International Conference on Regional and Urban Modelling

Brussels, June 1-3, 2006

THE PRICE OF THE RUSTIC LAND IN A PROTECTED NATURAL AREA: ECONOMETRIC ANALYSIS IN THE CASE OF THE URDAIBAI BIOSPHERE RESERVE

Inmaculada Astorkiza Ikazuriaga* (e-mail: inma.astorkiza@bs.ehu.es)

Ana Ferrero Rodríguez* (e-mail: ana.ferrero@bs.ehu.es)

Patricia Abelairas Etxebarria* (e-mail: ebbabetp@bs.ehu.es)

*Dept. of Applied Economics V

University of the Basque Country (UPV-EHU)

Av. Lehendakari Agirre, 83

48015 – BILBAO

SPAIN

Tel.: +34 - 94 601 70 96

Abstract

This paper has the objective of explaining the variables that have determined the evolution undergone by the prices of the rustic land in the last years in a Protected Natural Area. The high prices of the land in the urban zones and the high valuation of the countryside as a residence place, are increasing the demand of rustic land with residential aims, whereas the expectable purpose is the agrarian one. In the Urdaibai Biosphere Reserve, the pressure is greater due to its natural qualities and to her more restrictive legislation when controlling the uses of these zones. These conditions are causing tensions between the supply and the demand of rustic land in Urdaibai and they are originating non desirable distortions in the market. A GIS database with geographically located lands has been created to make a spatial econometric analysis that clarifies the factors that originate the prices of the rustic land market.

Keywords: land use, rustic land price, protected natural area, spatial econometric model

1.- Introduction

The main objective of this paper is to explain the variables that have determined the price evolution of rural land in a Protected Natural Area, such as the Urdaibai Biosphere Reserve (RBU) located in the province of Bizkaia. The interest in this type of land is partly due to the urban sprawl phenomena that is increasingly becoming common in areas relatively close to major cities. The problems of the high housing prices in urban zones together with the high desirability of the countryside as a place of residence are the reasons for this type of phenomena. A segment of urban dwellers are looking to areas not far from the city, but with ecological and landscape qualities and in short, a quality of life much higher than that enjoyed in cities. These rural areas close to social and economic centres are witnessing how the prices of their land are progressively increasing and, therefore, of the dwellings built there. Despite the steady increase of prices in these non-urban zones, the rural areas continue to register a significant differential with the over-valued prices of city dwelling. This price differential together with the current boom of the rural world means that these zones are witnessing a significant increase in the demand for land.

On the other side of the market, the land supply in rural areas is being highly controlled by public policies in general, as there is a tendency from different spheres to preserve the values still to be found in the environment. Thus, territorial policies reflect this trend with strict land use planning and policies. The restrictions imposed by current legislation regarding the functions that can be developed on land with interesting characteristics from the landscape, ecological or agrological points of view, for example, greatly limit the supply to an extent.

The two circumstances, explained on the one hand by the growing demand for land in rural area and on the other hand, by the supply restrictions for the sake of defending the rural and natural environment, require the variables leading to the raising of the prices of rural land to be analysed.

The unique features of our study area, the Urdaibai Biosphere Reserve, hone the aforementioned trends even further for various reasons. First of all, the characteristics of this Natural Area are even more appreciated than those of the majority of rural areas and its location is exceptional as it is only 30 minutes away by car from the Greater

Bilbao area. Secondly, the fact that it was declared a Biosphere Reserve by the UNESCO (1984) meant that specific legislation was enacted that zealously protects the activities carried out on the land that form this natural enclave.

By analysing the land market of the RBU rural areas and of various nearby rural towns, this study aims to explain the variables that affect the land prices of these non-urban areas. The paper is therefore organised as follows. The following section considers studies that have tried to answer the causative factors of the evolution of land prices from the perspective of spatial econometrics. The third section provides an overview of the database specifically created for that purpose, each of the variables that have been included in the model and the methodology used for the analysis. The following epigraph sets out the main results of the spatial econometric study beginning with the spatial data visualisation techniques, followed by an exploratory analysis where the existence of spatial autocorrelation will be observed. The confirmatory analysis will then be carried and will confirm the existence of spatial autocorrelation and its typology. The specific spatial autocorrelation of the model will then be introduced in the basic model used as the starting point and resulting in the final model. The different aspects of the model will be interpreted according to the results of the estimation of the full model. The paper ends with the conclusions extracted from the obtained results.

2.- Background

The analysis aims to explain the evolution of the price of land that was initially reserved for agrarian purposes. The fact that the primary sector and the population of the majority of the rural towns in the RBU are declining, as verified by Murua et al. (2001), would lead us to presume that the pressure on the property market is small or non-existent. However, the situation is truly different and residential development in those zones is becoming more important. As has already been explained, those towns and their rural land are increasingly attractive for building first and second homes due to their proximity to densely populated areas, the attraction of living in an exceptional natural setting, the improvement to its infrastructures and the price differential existing with respect to similar dwellings in urban municipalities. The majority of the land transactions analysed, even though they are classified as rural, are for high prices that do not reflect the agrarian income that could be obtained from them but rather reveal their residential purpose, as the price of land reflects what its use is or what can be potentially developed there (Plantinga, 2002).

The interest of the study stems from three points: first of all, the dual consideration of those plots as the land studied is classified as rural but the majority of it is not used for agrarian purposes. This means that it is not possible to apply the same techniques or models that are applied to the prices of a rural zone in the true sense of the term where only the agrarian, livestock or forestry uses determine the final price of each piece of land. The apparent residential goal of those rural land transactions recommend that they should be treated from a real estate perspective.

Secondly, it is of interest to describe the environment on which the study is centred, a Protected Natural Area close to a densely populate area, like the Great Bilbao metropolitan area. There are very few studies focusing on the land market of periurban and Protected Natural Areas.

Finally, the study analyses a spatially referenced database with variables such as the appraised price, its surface area, typology, location, etc., which require a spatial treatment using spatial econometric techniques. This recently created and developed analysis technique means that only publications from the last year or so referring to the land prices use this technique.

The papers that have explained the prices of strictly agrarian land include the one by Plantinga et al. (2002) where they estimated a cross section model to determine the values of agrarian land in order to be able to understand how the present prices of the agrarian land are influenced by its future development. They used a spatial error model that they applied to their spatially referenced database.

Special mention should also be made of the Patton and McErlean study (2002), where starting from a hedonic price model and using spatial econometric techniques, they arrived at a spatial lag model and confirmed the existence of spatial heterogenity. They set out the squared inverse of the distance as the weights matrix and the starting functional form is log-log. They concluded that the price of the agrarian land is not only determined by the characteristics inherent to the plots of land, but the price of the nearby land also significantly affects it.

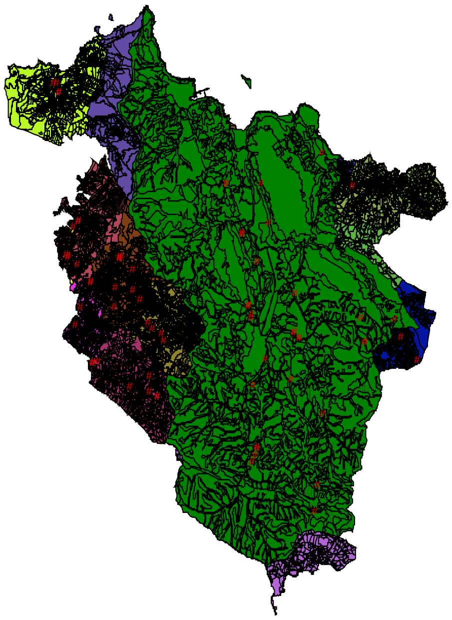
With respect to studies that analyse the land prices in an urban sphere, special mention should be made of Páez et al. (2001) which analysed the prices of the spatially referenced land in the city of Sendai (Japan). It used variables relating to the distance to various urban centres as explicative variables. It performed an autoregressive econometric model using a contiguity matrix based on the distance matrix (this type of matrixes are interesting for non-adjacent but nearby observations), which concluded that the city studied is mainly monocentric.

3.- Database and model

The study is based on analysing the prices of the land transactions located in the rural zone of the municipalities of the Urdaibai Biosphere Reserve and of five similar and nearby towns, but which do not belong to the RBU.

The database, developed by author, began by gathering the data of transactions recorded between 1992 and 2003 in the municipalities in question. Variables such as the surface area of the plot, the type of land, its price and the year of the sale were extracted. As far as the price is concerned, it should be emphasised that the prices used as those that the credit institutions and banks stipulate as "Appraised Price" as it is the best reflection of the market price. During a second phase, the location of the land studied in the rural land plots of each town was gathered from the Cadastre and Valuation Section of Bizkaia Provincial Council. Finally and using a GIS program (ArcView 3.2), these land plots were overlapped with the transacted plots of interest and zoning of the RBU (in green on the map below) and the municipalities outside Urdaibai (in various colours on the map). Each piece of land, which is represented by a red point on the map, was associated in this cartographic information, with all the variables being studied.

Illustration 1 GRAPHIC DATABASE IN GIS FORMAT



Source: By author.[1]

Even though in principle there were 725 land sales observations, for various reasons, including the lack of appraised price or the impossibility to locate the plot, the observations that have all the characteristics required for them to be incorporated in the spatial econometric analysis were reduced to 78.

The model used is based on a log-linear uniequational demand model. As far as the supply is concerned, the main feature is its total rigidity as, on the one hand, the orography limits it to a great extent and, on the other hand, the general and specific legislation determines its zoning and the uses permitted in each one. Therefore, the econometric model that best collects the information supplied by the variables is the following:

$$PRICE_i = AREA_i^{\beta 1} . DATA_i^{\beta 2} . \beta_3 DISTN_i^{\beta 3} . GDP_i^{\beta 4} . MINLOT_i^{\beta 5} . NRNONR_i^{\beta 6} . RESERVE_i^{\beta 7} . LANDTYPE_i^{\beta 8} . CONS_i^{\beta 9} . \mu$$

[1] All the illustrations and tables herein were prepared using the data supplied by the Gernika Land Registry, the RBU Board and by the Casastre and Appraisal Unit of Bizkaia Provincial Council's Treasury and Financial Department.

To which we apply logarithms to linearize it:

$$lnPRICE_i = \beta_1 lnAREA_i + \beta_2 lnDATA_i + \beta_3 lnDISTN_i + \beta_4 lnGDP_i + \beta_5 lnMINLOT_i + \beta_6 RCNORC_i + \beta_7 RESERVE_i + \beta_i\, LANDTYPE_i + \beta_9 CONS_i + \mu$$

The definition of the variables is as follows:

$PRICE_i$ = real appraised price in Euros per m^2 of land i
$AREA_i$ = m^2 of land i
$DATA_i$ = transaction year of the plot of land i
$DISTN_i$ = distance in metres to the centre of the municipalities of land i
GDP_i = GDP per capita in Euros in Bizkaia at the time of land i transaction
$MINLOT_i$ = minimum surface area (m^2) required to build in the zone where land i is located
$RCNORC_i$ = variable dummy “Rural Centres” (1) or “Non-rural Centres” (0)
$RESERVE_i$ =variable dummy “RBU” (1) or “Non RBU” (0)
$LANDTYPE_i$ = variable dummy “Farm Property” (1) or “Non Farm Property” (0)
$CONS_i$ = variable dummy “Built Land” (1) or “Non-built Land” (0)

4.- Spatial econometric analysis

Working with spatial data has some specific characteristics that prevent the use of some conventional econometric techniques. Spatial econometric techniques attempt to modelize data that are spatially located and that therefore usually raise problems that are not dealt with in general econometrics, such as heterogenity or spatial autocorrelation. Heterogenity occurs when data of very different spatial units are used to explain a single phenomena and does so in the form of heterocedasticity or structural instability. Spatial autocorrelation appears provided that the value of the variable in a place of the space is related to its value in other places of the space and, this problem only has a solution within the context of spatial econometrics. These techniques began with exploring the data in order to check if there is spatial autocorrelation using the Exploratory Analysis in order to then use the Confirmatory Analysis to confirm its existence, determine its typology and enter it in the regression model in such a way that the spatial autocorrelation does not distort the results or conclusions obtained.

It was during the Seventies when the term Spatial Econometrics first appeared in the work of Paelink and Klaasen (1979) referring to the techniques to study spatial autocorrelation in error term. Subsequently, Anselin (1988) defined the term as the techniques that deal with the unique features resulting from the space in the statistical analysis of the regional science models. During the 80s and 90s, great progress was made in spatial econometric techniques. The work of Cliff and Ord (1981), Blommestein (1983) and Anselin (1980, 1988) analysed the most important methodological aspects for the first time. Since then, journals have been covering the new contributions to this field. The fact that spatially referenced databases are available, the development of the computer tools that enable the systematic spatial treatment of the data and the growing interest in the space and the spatial interaction in regional science has resulted in increasingly greater spatial econometrics. As far as the computer tools used in this field are concerned, special mention should be made of Anselin's great contribution (1992) when he implemented the *SpaceStat* program where the different spatial autocorrelation detection contrasts in the econometric models and the appropriate estimation methods given the presence of that dependency operate under Gauss. As well as developing this software, Anselin also developed extensions to GIS programs such as *ArcView* that allow the results obtained in the *SpaceStat* econometric package to be viewed in space.[2]

4.1.- Weights Matrixes

When autocorrelation appears in the time context, this occurs unidirectionally and the observations have some type of time interrelation but in a single direction. In the spatial case, however, the dependency between the observations or autocorrelation is more complex as it is multidirectional, in other words, each one of the observations may be

[2] Anselin’s last contribution as regarding spatial econometrics software is the GeoDa. program. For more information : https://www.geoda.uiuc.edu/default.php

related to various neighbouring observations[3] and also, be so in a different way to each one. Spatial econometrics has provided a solution to this problem by means of creating to the so-called Weights Matrix or W Matrix which is a squared matrix of an identical side to the number of observations. The elements that form that matrix outline the manner in which each observation is related with the others. The most common weights matrixes are the binary contiguity that assigns an 1 to the neighbouring observations and a 0 to those that are not, but there are other types of weights matrixes for databases with non-bordering observations such as economic distances or the matrix proposed by Anselin that assigns a weight to the interrelation of the observations by means of the squared distance inverse.

The weights matrix allows the so-called Spatial Lag Operator to be calculated, which comprises a weighted average of the values in the neighbouring locations (with the term neighbouring being taken in the widest sense), with exogenous and set weightings. The lag operator is obtained as the product of the spatial weights matrix by the observations vector of a random variable. Each element of the spatial lag operator is the weighted average of the values of the variable in the neighbouring observations sub-group. If the spatial weights matrix is standardized by rows, the spatially lagged variable would represent the softened neighbouring values, given that the sum of all the weights of a specific row is equal to 1 (Moreno and Vayá, 2000).

In this analysis, various matrixes have been used despite the fact that the article considers the results associated to the inverse distance matrix due to the need for synthesis. This matrix was chosen due to the spatial distribution of the data, as the prices of plots of land will have a closer relationship the closer they are, in other words, each weight between observations i and j will be inversely proportional to the squared distance that separates them, in other words, when the distance is increased, the weight of its interrelation will reduce. The inverse matrix of the distance will be designate in this study as RSINVER (Row Standardised Inverse) and is defined as follows:

$$W_{ij} = 1/d_{ij}^2$$

[3] The term "neighbouring" is frequently used in the use of spatial econometric techniques to describe regions that are adjoining or which according to the determining criterion, we consider them as nearby in the widest sense of the world and not only referring to physical proximity.

Other matrices entered in the model were the contiguity matrices based on a distance matrix that allocates a 1 to observations that are at an equal or lower distance than the one that means that each observation has a determined number of neighbours (5, 6 or 7 specifically) and a zero for those that are at a great distance (k-nearest neighbour weights). The results have also been checked that were obtained from their relevant second order matrixes. All the matrixes used were standardised to make it easier to carry out statistics and comparisons.

4.2.- Spatial Data Econometric Analysis (AEDE)

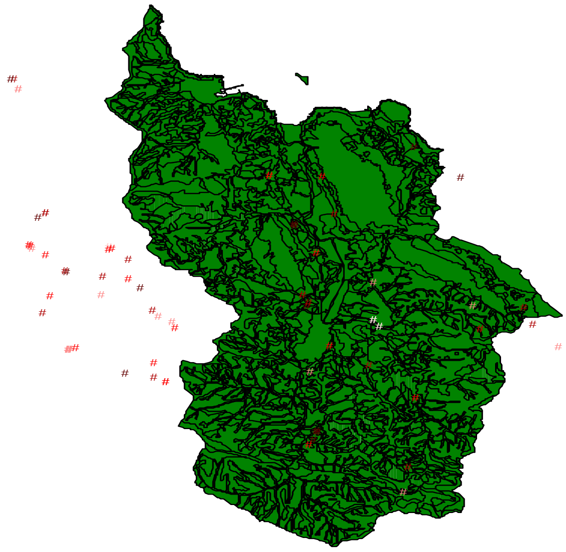
Spatial Data Econometric Analysis consists of a series of techniques that tries to describe spatial distributions, spatial outliers, spatial clusters and different spatial systems or forms of spatial instability. Therefore, before considering the exploratory analysis as such, a series of spatial data visualisation techniques are applied and which begin to throw light on the existing spatial relations. Then and as part of the exploratory analysis, the spatial global association is firstly studied before moving on to the analysis of the spatial local association.

The global approach studies the existence of spatial dependency summarised in indicators such as Moran's I, Geary's C or the G of Getis and Ord. The local indicators, known as LISA, Local Indicator of Spatial Association, (Anselin, 1995) are those that achieve two objectives: that the values of the information statistics regarding the relevance of a spatial grouping of similar values close to the observation and that the sum of the value of the statistic for all the observation is proportional to a global indicator of spatial association. The local contrast of Moran's I and the local contrast of the G of Getis and Ord are situated within this type of indicators.

The spatial econometric analysis begins with the spatial data visualisation techniques. Illustration 2 depicts the Distribution Map of the Variable to explain LNPRICE where the spatial association is not clear.

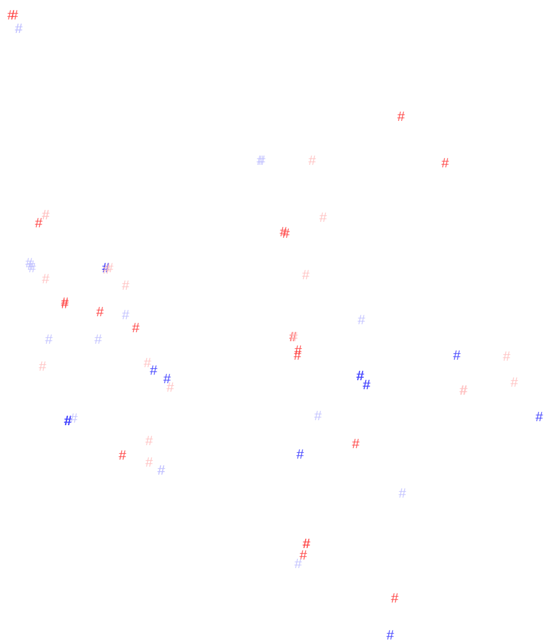
The BoxMap is then prepared to identify spatial outliers. The BoxMap is an extension of the map that represents the quartils together with the upper and lower outliers, in other words, observations outside the limits of a BoxPlot. The BoxMap of the LNPRICE variable indicates that there are no spatial outliers. In order to complete the spatial data visualisation techniques, illustration 4 depicts the spatial lag graphic. This graphic shows the value of the variable and the value of the spatially lagged variable or what is the same, the weighted average of the variable in the neighbouring observations using the weights matrix. If these values are similar, it indicates that the values of the variable at a point is similar to that of the variable at neighbouring points, in other words, it indicates the presence of spatial autocorrelation. As can be noted in the spatial lag graphic of the LNPRICE variable, the existence of spatial autocorrelation begins to be discerned in the prices of the plots of land both inside and outside the Reserve.

Illustration 2 SPATIAL DISTRIBUTION MAP OF THE LNPRICE VARIABLE



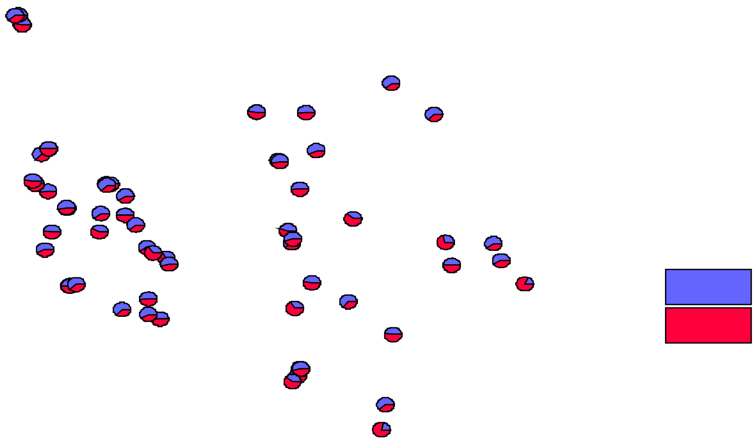
Source: By author.

Illustration 3 BOXMAP OF THE LNPRICE VARIABLE



Source: By author.

Illustration 4 GRAPHIC OF THE SPATIAL LAG OF THE LNPRICE VARIABLE WITH THE RSIVER MATRIX



Source: By author.

The Global Exploratory Analysis globally detects spatial dependency. The Moran's I (Moran, 1948) and Geary's C (1954) contrast, which have the lack of spatial dependence, (i.e., a random distribution), as a void hypothesis, compared to the alternative hypothesis of the presence of spatial autocorrelation, are widely used for that purpose. Another interesting statistic is the G of Getis and Ord (1992) which contrasts the type of autocorrelation that exists. If the value of the statistic is positive and significant, there is a high value association. On the other hand, if the value is negative and significant, there will be a low value association and finally, if the value is not significant, there is no spatial association. Together with these statistics, the ScatterMap is analysed. It uses different colours to classify each observation according to the value that the variable has and the lagged variable at that point. In this map, the observations are situated geographically using different colours according to 4 categories: high-high (high value positive autocorrelation), low-low (low value positive autocorrelation), high-low (negative high and low value autocorrelation) and low-high (negative low and high value autocorrelation). If any colour is clearly predominant, it will indicate the presence of the type of autocorrelation associated to that colour.

As can be seen from Table 1, the statistics clearly show that there is positive spatial autocorrelation (high values with high values or low values with low ones) as the value of Moran's I is positive and of Geary's C is negative and the two are highly significant.

Table 1 GLOBAL STATISTICS OF THE LNPRICE VARIABLE[4]

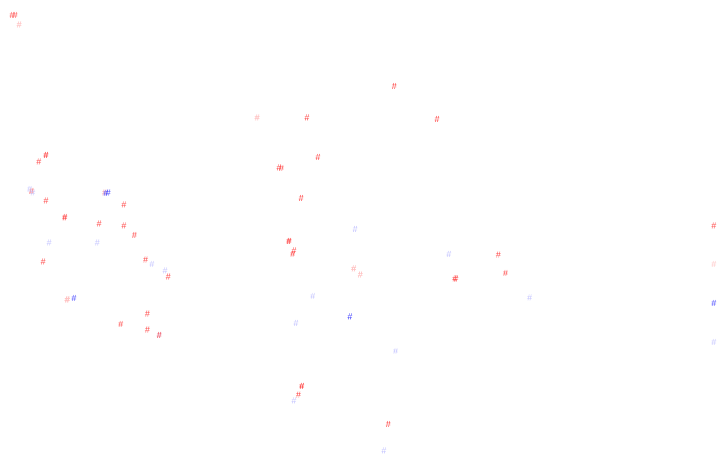
	Z (I)	Prob.
I de Moran	6.8717	0.0000
C de Geary	-5.4598	0.0000

Source: By author.

In the ScatterMap, it can be seen that red is the predominant colour followed by pink. That means that the majority of the land prices have similar prices to the surrounding area. In other words, they depict a positive spatial association. As the map illustrates, the most frequent is that the land has high prices and that the average of its "neighbours" (spatially lagged variable) is also high.

In the Global Exploratory Analysis, it can be discerned that there is positive autocorrelation in the LNPRICE variable. The following step will be the Local Exploratory Analysis where the existence of similar land price concentration in the space is considered.

Illustration 5 MORAN SCATTERMAP OF THE LNPRICE VARIABLE



Source: By author.

[4] The G of Getis and Ord is not considered as when the variable is studied in logarithms, negative values appear and they are not accepted by the calculation to find that statistic.

The Global Exploratory Analysis tests are not sensitive to situations where there are clusters or area/region groups that have a concentration or high or low values located in specific areas of the territory. In order to perform the Local Exploratory Analysis, the Local Moran's I, which considers the existence of region groups, is first of all calculated. It has the absence of spatial association as void hypothesis. Therefore, if the statistic is positive and significant, it indicates that there are similar value clusters of the variable analysed in region i. If, on the other hand, the statistic has a negative value, the concentration will be of dissimilar values. Once the statistic is calculated, the results can be depicted on a signification map where we see how the clusters are located on the map and spatial distribution, as well as obtaining the typology of each cluster. Another interesting graphic in this part of the analysis is the one representing the New-Gi statistic. This graphic detects concentrations of similar values and if they are high or low values.

When calculating the Local Moran's I statistic for the LNPRICE variable, observations 2, 3, 4, 5, 6, 41, 54, 57, 58, 59 and 60 have similar value clusters while observation 43 has a group of dissimilar values, or what is the same, high values with low values or vice versa. On the signification map, we see how the clusters or local spatial groups are located that are significant and that only a small number of them appear. On the distribution map of the Moran's I contrast of the LNPRICE variable, it is noteworthy that the majority of the significant clusters are low with low values. High with low clusters also appear as significant. On the New-Gi contrast map, we corroborate that the clusters that appeared in the Local Moran's I are significant and low with low values (it can be seen in the New-Gi statistic values table that this last term has not been included)

Following the Exploratory Analysis of the LNPRICE variable, it could be seen that it shows a spatial autocorrelation and the hypothesis of its random spatial distribution is rejected. Confirmatory analysis is the most suitable tool to confirm if spatial autocorrelation exists, in order to establish the type and to model this spatial autocorrelation in such a way that it does not distort the results that we extract from the econometric model.

4.3.- Confirmatory Spatial Analysis

The Exploratory Analysis prepared in the above sub-section has allowed us to conclude that there is a high degree of spatial dependency in the prices of the rural land both inside and outside the RBU. This spatial dependency problem may be reason why autocorrelation exists in the remainders of the previous model that would spark two problems. First of all, the existence of this autocorrelation would invalidate the traditional estimation methods, such as the OLS estimation. On the other hand, if the model incorrectly incorporates spatial autocorrelation, the subsequent conclusions will also be erroneous. Confirmatory Spatial Analysis comprises a series of specifications, the specification contrasts, estimation models and specific validation procedures for the spatial models. Using these instruments, the existence of spatial autocorrelation will be confirmed, its typology determined, it will be included in the model in the most appropriate way to its typology and the final model validated.

Spatial autocorrelation may be of two types: substantive or residual. In case of a spatial lag of the explained variable or of any explanatory variable being omitted by error, the spatial dependency would be transferred to the error term, which would be spatially correlated. This type of autocorrelation is substantive and is solved by including the spatially lagged correlated variable in the model. When spatial dependency is not caused by the erroneous omission of the lag of a variable, the existing autocorrelation is residual. In this case, the solution consists of including spatial dependency in the error term.

The starting model is as follows:

$$lnPRICE_i = \beta_1 lnAREA_i + \beta_2 lnDATA_i + \beta_3 lnDISTN_i + \beta_4 lnGDP_i + \beta_5 lnMINLOT_i + \beta_6 RCNORC_i + \beta_7 RESERVE_i + \beta_8 LANDTYPE_i + \beta_9 CONS_i + \mu$$

Taking into account the presence of spatial autocorrelation that is revealed by the Exploratory Analysis and the absence of a theoretical model that includes the existence of spatial dependency in that starting model, it is proposed to follow the strategy of *Spatial Variable Expansion 2* (EEV2) created by Florax (1992) and Folmer and Florax (1992) to correct the existence of spatial dependency by means of including a set of spatial lags of the erroneously omitted variables. In order to develop this strategy, in first place, the base regression model has to be estimated using OLS. The Moran's I, LM-ERR (Lagrange Error Multiplier), LM-LAG (Lagrange Lag Multiplier) contrasts and the relevant robusts, LM-LE and LM-EL, are then obtained. These contrasts will have the following interpretation:

- If the contrasts lead us to accept the void spatial autocorrelation absence hypothesis, we will consider the starting model valid without presence of spatial autocorrelation asserting the randomness of the variable and the absence of spatial dependency clusters.
- If the Moran's I contrast and the LM-ERR and its LM-EL robust are significant, or at least the latter two have a lower probability than the LM-LAG and its robust, it would indicate the existence of residual spatial autocorrelation and the spatial error model would be estimated by maximum likelihood.
- If the Moran's I contrast and the LM-LAG and its LM-LE robust are significant, or at least the latter two have a lower probability than the LM-ERR and its robust, it would indicate the existence of substantive spatial autocorrelation and the spatial lag model would be estimated by maximum likelihood.

The results arising from the aforementioned EEV2 application are set out below. The following tables summarises the results obtained from the estimate using the Ordinary Least Squares. As can be noted, the R^2 and the adjusted R^2 are not very high but it must be taken into account that in the models where spatial dependency (a point that will be confirmed in the following step) is present, this adjustment measure is not reliable. With respect to the significance of the variables of the model, it should be stressed that the RCNORC and RESERVE variables are not significant. In the case of the RCNORC dummy variable, which determines the location inside or outside the rural centre, it could a priori be thought that the location of the land in a zone that is nearly suitable for development, such as Rural Centres or Population Centres, would be fundamental when setting the price. As can be seen, this variable is not significant, with the probability of accepting the void hypothesis (H_0 = non-significant variable) 0.3073. A possible explanation that we will consider further on (when we have the final model with the necessary corrections regarding the spatial autocorrelation presence) would lie in the fraudulent residential use of the rustic zones outside the Rural Centres, pursuant to

Articles 90 and 102 of the Master Plan for Using and Managing the Urdaibai Biosphere Reserve where the construction of a dwelling associated with an agricultural operation is allowed. This practice would justify the price not being significantly determined by the location inside or outside of the Population Centres.

With respect to the RESERVE variable, which includes whether the land in question belongs or not to the RBU, it makes the absence of explanatory capacity of this variable even clearer. In principle, the fact that a plot of land is located within the Urdaibai Biosphere Reserve leads one to suppose that there would be a greater readiness to pay a higher price for the same. The results show that this is not the case as the probability of accepting the void hypothesis of non-significance is 0.97. The lack of explanatory power of this variable may be due to the proximity to Greater Bilbao, the capital of the province, of those studied outside the RBU municipalities. The lack of availability of data from similar municipalities to Urdaibai, but which do not suffer so clearly from the pressure of a large city, has meant that the study of the prices of the plots of land taken from municipalities such as Arrieta, Fruniz, Meñaka or Morga may be biased due to their proximity to Bilbao and to rapid access roads. The strict regulation of agricultural land situated outside the Population Centres that has developed in the RBU means that the supply of that land is stricter, if possible, than outside the Reserve, which increases its prices. Due to this, and despite the exceptional natural and landscape values of the RBU, numerous people seeking land for residential purposes are put off from buying plots in its rural land due to the significant difficulties facing its subsequent property development and opt for alternative locations that also, in our case, are closer to the capital.

Another point to be mentioned of the model, before considering the spatial econometrics as such, is that it has had various variables that despite the fact that they would seem to have an explanatory power, this has not been the case given the results of the model. These variables are the distance to rapid access routes, the mortgage market interest rate, the return on alternative investments and the prices of agricultural land. They have not been introduced in the model due to the clear non significance with very high probabilities associated to accepting the void hypothesis.

The result of the variables are significant and with the excepted signs except LNGDP that appears with a negative sign, which would mean that increases in the income would be linked to drops in the prices of the land studied.

Diagnosing the regression made the high multicollinearity clear that exists between the variables, that the errors follow a Normal distribution (with an associated probability of 0.8936) and that there is heterocedasticity in the remainders (the probability of accepting the heterocedasticity hypothesis is 0.0734). It is interesting to point out that in the models where spatial autocorrelation exists, the heterocedasticity contrasts are not reliable because it may be generated by the spatial autocorrelation itself. We will return to this subject in the diagnosis of the final model where spatial autocorrelation will already have be taken into account and the existence of heterocedasticity will again be considered.

The following table sets out the results obtained from the spatial dependency diagnostics. The Moran's I makes clear the existence of spatial autocorrelation with a probability of 0.0000 that the void hypothesis of non spatial dependency is certain.

The following contrasts will reveal the type of autocorrelation. The two non robust statistics, LM-ERR and LM-LAG, are highly significant (probability = 0.0000). The robust statistic of the spatial error, LM-EL, is not significant while the spatial lag one, LM-LE, is highly so (probability = 0.0002). The LM-LE statistic reveals the presence of substantive spatial autocorrelation.

Therefore, on the one hand, the existence of spatial autocorrelation in the model and, on the other hand, the autocorrelation type has been determined. In conclusion, the results of the contrasts indicate the possibility that there are important interdependencies between the observations, which are omitted in the basic initial model. The results suggest that the model should be re-specified including a spatial lag cluster. The model will include the variable to explain the spatially lagged LNPRICE as an explanatory variable due to the spatial dependency cluster that was described in the previously developed Exploratory Analysis.

Table 2 ESTIMATION OF THE BASIC MODEL BY MCO

GOODNESS	VALUE
R^2	0.5346
R^2 ADJUSTED	0.4730
LIK	-116.410
AIC	252.819
SC	276.386
VARIABLE	**COEFFICIENT**
CONSTANT	-20743**
	(9375.29)
LNAREA	-0.3031*
	(0.1761)
LNDATA	2753.82**
	(1244.57)
LNDISTN	-0.3360***
	(0.1220)
LNGDP	-17.69**
	(8.7713)
LNMNLOT	-0.8451*
	(0.5019)
RCNORC	-1.0427
	(1.0137)
RESERVE	-0.0189
	(0.5038)
LANDTYPE	1.6012***
	(0.4669)
CONS	1.4066***
	(0.2851)
REGRESIÓN DIAGNOSTIC	**VALUE**
MULTICOLINEALITY COND.NUMBER	289753.4479
JARQUE-BERA TEST	0.2249 Prob = 0.8936
BREUSCH-PAGAN TEST	15.6990 Prob = 0.0734

- *, ** and *** denotes significance at 10%, 5% and 1%.

- The typical deviations appear in parenthesis.

Table 3 STATISTICS FOR SPATIAL DEPENDENCY DIAGNOSTICS

SPATIAL DEPENDENCE DIAGNOSTIC	VALUE	PROBABILITY
MORAN´S I	5.7033*	0.0000
LM-ERR	23.1157*	0.0000
LM-EL	0.1785	0.6725
LM-LAG	36.4545*	0.0000
LM-LE	13.5173*	0.0002

- * denotes significant to 1%

The following step consists in re-estimating the model including the spatial lag by means of the Maximum Likelihood (MV) model as the OLS estimation in this type of models generate biased and inconsistent estimators (see Moreno and Vayá, 2000). The Maximum Likelihood estimation obtains the estimators using the maximization of the logarithm of the likelihood functions associated to the specified spatial model. The spatial lag model to be estimated is:

$lnPRICE_i = \rho WlnPRICE_i + \beta_1 lnAREA_i + \beta_2 lnDATA_i + \beta_3 lnDISTN_i + \beta_4 lnGDP_i + \beta_5 lnMINLOT_i$
$+ \beta_6 RCNORC_i + \beta_7 RESERVE_i + \beta 8LANDTYPEi + \beta_9 CONS_i + \mu$
$\mu \sim N(0, \sigma^2 I)$

First of all, Table 4 reveals the R^2 goodness of fit that as has been explained is not useful in models with spatial autocorrelation. The correct method for selecting models in the sphere of spatial models involves comparing the logarithm of the likelihood functions (LIK), of the Akaike information criterion (AIC) and the Schwarz criterion (SC). A model will be better the higher the LIK value and the lower the ACI and SC values are. It can be seen from the results how the logarithm of the likelihood function increases and goes from -115.093 to -96.1848 and the AIC and the SC decreases and falls from 252.186 to 214.370 and from 278.110 to 240.293 respectively. In general terms, it is clear that the spatial lag model is more adjusted to the study date due to the confirmed existence of spatial autocorrelation.

The W_LNPRINCE lag variable has a positive coefficient (p = 0.6551) and is highly significant by supporting the selection of the spatial lag model.

The signs of the coefficients of the rest of the variables do not vary with respect to the basic lag model estimated by OLS but the value of the coefficient do. The absolute values decrease in the spatial lag model which implies that its explanatory power was inflated by neighbouring observations. The W_LNPRINCE lag variable has assumed this overvaluing. The significant variables in the basic model continue being so in the spatial lag one and the non-significant ones continue not to be so. There is only one exception: the LNMINLOT variable, which represents the minimum surface area required to be able to construct a dwelling in the zone where the plot of land is located. This variable was significant in the significant basic model, but it looses its explanatory power in the final model and stops being so. This variable could be said to be linked to the so-called RCNORC (which does not appear significant in any of the cases) and that, therefore, it confirms that the location inside or outside the population centres and the requirements with respect to the minimum surface area required to be able to construct are not determining when setting the price of the rural land studied.

The diagnosis of the spatial lag regression is set out below. The two heterocedasticity tests (Breusch-Pagan and spatial Breusch-Pagan) clarify that heterocedasticity does not exist. In the initial basic model, heterocedasticity appeared in the remainders, but, as was explained in that section, it cannot be taken greatly into account as that heterocedasticity was sometimes caused by the presence of spatial autocorrelation. In the case in question, that was true. Once the relevant spatial autocorrelation was introduced in the model, those tests accept the void homocesdasticity hypothesis of the remainders.

The likelihood test of the auto-regressive spatial coefficient (p) was also included and confirmed its significance.

If the specified spatial lag model is the appropriate one, the spatial dependency should be solved and the LM-SPATIAL ERROR test is calculated to check this point. As can be seen in Table 4, the void hypothesis of the spatial error model is rejected, which means that the spatial autocorrelation existing in the model data was well adjusted and modelled with the spatial lag.

In order to complete the analysis, the following statistics have to checked in the following order:

$$W \geq LR \geq LM$$

where

W = Wald Test which tallies with the square of the asymptotic t test of p (Z value of lag variable)

LR = Likelihood Ratio

LM = LM-LAG test value

It was seen that the following is fulfilled:

$$9.4912^2 \geq 40.4496 \geq 36.45$$

If that is not true, it would indicate a potential specification error of the model, that the error did not follow a normal distribution or that the matrix chosen to collect the spatial dependencies existing between the observations had been badly chosen.

Despite all the matrixes used indicating to us to follow the spatial lag model and that it has been confirmed as true, the spatial error model was calculated to confirm that it would be the appropriate one. In fact, the LIK, AIC, and SC tests indicate a poorer adjustment of the spatial error model compared to the spatial lag (LIK = -100.12 AIC = 220.240 and SC=243.807). On the other hand, the COMFAC tests make it clear that the spatial error model is inadequate with probabilities of 0.0366 and 0.0359. In addition, the LM-SPATIAL LAG test is highly significant which means that should the spatial error model be used, it would be spatial autocorrelation that would not have to be taken into account. Lastly, it should be pointed out that if this model would not fulfil the expression $W \geq LR \geq LM$ would take us to accept the applied spatial lag model.

Table 4 ESTIMATION OF THE SPATIAL LAG MODEL BY MV

GOODNESS	**VALUE**
R^2	0.6722
LIK	-96.1848
AIC	214.370
SC	240.293
VARIABLE	**COEFFICIENT**
W_LNPRICE	0.6551 *** (0.0690)
CONSTANT	-12239.8 ** (6303.91)
LNAREA	-0.3786*** (0.1182)
LNDATA	1624.65 ** (836.841)
LNDISTN	-0.1811 ** (0.08486)
LNGDP	-10.1855 * (5.8932)
LNMINLOT	-0.4835 (0.3374)
NRNONR	-0.6881 (0.6800)
RESERVE	-0.3071 (0.3385)
LANDTYPE	1.3135 *** (0.3176)
CONS	0.9010 *** (0.1933)
REGRESIÓN DIAGNOSTIC	**VALUE**
BREUSCH-PAGAN TEST	4.6685 Prob.= 0.8621
SPATIAL BREUSCH-PAGAN TEST	4.6958 Prob.= 0.8599
LIKELIHOOD RATIO TEST (LR)	40.4496 Prob.= 0.0000
LM SPATIAL ERROR TEST	1.2652 Prob.= 0.2606

- - *, ** and *** denotes significance at 10%, 5% and 1%.
- The typical deviations appear in parenthesis.

5.- Conclusions

The submitted article studies the demand for rural land in a natural enclave such as the RBU and in various nearby and similar municipalities that despite their rural nature and agricultural vocation, often reveal a real estate/residential interest.

When reviewing the literature that analyses the prices of land with econometric techniques, the lack of studies that analyse the problems in Protected Natural Spaces can be seen. This analysis is focused on an exceptional natural setting, but periurban in nature. Over recent year, the study of the formation of the price of land using spatially referenced databases has involved using Spatial Econometrics.

In the Spatial Exploratory Analysis of the Data, the existence of spatial autocorrelation between the prices of the land, both inside and outside the Urdaibai Biosphere Reserve, is revealed. This autocorrelation is high and positive, indicating that the price of the land that has been sold and bought inside and outside of the RBU are generally high, party due to the price of the neighbouring land being so. Local statistics (LISA) revealed that there is a small number of plots of land that form clusters or groups around them. In particular, the most significant are the plots of land with low prices and are associated to neighbouring land with low values.

In order to be able to confirm the autocorrelation that is revealed in the Exploratory Analysis, a basic model was initially put forward where the coefficients of the explanatory variables were calculated by means of OLS method. The following appear as variables with explanatory power: the surface area of the land, the year of the same, the distance to the centre of the municipality's urban centre, the GDP per capita of the province, the minimum surface area required to construct in the zone where the land is located, the typology of the plot of land and if it has been built on previously or not.

All these variables appear with the expected signs, except the GDP that has a negative coefficient contrary to the supposition of the theory of demand of a normal asset. The coefficient of the AREA variable appears as negative as the prices per m^2 of the larger plots of land are less as their theoretical use would be agriculture. On the other hand, the DATA variable reflects a positive coefficient as the value of the land has been adjusted over time by a percentages significantly higher than that of the inflation. With respect to the distance to the centre of the town (DIST), the sign is negative as expected due to the fact that the proximity to the service and leisure centre is an important feature when appraising a plot of land. The minimum surface area required to build on rural land outside the Population Centre (MINLOT) has a negative value as, in principle, if the surface area required to build in a zone is high, the price drops due to the land being used for agricultural (as is explained below, this variable stops being significant in the final model). The land type (LANDTYPE) determines its price to a great extent and reflects the fact that if the land belongs to the “Heredad” type is more valued that if it is not. Finally, the fact that the property has been built on before (CONS) also has explanatory power and, as is logical, it is more expensive than the land that has been built on subsequently. The non-significance of the variables of belonging to Rural Centre (RCNORC) and to the Urdaibai Biosphere Reserve (RESERVE) will be explained in detail in the final model.

It is highly important to stress that the basic model described above has serious problems of multicollinearity and heterocedasticity. When checking whether there are signs of spatial autocorrelation in it, it is clear that there are. It specifically indicates the existence of substantive spatial autocorrelation which led us to put forward a final model based on spatial lag, in particular by including in it the variable to explain spatial lagging as an explanatory variable. This specification implies that the prices of the land both in and outside of the RBU are partly determined by the prices of the plots of land considered as neighbouring.

As has been previously explained, the presence of spatial autocorrelation between the observations invalidates the OLS estimation and taking into account that the error term follows a Normal distribution, the recommend method is the Maximum Likelihood one which is applied to the spatial lag model. When calculating the model in this way, its was obvious that the adjustment of the model to the studied data had improved. The LIK, AIC and SC tests improved considerably.

The spatially lagged price appears significant with a probability equal to 0 and with a positive coefficient value supporting the positive spatial autocorrelation extracted in the previous Exploratory Analysis.

With respect to the significances and signs of the rest of variables, they follow the same outline as in the basic model, but the majority of its coefficient decrease in absolute values indicating that the explanatory power that was obtained in the basic model was overvalued by the neighbouring observations. This excessive explanatory power of the variables is trapped by the spatially lagged variable. The only variable that changes its significance is the SUPMIN variable that stops being so. This point may not be logically joined with the non-significance is either of the two models of the RCNORC variable (which indicates belonging to a rural centre). Given the obtained results, it may be thought that, although the location in or outside the rural centre was expected a priori to be determining, together the minimum surface area that is required to be able to build on each of them, it was shown not to be case. Taking into account that the population centre zones are the only ones where building is allowed even though with restrictions and that outside them, dwellings for farm workers associated to an agricultural operations only may be built is an interesting result. The abandonment of the agricultural activities initially linked to the building of new dwellings linked to this activity outside the population centres that has been observed on numerous occasions may be the cause of this result.

Rural land legislation is not effective at preserving agriculture land and may be causing to some extent distortions and fraudulent uses in the rural land market, in such a way that on agricultural land where it is only possible to build farm dwellings, first or second homes are being built justified by the agricultural operations that in the medium term disappear as such.

Given the need to preserve the natural environment is a clear fact and the planning policies are trying to achieve that objective, there should be the possibility to adjust that legislation to avoid fraudulent use of the rural land which should be use for agrarian purposes due to its natural, ecological and agronomic characteristics.

Finally, the non-significance of the RESERVE variable should be highlighted. The strict regulation of agrarian land located outside the Rural Centres in the RBU means that the supply of that land is even more strict that outside the Reserve, which increases its price. Nonetheless, this effect is counteracted by the closer proximity and better access to Greater Bilbao of the rural land in the municipalities adjoining the Reserve that have been studied. Despite it being an enclave with exceptional natural and landscape values, numerous people seeking land for residential purposes are put off from buying plots in its rural land in the RBU due to the significant difficulties facing its subsequent property development and opt for alternative locations that also, in our case, are closer to the capital.

6.- Bibliography

ALBERDI, J. C. (2002). "Hacia el abandono de la función agraria en el caserío vasco". *Estudios Geográficos* 247: 209-235.

ANSELIN, L. (1988). *"Spatial Econometrics: methods and models"* Ed. Kluwer Academic.

ANSELIN , L. (1992) "*SpaceStat Tutorial"* Ed. National Center for Geographic Information and Analysis, University of Santa Barbara.

ANSELIN, L. (1995) *"New directions in Spatial Econometrics"* Ed. Springer-Verlag

ANSELIN , L. (1999) *"Spatial Data Analysis with SpaceStat and ArcView"* Ed. Department of Agricultural and Consumer Economics.

ANSELIN , L. (2004) "*GeoDa 0.9.5-I Release Notes*" Ed. Center for Spatially Integrated Social Science.

ANSELIN , L. (2005) "Exploring Spatial Data with GeoDa: A Workbook" Ed. Center for Spatially Integrated Social Science.

ANTOLIN, J. E. (1999)."El proceso de urbanización del suelo rural: la vivienda familiar sobre el suelo no urbanizable en Bizkaia 1981-1992". Ed.UPV/EHU.

ASABERE, P. K. and HUFFMAN, F.E. (2001). "Building permit policy and land price distortions: empirical evidence" *Journal of Housing Economics* 10: 59-68.

ASHWORTH, J. and PARKER, S.C. (1997). "Modelling regional house prices in the UK" *Scottish Journal of Political Economy* 44(3):225-246.

BARRENETXEA APRAIZ, J. C. (2000). "*Situación actual y regulación jurídica de los núcleos rurales en la Reserva de la Biosfera de Urdaibai"*. Estudio encargado por el Patronato de la Reserva de Urdaibai. Gernika-Lumo.

BASU, S. and THIBODEAU T. (1998) "Analysis of spatial autocorrelation in house prices" *Journal of Real Estate Finance and Economics* 17(1) : 61-85

CAN, A. (1990) "The measurement of neighborhood dynamics in urban house prices" *Economic Geography* 66(3):254-272.

CAN, A. and MEGBOLUGBE I. (1997). "Spatial dependence and house price index construction" *Journal of Real Estate Finance and Economics* 14:203-222.

CHESIRE, P. and SHEPPARD, S. (2005) "*The Introduction of Price Signals into Land Use Planning Decision-making: a proposal"* Paper presented at the 45th

Congress of the European Regional Science Association, Amsterdam 23-27th 2005.

CHI-MAN HUI, E. and SZE-MUN HO, V. (2003). "Does the planning system affect housing prices? Theory and with evidence from Hong Kong". *Habitat International* 27(3):339-359.

DALE-JOHNSON, D. and BRZESKI, W.J. (2001) " Land value functions and land price indexes in Cracow, 1993-1999" *Journal of Housing Economics* 10:307-334.

GREEN H. A. J. (1982) *"La teoría del consumo"* Ed. Alianza.

GREENE, W. (1998) "*Análisis Econométrico*" Ed. Prentice Hall

MAYER, C. J. and SOMERVILLE, C. T. (2000). "Land use regulation and new construction". *Regional Science and Urban Economics* 30: 639-662.

MORENO, R. and VAYÁ, E. (2000) *"Técnicas econométricas para el tratamiento de datos espaciales : la Econometría Espacial"*. Ed. Universitat de Barcelona

MURUA, J. R., ASTORKIZA, I., FERRERO, A. and ETXANO, I. (2001). "Retos del Plan Rector de Uso y Gestión de la Reserva de la Biosfera de Urdaibai". *XXVII Reunión de Estudios Regionales*, Madrid 28-30 de noviembre de 2001.

PÁEZ, A., TAKASHI, U. and KAZUAKI, M. (2001). "Spatial association and heterogeneity issues in land price models" *Urban Studies* 38(9): 1493-1508.

PATTON, M. and McERLEAN, S. (2003) " Spatial effects within the agricultural land market in Northern Ireland". *Journal of Agricultural Economics* 54(1):35-54.

PLANTINGA, A. J., LUBOWSKI, R. and STAVINS, R. (2002). "The effects of potential land development on agricultural land". *Journal of Urban Economics* 52: 561-581.

RAMANATHAN, R. (1989) *"Introductory Econometrics with applications"* Ed. Harcourt Brace Jovanovich.

THOMAS, R.L. (1987). "*Applied demand analysis*" Ed. Longman