aggregated over space, hence of little use to discriminate between different states of nature at the farm level. An alternative is to specify and test a latent class model, also referred to as finite mixture model (Orea and Kumbakhar, 2004; O'Donnell and Griffith 2006;), which treat heterogeneity in technology as generated by a latent discrete distribution.

In this paper we extend to this latter stream of literature by taking explicitly into account the inherent spatial heterogeneity in technologies. Spatial heterogeneity refers to those cases in which the same stimulus provokes a different response in different parts of the study region. Variations in relationship over space are referred to as a particular case of non-stationarity. If a spatial non-stationary relationship is modeled using global models, possible wrong conclusions might be drawn. In order to overcome the biasedness due to spatial heterogeneity, we propose an innovative method, i.e. the Iterative Geographically Weighted Regression (IGWR), to identify spatial technology clusters, i.e. groups of firms using a common technology.

The rest of the paper is organized as follows. In the second session we present the methodologies used to identify the spatial clusters of olive growing farms that is the subsample of farms in which a single local econometric model is justified. In the third session we present an application to the olive growing farms in Italy. Finally, some concluding comments are provided.

2. Methodology

In this section we briefly describe the Iterative Geographically Weighted Regression algorithm. In order to account for local parameter estimates over space GWRs are typically used (Fotheringham et al., 2002). The basic idea behind local geographical estimations is that simple linear functions may fit well for observations close to site i, but they will probably be inappropriate when more distant observations are included. In other words, this means that simple econometric models might provide a reasonable approximation of the local estimate as long as we use the information in a group of observations close to i. Therefore, simple linear functions can be written to account for local parameter estimates in the following way

$$y = (\beta \otimes X)\mathbb{1} + \varepsilon, \quad \varepsilon \sim MVN(0, \sigma^2 I) \qquad [2.1]$$

where y is an n-dimensional column vector, β is an n by $k + 1$ matrix with i-th row $\beta_i = (\beta_{i0}, \beta_{i1}, \beta_{i2}, \dots, \beta_{ik})'$, $\otimes$ is a logical multiplication operator in which each element of β_i is multiplied by the corresponding element of the i-th row in X, $x_i = (x_{i0}, x_{i1}, x_{i2}, \dots, x_{ik})$, $\mathbb{1}$ is a $(k + 1)$-dimentional column vector of ones, and ε is an n-dimensional column vector. Then, a geographically weighted estimator for each observation is simply obtained by repeated Weighted Least Squares (WLS) estimations

$$\hat{\beta}_i^{GWR} = (X'W_iX)^{-1}X'W_iy, \ \ i = 1, \dots n \qquad [2.2]$$

where W_i is an n-by-n matrix whose off-diagonal elements are zero, while those along the diagonal denote the geographical weights $(w_{i1}, w_{i2}, \dots, w_{ij}, \dots, w_{in})$ of each of the n observed data for regression point i. We will therefore have n diagonal spatial weighting matrices and n sets of local parameter estimates that correspond to the marginal effects.

Given our aim of identifying spatial clusters, that is subsamples of observations (farms in our case) in which a single local econometric model is justified, we iteratively extend the GWR approach till when all the observations with similar beta coefficients will belong to the same homogeneous cluster. This is obtained by computing new weights in the main diagonal of W_i at each iteration (see Polzehl and Spokoiny, 2000) and then comparing the estimated beta coefficients in [2.2] by using distance criteria. During this procedure, it is important to provide enough observations for each group since those characterized by a scarce quantity will be automatically excluded from the analysis due to identification problems, and treat as outliers. This can be avoided by setting the number of observations sufficiently larger than the number of parameters to be estimated. It is worth noting that the more the regressors, the higher the minimum number of observations required to construct the spatial clusters.

Given a production function

$$y = (X)\beta + \varepsilon \qquad \varepsilon \sim MVN(0, \sigma_\varepsilon^2 I_n) \qquad [2.3]$$

where y is a n by 1 vector of the quantity produced , X is an n by k matrix of inputs, β is a k by 1 vector of coefficient estimates, ε is the column vector of independently and identically distributed normal error terms, to iteratively identify spatial clusters of farms we need first to specify the diagonal matrix W_i in [2.2]. For this purpose we use the bi-square kernel weighting function for the initial weights and the Gaussian kernel weighting function for the iterative procedure (i.e. to update the weights), which have the following forms

$$w_{ij}^{Bisq} = \left(1 - \left({d_{ij}}/{b}\right)^2\right)^2 I(d_{ij} < b); \; w_{ij}^{Gauss} = \exp\left(-0.5\left({d_{ij}}/{b}\right)^2\right) \qquad [2.4]$$

where w_{ij} is the weight given to observation j in constructing the estimate for observation i, d_{ij} is the Euclidean distance between observations i and j, $I(.)$ is the indicator function that equals 1 if the condition in parenthesis holds and b is the so-called bandwidth. The choice of the bandwidth value is crucial: nearby observations of i are defined by the value of b that, therefore, determines which observations receive weight in constructing the estimate for observation i. Moreover, the value of b determines how rapidly the weights decline with distance. In general, higher values of b put more weight on distant observations leading to results similar to those obtained by OLS. To solve the problem of choosing an optimal bandwidth value, b^{opt}, a cross-validation (CV) method has been proposed and still currently used, which minimizes the overall residual sum of squares obtained when observation i is deleted in forming its own forecasted value

$$b^{opt} = \min_b CV = \min_b \sum_{i=1}^{n} \left(y_i - \hat{y}_{\neq i}(b)\right)^2 \qquad [2.5]$$

(Fotheringham et al., 2002). Another important issue is that the bandwidth value can be defined as fixed or variable. A fixed bandwidth presumes a fixed distance in which all the observations located at lower-distances are taken into account for the neighborhood definition. On the contrary, a variable or adaptive bandwidth assumes a fixed number of neighboring observations, so defining a k-nearest neighbor approach. In spatial economic analyses it has been suggested to use an adaptive bandwidth (McMillen and Redfearn, 2010). Thus, in this paper we adopt an adaptive bandwidth and we define, for the used kernel weighting function, the optimal variable bandwidth (b^{opt}) by using the minimizing CV criterion in [2.5]. The kernel function used to calculate the bandwidth value for each data set is the same kernel chosen to define the initial weights (i.e. the bi-square function). Moreover, we define the distance measure as a simple Euclid distance between farms in space[1].

Once obtained initial GWR estimates we can start the iterative procedure. The IGWR algorithm is based on different steps. The first step starts with the definition of a starting weighting vector, $w_{ij}^0 = K(d_{ij}; b)$, which is a Bisquare kernel function of the form in [2.4] based on both the distance between two units in space, d_{ij}, and the optimal bandwidth value obtained by the combination of the model in [2.3] with the criterion in [2.5]. The initial parameter and variance estimates, $(\hat{\beta}_i^0, \hat{\sigma}_\varepsilon^0)$, are then calculated by using the GWR model in [2.1] and the production function in [2.3]. From the second step and until the condition $\max\left|w_{ij}^{l-1} - w_{ij}^l\right| < \omega \; \forall ij, i \neq j$ holds, with ω a fixed small value, updated weights w_{ij}^l are calculated for each iteration l as follows. At the same iteration l the GWR parameter estimates $(\hat{\beta}_i^l, \hat{\beta}_j^l) \; \forall ij, i \neq j$ are compared by using the following quadratic forms

$$T^2{}_{ij}^{l} = (\hat{\beta}_i^l - \hat{\beta}_j^l)' (\Sigma_p)^{-1} (\hat{\beta}_i^l - \hat{\beta}_j^l) \qquad [2.6]$$

---

[1] The package *GWmodel* in R (Lu et al., 2014) was useful to obtain the optimal bandwidth values. We needed only to specify an Euclid distance matrix, the used kernel function for the initial weights and the same model as in [2.3].

where Σ_p is the pooled variance-covariance matrix obtained as a weighted average of the two variance-covariance matrices, i.e. $\Sigma_p = \left[tr(W_i)\Sigma_i + tr(W_j)\Sigma_j\right]/\left[tr(W_i) + tr(W_j)\right]$, and $(\hat{\beta}_i^l, \hat{\beta}_j^l)$ are $(k+1)$-dimensional column vectors. The updated weights are then calculated as a product of the previous weights with the updated ones,

$$w_{ij}^l = K(d_{ij}^l; b)K(T_{ij}^l; \tau) \qquad [2.7]$$

where $d_{ij}^k = d_{ij}/k$ in order to guarantee that further iterations do not decrease in the estimation accuracy, whereas τ is a parameter that scales the value of ${T^2}_{ij}^l$. In particular, the second factor in equation [2.7] is equal to $K\left({T^2}_{ij}^l; \tau\right) = \exp\left(-0.5\left(T_{ij}^k\tau\right)^2\right)$. The choice of τ is usually arbitrary and depends on the weight given from researchers to the criteria in [2.6]. The higher is the value of τ, the higher is the sensitivity to structural changes and therefore the number of clusters over space. Generally, a sensitivity analysis is required. Polzehl and Spokoiny (2004) found different default values by using Monte Carlo simulations and showed that this value is a function of the α-value. Finally, in order to stabilize the convergence procedure, the weights w_{ij}^l are then re-updated by averaging them with those obtained in the previous step

$$\breve{w}_{ij}^l = (1-\eta)\breve{w}_{ij}^{l-1} + \eta w_{ij}^l \qquad [2.8]$$

where $\eta \in (0,1)$ is a control parameter (also called memory parameter in Polzehl and Spokoiny, 2004) which is equal to 0.5.

3. An Application To Olive Farms In Italy

We use the above technique to examine the production of olives in Italy. In the Mediterranean area, Italy represents the central point of olive production, either in terms of history and environmental conditions, and it ranks second in the world after Spain for olive-oil production. The Italian olive sector is still characterized by a large number of small operations, more in detail Italy has the highest number of holdings (776 000) with the smallest average size (1.3 ha) in the EU Mediterranean countries. Over time the different microclimate conditions, soil formations and elevation have led to natural or man assisted modifications (breeding and selection) of the olive tree into many location-specific varieties, each with different productivity, agronomic needs and adaptability to irrigation and mechanization.[2] While in the case of arable crops, e.g. maize and oil seeds, local varieties have been widely substituted by industrial global varieties, whose production response does not vary much over space, on the contrary in the case of olive tree the local-specific varieties are still largely in use. The territorial anchorage of the production of location-specific olive varieties is further enforced by social and marketing considerations since, similarly to the case of wine and grapes, farmers choose the varieties to grow not only on the basis of their agronomic characteristics (disease resistance, climate preferences, high productivity, etc.) but also for their aptitude to preserve local knowledge (flavor, suitability for curing, etc.) and guarantee the production of high quality oil.

For example, in the case of olive trees farmers grow different cultivars that have different yields, aptitude to the mechanization of the harvesting and input needs. As we mentioned before the decision to grow a low-yielding or higher cost variety, hence not to operate on the most efficient production function, is partly connected to the characteristics of the local natural environment (climate, water, soils, etc.) and partly to the preservation of the local cultural heritage (flavor and quality of oil, landscape, etc.). Under these circumstances it is very difficult, when not impossible, to collect all the information needed to define the boundaries of the area within

---

[2] Two main methods of olive growing can be distinguished: traditional processes, generally used in mountainous or hilly areas which are not irrigated, and modern processes which involve elevated planting density, irrigation and mechanization. Over the last 30 years, the Italian olive growers have moved from traditional olive groves to new (i) medium-density plantation, adapted for mechanical harvesting; and (ii) high-density plantation adapted for full mechanization and especially for continuous harvesting using straddle machines. In both models (traditional and modern), the selection of varieties is fundamental to obtain the maximum productivity.

which a specific technology results efficient. Since the area, hence the groups, within which farms use the same production technology, are not known a priori, then the meta-production function is no more useful. A solution is offered by the application of the IGWR method, which allows us to identify local technology clusters of farms, that is groups of farms that follow a similar local production econometric model, by taking unobserved spatial heterogeneity explicitly into account.

3.1 Data

This study relies on cross-sectional data collected by the 2012 Italian Farm Accountancy Data Network (FADN) survey. The FADN sample is stratified according to criteria of geographical region, economic size (ESU) and type of farming. The field of observation is the total of commercial farms.[3] The survey gathers physical and structural data (i.e. location, crop areas, livestock numbers, labor force, etc.) and the economic and financial data needed for the determination of incomes and business analysis of agricultural holdings. It is worth noting that, starting from year 2009, the Italian FADN collects the geographic longitude and latitude of the farm, in this way allowing us to use spatial econometric models and, in our specific case, to account for farm spatial heterogeneity. Another advantage of the Italian FADN is that the information on input use and production results are collected by activity. As a consequence, differently from most of the previous SFA studies[4], we do not need rely on data referred to the whole farm production, hence we do not need to focus on farms highly specialized in olive

---

[3] In the FADN a farm is defined commercial when is large enough to provide a main activity for the farmer and a level of income sufficient to support his or her family.

[4] For example, in the case of olive production see Dinar et al. (2007) and Karagiannis (2009).

growing. The advantage of using information of farms either specialized and not specialized in olive growing is that the regional samples are large enough to perform separate regional analysis. The study focuses on three samples of olive-growing farms located respectively in the Marche (268) region, Apulia (270) and Tuscany (317). The dependent variable in the production function is the olive production measured in kilograms. The inputs included as explanatory variables are (a) land (measured in hectares), including only the share of utilized agricultural area devoted to olive-tree cultivation; (b) labor, comprising hired (permanent and casual) and family labor, measured in working hours; (c) capital, proxied by the hours of mechanical work employed in olive growing and harvesting; (d) other intermediate inputs including expenses for water, fertilizers, pesticides, fuel and electric power and other miscellaneous expenses, measured in euros, augmented with the expenses for contract work.

3.2 Empirical results

Here we present the results of the application of the IGWR approach to three samples of olive growing farms, the first referred to the Marche region, the second one to Tuscany and the third to Apulia.

We use the Cobb–Douglas (CD) functional form since the coefficients of this production function are easy to interpret. Furthermore, the use of the CD function avoids the multicollinearity problem that arises with more flexible functional forms such as the translog and the Fourier functional forms. In addition to this, flexibility is not an issue in our case given our coefficients are local specific.

More in detail, with the IGWR method we identified 7 spatial technology clusters of olive growing farms in the Marche region, 4 in Apulia and 2 in Tuscany (Fig. 1). The optimal bandwidth values suggested by the Cross-validation (CV) criterion in [2.7] are 174261 for Marche, 261 for Tuscany and 78 for Apulia, whereas the values of τ suggested by the sensitivity analysis performed on each sample of observations are 0.1, 0.01 and 0.001 for the Marche region, Tuscany and Apulia respectively.

We have computed the AIC statistics in order to select the model that fit the data best. The best model is the one with the lowest AIC. Both in Marche and Tuscany the model with spatially varying coefficient fits the data best, while in Apulia even if we find no evidence of spatial dependence the best fitting model is the spatial model with spatially varying parameters. These results give rise to the hypothesis that the effects of farm localization on production are mainly represented by heterogeneity in technology, and not by the presence of spatial autocorrelation.

These results allow us to conclude that accounting for unobserved spatial heterogeneity is an important issue when modeling the production function of olive growing farms. It is interesting to note that in all three regions the parameter estimates for all inputs are statistically significant in the global model, while in the model with spatially varying coefficients this is no longer true. For example, in the Marche region the coefficient estimate associated with labor is significant only in one cluster, while changes in levels of labor do not have a statistically significant impact on olive production in the remaining clusters. As for the other inputs, according to the global model a 1% increase in their use leads to approximately a 0.18% increase in output, while the results of the IGWR model show that in three clusters the output elasticity of other inputs is greater than 0.20%, while in other two clusters it is not statistically different from zero. In other word the results of the global model provide misleading signals since they fails to recognize the geographical variation in production technology.

Finally, in order to test the validity of our results, we overlaid the regional maps of the olive technological clusters to the maps of olive varieties obtained on the basis of a priori information available at the territorial level. It is very interesting to note that there is a high degree of overlapping between these two map types. The results of this cross-validation test show that farm geographic longitude and latitude can be used to proxy soil types, climate and variety variations, hence to account for the spatial heterogeneity in technologies arising from the adaptation of production techniques to variables that are not usually gathered by agricultural surveys at the farm level.

Conclusions

The article presents an empirical analysis of production function within the framework of heterogeneous technology with specific attention to the role of spatial non-stationarity in the identification of the production function and its impact on the choice of estimators. We suggest a new approach for the simultaneous estimation and stationary test of the regression models, to see whether or not the model parameters show to be spatially clustered and thus locally homogeneous.

The proposed partitioning algorithm is based on an iterative version of the geographically weighted regression methods to identify groups of farms, which obey a common production function. The proposed IGWR procedure explicitly allow for local parameter heterogeneity.

The framework is applied to estimate the production functions of olives on geocoded individual farm data as collected by the RICA sample survey in 2012 in 3 Italian Regions. The empirical results confirmed the existence of technological heterogeneity; therefore empirical analysis that fails to incorporate parameter heterogeneity can produce misleading results. In our empirical analyses we always identify at least two groups, with significantly different regression parameters. Once partitioned the study area, we observed that the spatial interaction between farms belonging the same cluster is not anymore significant, giving rise to the hypothesis that the effects of farm localization on production technology are mainly represented by heterogeneity, and not by the presence of spatial autocorrelation. The presence of these different production regimes is related to the existence of some latent, not observed, factors that are closely related to the spatial position of the observed farms as climate, soil types, varieties.. A specific feature of these factors is that they are locally approximately homogeneous, but they rapidly change as soon as a natural, social or economic border is crossed.

The results are encouraging even if some additional research is needed. In particular, the conditions under which the proposed partitions imply a better fit of the regression model, should be better investigated. Finally, the proposed strategy performs very well with large datasets, but we need to explore the computational burden when huge sample sizes are used.

References

Battese G. E., and Prasada Rao D. S. 2002. “Technology Gap, Efficiency, and a Stochastic Metafrontier Function.” *International Journal of Business and Economics* 1(2):87-93.

Fotheringham, A. S, Brunsdon C., and Charlton M. 2002. *Geographically Weighted Regression. The analysis of spatially varying relationships.* Chichester: John Wiley & Sons, Ltd.

Dinar A., Karagiannis G., and Tzouvelekas V. 2007. “Evaluating the impact of agricultural extension on farms' performance in Crete: a nonneutral stochastic frontier approach.” *Agricultural Economics* 36(2):135-146.

Hayami, Y., and Ruttan V. W. 1971. “*Agricultural Development: An International Perspective*.” Baltimore: John Hopkins University Press

Karagiannis G., and Tzouvelekas V. 2009. "Measuring technical efficiency in the stochastic varying coefficient frontier model,” *Agricultural Economics* 40(4):389-396.

Lau, L. J., and Yotopoulos P. A. 1989. “The Meta-Production Function Approach to Technological Change in World Agriculture.” *Journal of Development Economics* 31: 241-269.

Lu B., Harris P., Charlton M. and Brunsdon C. 2014. “The GW model R package: further topics for exploring spatial heterogeneity using geographically weighted models.” *Geo-spatial Information Science* 17(2):85-101.

Mundlak Y., Butzer R., and Larson D. 2012. “Heterogeneous technology and panel data: The case of the agricultural production function.” *Journal of Development Economics* 99(1):139-149.

Mundlak, Y., and Hellinghausen R. 1982. “The Intercountry Agricultural Production Function: Another View.” *American Journal of Agricultural Economics* 64:664-672.

O'Donnell C. J., and Griffiths W. E. (2006). Estimating State-Contingent Production Frontiers, *American Journal of Agricultural Economics* 88(1): 249-266.

Orea, L. & Kumbhakar, S.C., 2004. “Efficiency measurement using a latent class stochastic frontier model.” *Empirical Economics* 29(1):169–183.

Polzehl, J., Spokoiny, V. 2000. “Adaptive weights smoothing with applications to image segmentation.” *Journal of the Royal Statistical Society* 62, Series B: 335-354.

Polzehl, J., Spokoiny, V. 2004. Local Likelihood Modeling by Adaptive Weights Smoothing. Available at. http://opus4.kobv.de/opus4-matheon/frontdoor/index/index/docId/69.

Tables and Figures

Table 1. Marginal Effects – Marche Region

	Global Model	Model with spatially varying parameters	Spatial Global Model	Spatial Model with spatially varying parameters
Intercept	-0.519***	-	-0.086	-
Intercept (1)	-	-0.856**	-	-0.322
Intercept (2)	-	0.257	-	0.270
Intercept (3)	-	-0.547	-	-0.544
Intercept (4)	-	0.440	-	0.471
Intercept (5)	-	0.444	-	0.486
Intercept (6)	-	0.776˙	-	0.771˙
Intercept (7)	-	0.267	-	0.412
Land	0.433***	-	0.429***	-
Land (1)	-	0.515***	-	0.500***
Land (2)	-	0.499***	-	0.500***
Land (3)	-	0.210*	-	0.220**
Land (4)	-	0.373***	-	0.372***
Land (5)	-	0.882**	-	0.870***
Land (6)	-	0.729***	-	0.719***
Land (7)	-	0.297***	-	0.298***
Labour	0.089*	-	0.094*	-
Labour (1)	-	-0.061	-	-0.056
Labour (2)	-	0.016	-	0.021
Labour (3)	-	0.395**	-	0.411***
Labour (4)	-	0.064	-	0.065
Labour (5)	-	-0.113	-	-0.101
Labour (6)	-	0.172	-	0.160
Labour (7)	-	0.029	-	0.007
Capital	0.062˙	-	0.064˙	-
Capital (1)	-	0.144	-	0.154˙
Capital (2)	-	0.125	-	0.120
Capital (3)	-	-0.089	-	-0.094
Capital (4)	-	0.098	-	0.096
Capital (5)	-	0.069	-	0.063
Capital (6)	-	-0.230	-	-0.209
Capital (7)	-	0.164	-	0.180˙
Other inputs	0.179***	-	0.180***	-
Other inputs (1)	-	0.230***	-	0.240***
Other inputs (2)	-	0.187*	-	0.185**
Other inputs (3)	-	0.360**	-	0.341**
Other inputs (4)	-	0.186**	-	0.185**
Other inputs (5)	-	0.039	-	0.032
Other inputs (6)	-	-0.024	-	-0.022
Other inputs (7)	-	0.311***	-	0.311***
rho	-	-	-0.168	-0.213
lambda	-	-	0.453˙	-0.442
AIC	290.215	279.002	292.169	282.016

Note: Asterisks *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 2. Marginal Effects – Tuscany Region

	Global Model	Model with spatially varying parameters	Spatial Global Model	Spatial Model with spatially varying parameters
Intercept	-1.641***	-	-1.478***	-
Intercept (1)	-	-2.791***	-	-3.229***
Intercept (2)	-	1.304*	-	1.380**
Land	0.167***	-	0.168***	-
Land (1)	-	0.108	-	0.119
Land (2)	-	0.175***	-	0.166***
Labour	0.273***	-	0.281***	-
Labour (1)	-	0.277*	-	0.244
Labour (2)	-	0.309***	-	0.327***
Capital	0.179***	-	0.176***	-
Capital (1)	-	0.182·	-	0.173·
Capital (2)	-	0.168**	-	0.153**
Other inputs	0.337***	-	0.336***	-
Other inputs (1)	-	0.570***	-	0.595***
Other inputs (2)	-	0.281***	-	0.275***
rho	-	-	-0.056	0.120
lambda	-	-	-0.028	-0.974
AIC	565.735	560.494	569.472	562.319

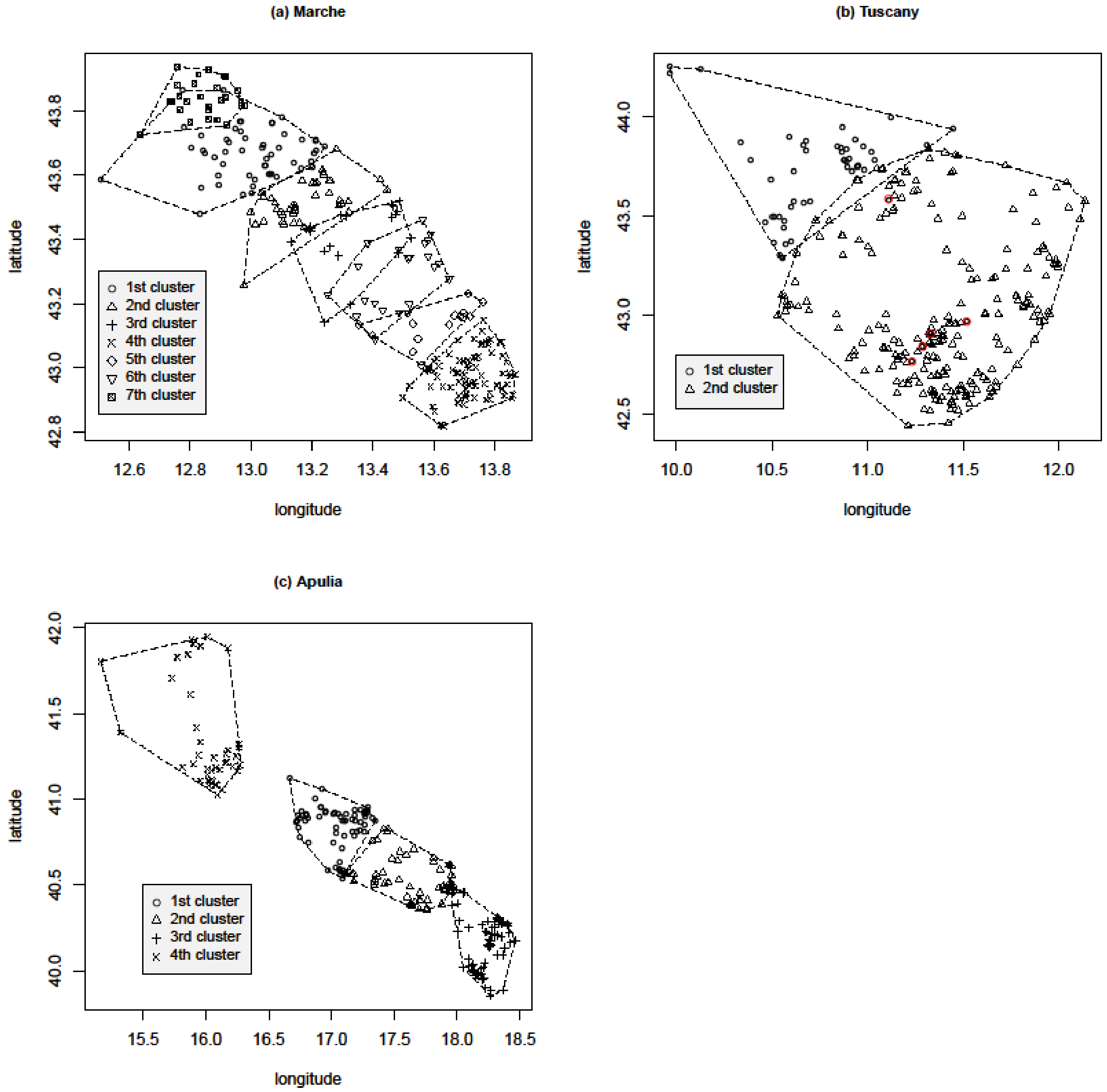
Note: Asterisks *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Table 3. Marginal Effects – Apulia Region

	Global Model	Model with spatially varying parameters	Spatial Global Model	Spatial Model with spatially varying parameters
Intercept	-1.318***	-	0.428	-
Intercept (1)	-	-2.230***	-	1.784
Intercept (2)	-	0.836	-	0.905·
Intercept (3)	-	1.332*	-	1.349*
Intercept (4)	-	1.081	-	1.442*
Land	0.056·	-	0.059·	-
Land (1)	-	-0.033	-	-0.031
Land (2)	-	0.140·	-	0.089
Land (3)	-	0.033	-	0.047
Land (4)	-	-0.075	-	-0.097
Labour	0.378***	-	0.412***	-
Labour (1)	-	0.503***	-	0.509***
Labour (2)	-	0.405**	-	0.422*
Labour (3)	-	0.236	-	0.235·
Labour (4)	-	0.564*	-	0.399·
Capital	0.217***	-	0.192***	-
Capital (1)	-	0.009	-	0.013
Capital (2)	-	0.297·	-	0.324*
Capital (3)	-	0.278*	-	0.290**
Capital (4)	-	0.153	-	0.081
Other inputs	0.346***	-	0.346***	-
Other inputs (1)	-	0.582***	-	0.566***
Other inputs (2)	-	0.221***	-	0.222***
Other inputs (3)	-	0.373***	-	0.359***
Other inputs (4)	-	0.319*	-	0.543***
rho	-	-	-0.369·	-0.812
lambda	-	-	0.203	-2.005
AIC	524.231	514.591	524.947	504.679

Note: Asterisks *, **, and *** denote significance at the 10%, 5%, and 1% levels, respectively.

Figure 1. Spatial clusters of olive growing farms



Note: Bisquare kernel function for initial weights and Gaussian kernel function for the updated ones are used.

Adaptive bandwidths are based on the CV criterion: (a) $b_{knn}^{CV} = 174$, (b) $b_{knn}^{CV} = 261$, (c) $b_{knn}^{CV} = 78$.