Estimation of Spatial Panel Data Models Using a Minimum Distance Estimator: Application

Théophile Azomahou

April 1999

Author affiliation

BETA-Theme, University Louis Pasteur
Economics Department
GSP, Engees-Cemagref
61, avenue de la Forêt Noire
F-67085 Strasbourg Cedex
France
phone: (+33) 3.88.24.82.33
fax: (+33) 3.88.24.82.84
e-mail: azomahou@cournot.u-strasbg.fr

Abstract

This paper is concerned with modelling and estimating panel data autoregressive spatial processes in the framework of minimum distance methods. A contiguity matrix based on distance between points relates observations spatially. The model is estimated in two stages. First, the cross-section parameters are consistently estimated by maximum likelihood, and a consistent asymptotic covariance matrix is computed for the second stage. Minimum distance estimators are derived under fixed slopes and all identical parameters restrictions. We used this specification to examine empirically spatial patterns of residential water demand for the French department of "Moselle", including electricity price effects.

Keywords: Minimum distance estimator, panel data, spatial dependence, water demand.

JEL Classification system: C13, C23, D12

1 Introduction

It is often relevant to consider the spatial distribution of phenomena such as diffusion patterns in counties or states of a country, or as a map of points of occurrences. This yields the spatial analysis of the so-called lattice data, i.e., observations for a fixed and given set of locations in space. Related applied econometric areas are various. LeSage and Dowd (1997) used this methodology to examine the spatial contiguity influences on state price level formation. A similar framework has been used by Case (1991) to analyse the spatial patterns in household demand for rice in some Indonesian districts. She also noticed that spatial modelling can be used in public economics for example to suggest the extent to which regions look to others in determining the appropriate composition of taxes or pricing, levels of expenditure and public good provision.

In a regression framework, spatial autocorrelation (more generally spatial dependence) is the situation where the dependent variable or/and the error term of a regression function, at each location, is correlated with observations on the dependent variable or/and values of the error term at other locations. As pointed out by Anselin (1988), ignoring this structure when it is actually existing results in misspecification and bias in estimation. Indeed, when studying the consistency of the ordinary least squares estimator (OLS), the presence of a spatial weight matrix in the dependent variable results in a quadratic form in the error terms. Therefore, the OLS estimator will be biased as well as inconsistent for the parameters irrespective of the properties of the disturbance.[1] In the case of spatial residual autocorrelation, OLS estimates will still be unbiased, but inefficient due to the non-diagonal structure of the disturbance covariance matrix (Kelejian and Prucha (1997)). The major estimation methods used to deal consistently with the multi-directional nature of spatial dependence are: maximum likelihood, method of moments and Bayesian analysis (Anselin (1988), Driscoll and Kraay (1998), Kelejian and Prucha (1998)). Recently, LeSage (1997) have shown that Gibbs sampling provides a consistent alternative framework.[2]

[1]The problem is analogue to the one induced by the simultaneity issues in least-squares. It arises from the introduction of dependence between neighbouring observations in y as additional regressor. Then, the multi-directional nature of the spatial dependence matters.

[2]Sometimes referred to as Markov Chain Monte Carlo sampling, Gibbs sampling allows to handle an analytically cumbersome Bayesian model. It provides estimates of spatial heteroskedasticity and estimates of parameters robust with respect to the presence of outliers.

While most studies focus on cross-sectional specifications, spatial panel data models have not received much attention. As outlined by Case (1991), fixed effect models can be used to control for spatial components using panel data. But in some cases, when there is no intra-regional variation in variables of interest, a spatial modelling approach may be more appropriate. This is the case for example when the variation in the variable depends upon distance between points. Then, there is a perfect correlation between the variables of interest and the fixed effects. The same paper discusses the gains in information and efficiency which are achieved by spatial random effects modelling, and shows that when specific effects are uncorrelated with right hand side variables, there are clear benefits to spatial specification. More generally, it can be argued that the equicorrelated structure of spatial dependence implied by the error components model does not allow for distance decay effects.

Here, I consider modelling and estimating panel data autoregressive spatial processes in the general framework of minimum distance estimators. The study starts specifying a mixed regressive spatial autoregressive model. This specification defines a class of random fields, i.e., models derived from processes indexed by space, time and cross-sectional dimensions. I work with a row-standardized contiguity matrix, i.e., the spatial weight matrix is normalized so that the rows sum to unity. This standardization produces a spatial lagged variable that represents a vector of average values from neighbouring observations. The specification is assumed to be the true data generating process which relates observations with reference to points in space and time. Then, we take advantage of various econometric studies to estimate the model in two stages.

First, assuming the errors to be normally distributed, the cross-section parameters can be consistently estimated by maximum likelihood. Under suitable regularity conditions, this stage provides both unrestricted consistent parameters estimates, including the spatial coefficient, and scores which are used to compute the consistent asymptotic covariance needed for the second stage. Then, two cases are considered: the fixed slopes case and the all identical parameters one. The minimum distance estimator is computed for each case by stacking the estimates from the first stage in a block vector on which moment conditions are imposed. In the second stage, we minimize a quadratic distance from zero in the norm given by the inverse of the block asymptotic covariance matrix. The resulting minimum distance estimator is consistent and asymptotically efficient as well. We used this specification for empirical purposes.

The empirical analysis consist in examining the spatial variations of residential water demand for the French department of "Moselle", including electricity price effects. As indicated by Hansen (1996), when estimating the determinant factors of residential water demand, we may expect to observe the indirect effects of energy variables, according to water consumption between different water-using tasks. Indeed, water is consumed by households jointly with different tasks which involves use of water and in most cases sizable amounts of energy and other goods (appliances, etc.). Table 9 in Appendix 5.1 reports the daily distribution (on average) of French residential water consumption between household tasks. About 40% of this distribution is concerned with water heating (mainly by electricity). We combine this consideration with spatial aspects. In this context, the specification described may be viewed as a model of endogenously changing tastes, which allows to check for social interdependence by testing the extent to which households look to a reference group when making water consumption decisions. It may also be thought of as indicating the magnitude and the direction of interactions between consumers with respect to the availability of water resources.

Section 2 presents the data. It describes the sampling and basic descriptive statistics. Spatial correlograms are computed to check for spatial patterns in consumption observations. I also use nonparametric density estimation to identify relevant features which occur with the water average price distribution during the data collecting period. Section 3 presents the model and estimation results. It extends the basic elements of spatial modelling to panel data specification using the minimum distance approach. Section 4 concludes the study.

2 Data

The department of "Moselle" is made of about 730 communities (municipalities) out of which 115 neighbouring communities have been selected for the empirical study of the households demand for drinking water.[3] Households living in these communities are supplied with drinking water by a private operator. The data considered here represent the first lattice collected from the French network of drinking water distribution. The data are collected with a biannual frequency (from 1988.1 to 1993.2). We have then a balanced panel of 1380 spatial observations. Some avail-

[3]The department of "Moselle" is located in the north-east of France. The communities selected for the study are those for which we succed in obtaining reliable informations.

able variables do not require previous important changes before being used. Others (municipal characteristics) have been constituted from informations available in the last municipal inventory.[4] This section describes the sampling and relevant features in variables. I also describe the possible measurement errors in data. Appendix 5.2 is devoted to the data sources and the definition of variables.

2.1 Sampling and descriptive statistics

The first step of this study has been the practical work which led to the data collected. Since this collection has never been undertaken before, two important issues arose from a closer look of consumption values. The first one was the identification of households' consumption. The network manager (private operator) provides water service to the so-called "subscribers" which terms as well the citizens, that is to say the users living in individual house or in collective blocks of flats (for instance council flats), as industrials and businesses. The households' demand gathers together individual user consumption and collective user consumption. Most of the households living in collective lodgings do not yet have meters that can give them an accurate indication about the amount of their consumption. It is also noticed for these consumers, that the water price is included in the rent charges. As a result, we can suppose that the households concerned are not aware of the necessity to control their budget with respect to water expenses.

Moreover, the example of collective blocks of flats which shelter small businesses is mentioned. In the case when a household living in a collective lodging gets a business linked to his subscriber regime, we cannot separate the household consumption of water from the business one. A similar issue occurs for some households living in individual houses. Indeed, for those among them who possess farms or small size exploitations linked to their "subscriber regime", the identification of purely domestic volumes is difficult. For all these reasons, and in order to reduce the evaluation errors and to be sure that the target sector corresponds to the residential one, I have selected the subscribers connected to the drinking water supply network with a main water capacity of 15mm in diameter, when this information was available. Despite of this choice, we cannot exclude that some marginal consumption values (coming from small businesses or other consumptions different from the domestic

[4]All informations related to the communities characteristics come from the last municipal inventory. The municipal inventory is a document which provides the characteristics of French communities. The study is conducted by the "National Institute of Statistics and Economic Studies". The last recording dates from 1988.

Table 1: Descriptive statistics of consumption and average price

	Consumption per house				Average price in FF			
Period	mean	std	min	max	mean	std	min	max
1988.1	69.68	27.75	1.11	153.15	6.28	2.11	3.24	11.29
1988.2	70.13	23.56	1.04	148.74	6.37	2.14	3.27	11.38
1989.1	72.28	28.37	1.00	186.78	6.69	2.31	3.09	11.52
1989.2	74.55	27.11	0.88	175.28	6.79	2.35	3.09	11.64
1990.1	73.47	27.43	0.96	162.52	7.05	2.44	3.40	12.40
1990.2	72.67	26.33	0.86	163.37	7.22	2.53	3.41	12.54
1991.1	75.56	29.17	0.90	179.48	7.70	2.65	3.47	13.08
1991.2	75.04	28.90	0.86	187.81	7.95	2.91	3.56	16.19
1992.1	71.94	27.07	0.73	155.81	8.66	3.40	3.63	17.59
1992.2	72.75	27.68	0.87	170.37	9.01	3.51	3.67	18.10
1993.1	72.14	26.51	0.81	157.33	9.97	3.97	4.09	19.46
1993.2	71.24	29.26	0.83	176.19	10.58	3.50	4.77	19.50

one) were found in collected data.

The second problem concerns the reconstruction of some consumption values: either because they disappeared during floods (it is the case of 1990's data), or they existed under high level of aggregation. This concerns only some (very few) unionized municipalities. The non-unionized municipalities display a semester volume. Unions result from the gathering together of municipalities. The union data are used for the estimation of the volume consumed when municipalities data were missing. Thus, the household semester volumes were not available to be used directly. The data used to reconstruct consumption values, as far as municipalities are concerned, come from a document termed "water products". The volumes looked for are semester values. When semester data are missing, I face two cases: either only some municipalities composing the union are considered and in this case I suppose that the consumption in the other municipalities varied in the same proportion, or the details of the volume consumed are not available and in this case, the average weight of each municipality in the union is computed. As a result, the data present two characteristics which make their biannual use delicate.

On the one hand, the water reading frequency ran from at least a quaterly period to a yearly one. In the meantime, the pricing remains biannual. The accurate biannual readings are available for 1988.1 for all the municipalities, as well as the readings

Table 2: Descriptive statistics of meteorological variables

	Rainfall in m				Mean temperature in C^0			
Period	mean	std	min	max	mean	std	min	max
1988.1	8.76	0.65	7.38	11.10	8.91	0.33	8.00	9.58
1988.2	7.29	0.70	5.94	8.87	11.47	0.29	10.68	12.15
1989.1	6.17	0.56	5.12	7.69	8.78	0.32	8.21	9.50
1989.2	6.55	0.50	5.69	8.92	11.86	0.36	10.83	12.55
1990.1	6.94	0.66	5.84	8.24	9.33	0.26	8.66	10.00
1990.2	6.93	0.77	5.50	9.41	11.48	0.32	10.56	12.20
1991.1	4.65	0.40	3.72	6.80	7.02	0.28	6.25	7.65
1991.2	5.95	0.89	4.76	7.51	11.99	0.27	11.15	12.58
1992.1	5.46	0.67	4.14	7.51	8.70	0.28	8.10	9.28
1992.2	7.97	1.20	5.26	10.25	11.92	0.18	11.30	12.46
1993.1	4.50	0.81	2.77	5.77	8.81	0.22	8.31	9.46
1993.2	10.14	0.72	8.78	12.35	10.71	0.26	9.90	11.28

Table 3: Descriptive statistics of disposable income (in thousands of FF)

Period	mean	std	min	max
1988	57.51	8.28	33.38	75.24
1989	59.07	8.65	37.47	79.23
1990	62.06	9.31	31.82	85.16
1991	63.82	10.31	33.66	92.14
1992	65.49	11.33	34.18	104.16
1993	66.97	11.76	34.33	97.68

Table 4: Descriptive statistics of characteristics in 1990

Variable	mean	std	min	max
Proportion of person <19 years	0.28	0.04	0.13	0.31
Density of population	1.10	2.60	0.0038	14.61
Proportion of Workers	29.96	4.39	11.92	37.82
Proportion of Unemployed	9.78	4.03	2.70	23.62
Index of equipment	61.87	6.86	30.24	76.84

of 1993.1. From 1990.1 to 1992.2 some municipalities adopted a yearly reading. In this case and to reduce the cost induced by meters readings, the volumes for one semester is estimated from the consumptions of a precedent year, whose duration between two readings does not always equal 52 weeks. Moreover, the calendar year is no longer taken into account, but a year running from June to June. On the other hand, we face a difference in the frequency of data collection. Indeed, the consumption reading frequency may vary from one year to another because of climatic hazards or other unforseeable parameters. To correct these bias, the consumption values presented in this study are corrected to lead to a 52 weeks frequency. These two characteristics, estimated values and difference in the reading frequency, are possible sources of errors.[5]

The explained variable is the aggregate water consumption per community expressed in cubic meter per house. Urban communities are larger than rural ones. So as to consider homogeneous observations and in order to reduce the community size effect, each consumption value has been divided by the total number of households per community in 1990, the year of the last inventory available. It also constitutes the last period when the population general census was conducted by the offices of the National Institute of Statistics and Economic Studies "(INSEE)". Descriptive statistics related to the variables are shown in tables 1, 2, 3 and 4.

National statistics indicate an average water consumption tendency around 120 m3 per house and per year. These figures vary from one house to another. Old houses are light on water consumption whereas high standing dwellings with gardens can consume around 180 m3. When we compare these indicators with those computed from the sample, we notice that the average consumptions recorded are of the same magnitude. Minimum values can be considered as the consumption of rural communities. These tendencies are also indicative of the standard of living of the population considered. As a whole, there is no outliers in consumption values. Note however some high values for 1989.2, 1991.1 and 1991.2 where we observe 74.55, 75.56 and 75.04 m3 respectively. This may result from extra consumptions in addition to purely domestic ones. It may be the case for households having small businesses or farms as described above. These statistics support, on average, the relative statility of our data.

[5]We compute nonparametric density estimation to look closer for the distribution of consumption values. The results show a main unimodal distribution around 60 and 80 m3 for each period. These estimates support our recording target sector, i.e. the consumption of residential subscribers. Density estimates are not presented here but they are available upon request.

Disposable income statistics are characterised by very low values. Consider for example the year 1990 where the minimum values are the lowest, that is, 31,820 FF per household paying tax. One obtains a monthly disposable income figure of 2,651.66 FF. Supposing that this household paying tax is made up of a single, the latter earns around the "minimum insertion income" in France. This shows the difficulty usually encountered in recording income data. Other reasons explain these low values. Indeed, various studies conducted by the "National Institut of Statistics" show that, in the department of "Moselle", taxable incomes under-estimate by 30% on average the actual household incomes.[6] This under-estimation is extremely high for the self-employed (43%), and even more for self-employed farmers (57%). Moreover, even if we know that the consequences of the economic crisis on the evolution of global wages has been compensated by a strong increase of social benefits and a slight increase in taxes, the "Moselle" departement is below the national indicators.

Average price values clearly indicate relevant patterns. The average price increases continuously on the twelve biannual periods. This increase shows three figures. From 1988.1 to 1989.2 the average price is below 7 FF; from 1990.1 to 1991.2 it is below 8 FF and from 1992.1, the tendency is higher than the previous ones. This last tendency indicates an important modification in the water price structure. As a whole, the price variable suggests a clustering pattern. It also presents an increasing dispersion within clusters with stable minimum values (around 3,5 FF). All these figures are examined more carefully in the next paragraph. Before, note that the meteorological values presented here are not dummy variables as in many studies, but the thrue values recorded by the Regional Center of Meteorological Studies.[7]

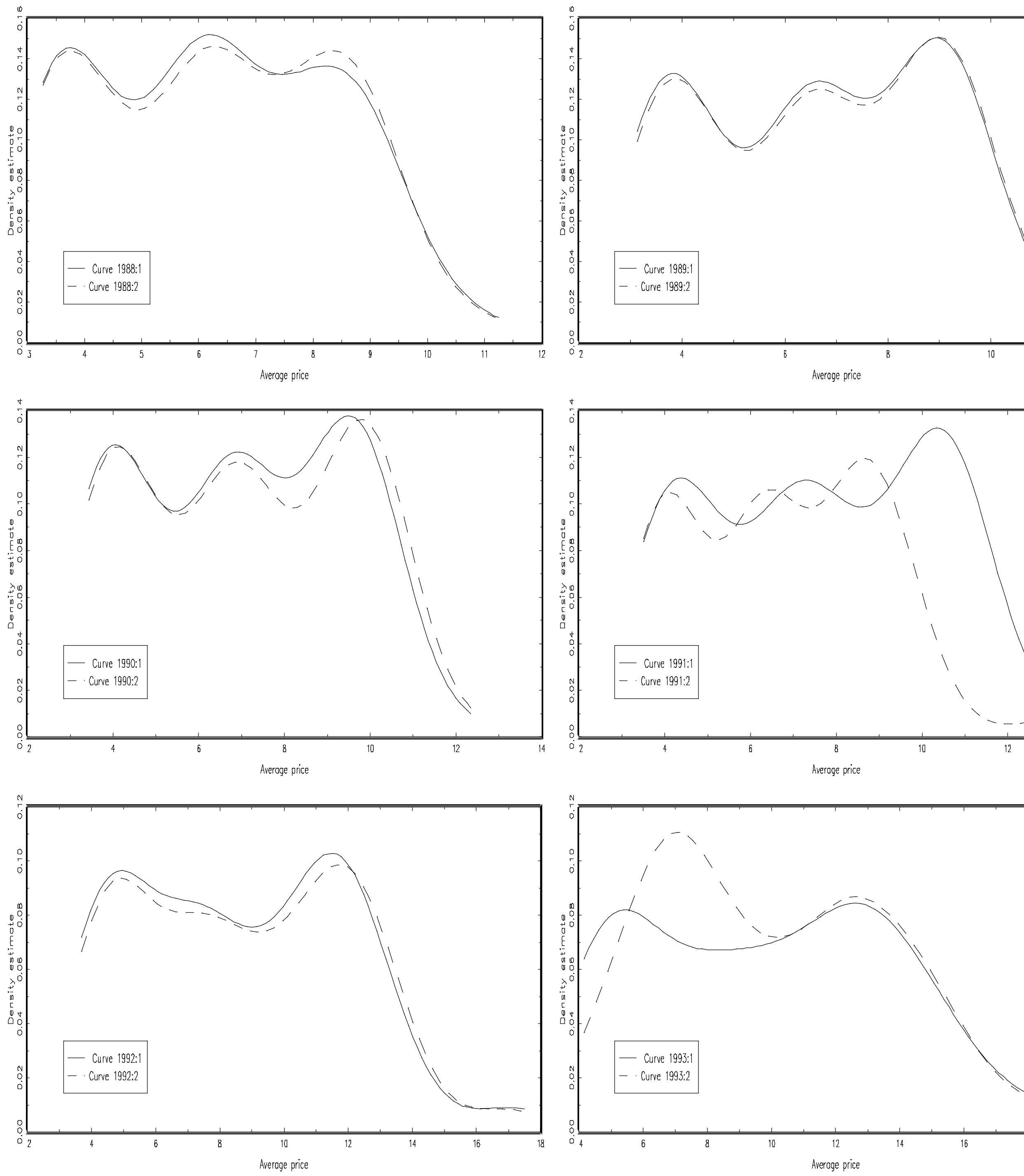
2.2 Distribution of average price

For various reasons described below, it seems relevant to study the distribution of the average price during the data collection period. Indeed, the organization and the management of water distribution in France pertain to public service liability. The price results from a negociation between local authorities and the water distributor which may be the local collectivity itself or a private company. Communities and households concerned by this study are supplied with drinking water by a private firm.

[6]Tableaux de l'Economie Lorraine 1997/1998 (Tables of Lorraine Economics), (INSEE (1998)).

[7]As we may expect, the first semester values are less than those of the second semester.

Figure 1: Distribution of average price, kernel density estimates (Three modes from 1988.1 to 1991.2 corresponding to three within sector pricing. In 1991.1 and 1992.2, the central mode starts disappearing as a result of the "M-49 directive". From 1992.1 on, only two modes remain.)



According to the water supplier, the communities are organized in two sectors, but there is some doubt about the exact number of sectors. We denote each sector by a dummy variable (dummy 1 for sector 1 and zero for sector 2). Out of 115 communities, 65.2% belong to sector 1. The sectors correspond to two distinct areas of water management. This spatial arrangement is mainly due to the network management issues (water transportation and various treatments to make water drinkable) and is closely linked to the various elements of water prices.[8] The marginal price of water is the same within a given sector but varies between sectors. Thus, we know that there is no intra-regional variation in the marginal price. But the average price varies from one community to another when the fixed charges of water are included, see appendix 5.2 for the computation of average price. Moreover, the laws on water of november 1992, by the so-called "M-49 directive", have strongly modified the working orders of water agencies.[9] This modification has been translated into a high increase in water prices. The aim is to let customers pay for the effective price of water, and no more for the water service.

To check for the persistency of sector design effects in the distribution of the average price (having incorporating the fixed charges of water), we use nonparametric estimation for data analysis and identification purposes (Silverman (1986) and Wand and Jones (1995)). Figure 1 shows the kernel density estimate of the average price for each time period.[10] Two main results follow. We notice that, up to 1991.2, the distribution displays three modes. From 1992.1 on, the central mode starts disappearing and by 1993.2 there are only two modes left. This distribution can be mainly explained by the modifications that occurred in water pricing in 1992. These modifications are due to the "M-49 directive" which resulted in a change in water pricing. Not only the price increased continuously as indicated by descriptive statistics, but now, two sectors appear clearly from 1992. The distribution reveals that there may be three sectors up to 1992.1. Thus, sector design effects remain in the average price. We may expect a within sector behavior regarding the water consumption and then a spatial effect.

[8]To make ideas clear, we computed the correlation coefficient between the average price variable and the sector dummy for each time period: (-0.33, -0.32, -0.38, -0.38, -0.36, -0.39, -0.35, -0.34, -0.35, -0.38, -0.35, -0.41). There is evidence of correlation.

[9]Set up on November 10th, 1992 (its implementation date) the "M-49 directive" imposes to water services (supply and cleaning up) the rule of budget balance. They are forced not to make their general budget support the water spendings (building up and maintenance of network, equipments, cleaning up...).

[10]Here, we use the Epanechnikov kernel and the cross-validation method for the choice of the bandwidth.

2.3 Test for specific spatial autocorrelation and spatial correlograms

I introduce here various analytic methods which are of value in assessing the spatial scale of a process. The variables of interest are: consumption, average price and disposable income. I use the G-statistic which provides a measure of overall spatial association, and observation-specific spatial association as well. These statistics are computed by defining a set of neighbour communities. For each location, these communities are considered as those which fall within a distance band. Two tests are computed.

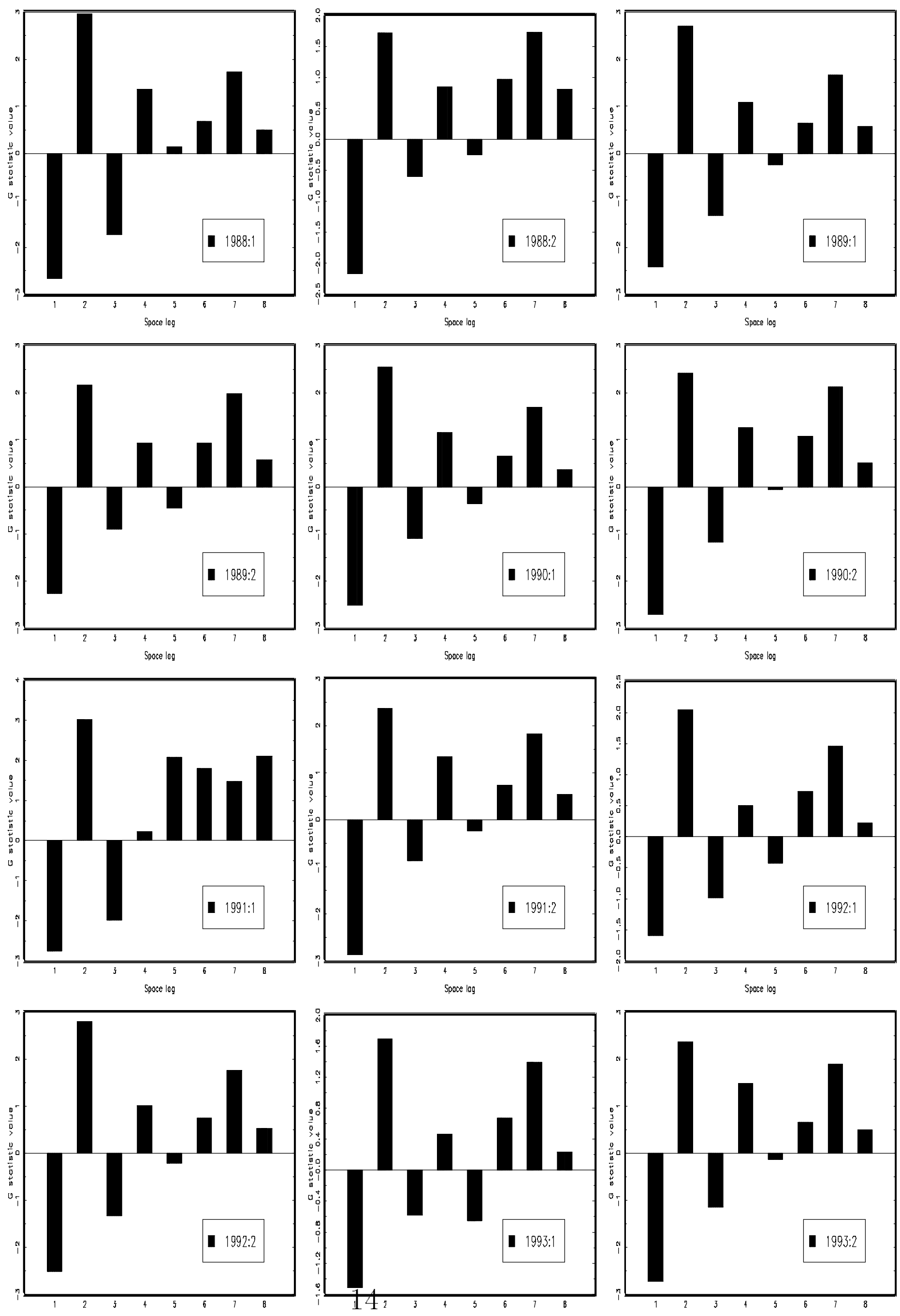
Table 5: G-test for specific spatial autocorrelation

	Consumption		Average price		Disposable income	
Period	G-stat	prob (%)	G-stat	prob (%)	G-stat	prob (%)
1988.1	0.329	0.7	0.389	00	0.358	48
1988.2	0.337	3.0	0.387	00	0.358	48
1989.1	0.332	1.6	0.395	00	0.361	15
1989.2	0.335	2.3	0.395	00	0.361	15
1990.1	0.332	1.2	0.395	00	0.359	30
1990.2	0.331	0.6	0.397	00	0.359	30
1991.1	0.329	0.5	0.399	00	0.362	9.7
1991.2	0.328	0.4	0.399	00	0.362	9.7
1992.1	0.341	11	0.405	00	0.363	7.5
1992.2	0.332	1.1	0.405	00	0.363	7.5
1993.1	0.342	13	0.405	00	0.365	2.6
1993.2	0.328	0.6	0.393	00	0.365	2.6

First, we test for a specific spatial association, i.e., the extent to which a location is surrounded by a cluster of high or low values for the variables of interest for each period.[11] The statistics are reported in tables 5. We observe a significant value for the consumption (except for 1992.1 and for 1993.1) which is indicative of a spatial clustering of low values. Average price G-statistic are all highly significant. Then, a spatial dependence for high values occurs. Except for 1993, spatial autocorrelation for income values is rejected. This test has a "static aspect" and do not provide informations on the spatial dynamics of the process. This issue is handled on the sequel using spatial correlograms.

[11]Note that measures of global spatial association such as Moran's I and Geary's c fail to detect such a pattern.

Figure 2: Estimation of spatial correlograms for consumption (Up to eight spatial lags on the X-axis and the t-value of the G-statistic on the Y-axis. The first two lags of each correlogram are highly significant indicating spatial dependence which decreases with lags, except for the seventh lag. The eighth lag for 1991.1 presents also significant spatial dependence.)



Although the interaction between spatial units may be strong between immediate neighbours, the strength of interaction will often vary in a complex way with distance. We test for the difference of spatial autocorrelation for the consumption variable over different weight matrices by using spatial correlograms.[12] For further technical details and discussions on spatial correlograms see Cliff and Ord (1981), and Cressie (1991).

The results of the estimated spatial correlograms for each time period are reported in figure 2. Spatial lags are reported on the X-axis (up to eight lags), and the t-statistics associated to G-values are indicated on the Y-axis.[13] A significant and strong indication of spatial clustering for the first and second orders of contiguity is evident (except for 1993:1). We notice a decreasing spatial autocorrelation with increasing orders of contiguity, which is typical of many spatial autoregressive processes. The significant and negative spatial autocorrelations at lag 1 contrast with the significant and positive spatial autocorrelations at lag 2. Then, at lag 1, low values of water consumption are likely to be spatially correlated, and at lag 2, it may be the case for high values. This result clearly indicates potential spatial dependence in consumption observations. Thus, it seems relevant to include the spatial dimension in the model specification.

3 Empirical specifications

In this section, I consider modelling and estimating panel data autoregressive spatial processes within the framework of minimum distance methods. I use the class of mixed regressive spatial autoregressive models as described in Anselin and Bera (1998).

3.1 Model

Consider a linear panel data regression model where the observations are stacked by time period $y_t = (y_1, \cdots, y_i, \cdots, y_N)'$ and $X_t = (X_1, \cdots, X_i, \cdots, X_N)'$ for $t =$

[12] Higher order contiguity is used to compute spatial correlograms. The contiguity matrices are obtained by taking powers of the unstandardized form of the first order contiguity matrix and by correcting for circularity. The spatial lag length is eight. It corresponds to the point where higher order contiguity result in unconnected spatial units, i.e., spatial units for which the corresponding row in the contiguity matrix consists only in zeros.

[13] To ease presentation, other statistics related to spatial correlograms (expectation, standard deviation, and significance level) are not reported here.

$1, \cdots, T$ and $i = 1, \cdots, N$. Such data organization is due to the introduction in the sequel of the matrix W which is the same over time. Moreover, it allows to take observations by cross-section. The dimension of y_t is $(N \times 1)$ and X_t is $N \times (K-1)$. We assume that each cross-section follows a spatial autoregressive process. Then, the model has the following structure:

$$y_t = [Wy_t, X_t]\,\theta_t + \varepsilon_t \quad \theta_t = (\rho_t, \beta_t)', \quad |\rho_t| < 1 \tag{1}$$

where y_t is the vector of dependent cross-sectional observations for an area (state, district, municipality etc.), W is a known $(N \times N)$ spatial weights matrix, usually containing first-order contiguity relations or functions of distance between spatial units. This is the matrix computed in Appendix 5.3. We work with a row-standardized version of W, that is W is normalized so that its rows sum to unity. This standardization produces a spatial lagged variable Wy_t (also termed regionalized variable) that represents an average of values from neighbouring y_t for each time period. X_t is the matrix of explanatory variables, θ_t is a $(K \times 1)$ vector of unknown parameters to be estimated. It contains the spatial coefficient (scalar) ρ_t and the vector β_t (of dimension $K-1$) of the other explanatory variables. $\varepsilon_t = (\varepsilon_1, \cdots, \varepsilon_i, \cdots, \varepsilon_N)'$ is the $(N \times 1)$ vector of disturbances.

Let consider that the G-variates y_t are generated by $y_t = f(y_t, X_t; W, \theta_0) + \varepsilon_t$ where $\theta_0 \in \Theta \subset \mathbb{R}^K$, $y_t \in \mathbb{R}^G$, $X_t \in \mathbb{R}^P$, $\varepsilon_t \in \mathbb{R}^G$. We assume that the conditional distribution of ε_t given X_t is equal to the product of the conditional distributions for $t \neq s$. This distribution is assumed to be gaussian with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_s') = \sigma_t^2 Id$, for $t \neq s$.

Relation (1) is analogous to the multivariate lagged dependent variable model for time series regressions, with a spatial parameter ρ_t indicating the extent to which variations in y_t are explained by the average of its neighbouring observations values. The hypothesis of normal errors allows to estimate the parameters of each cross-section separately by maximum likelihood. Estimation by asymptotic least squares (minimum distance) is conducted in two stages.

First, let $\hat{\theta}_t$ denote the unrestricted maximum likelihood estimates for parameters ρ_t, β_t and σ_t^2 for each cross-section. That is:

$$\hat{\theta}_t = \arg\max_{\theta_t \in \Theta} \sum_{\substack{i=1,\cdots,N \\ t \neq s}} \psi_{it}\left(y_t, X_t; W, \theta_t\right) \tag{2}$$

where $\psi_{nt}\left(y_t, X_t; W, \theta\right)$ denotes the log likelihood function computed as:

$$\psi_{nt}(y_t, X_t; W, \rho_t, \beta_t, \sigma_t^2) = -\frac{N}{2}\ln(2\pi) - \frac{N}{2}\ln\sigma_t^2 + \ln|A| - \frac{1}{2\sigma_t^2}e_t'e_t \tag{3}$$

with $e_t = Ay_t - X_t\beta_t$, $A = I - \rho_t W$ and $|\cdot|$ denotes the determinant. Such a likelihood is non-linear in parameters and is usually handled numerically. For the details see Anselin (1988).

In a second stage, we use the unrestricted maximum likelihood estimator of the first stage to obtain the restricted asymptotic least squares estimates by imposing several restrictions of the form $g(\hat{b}(\theta), a) = 0$. The ALS estimator is obtained by choosing $\hat{a}_n$ to minimize a quadratic form for the norm given by the inverse of the asymptotic covariance matrix of $g(\hat{b}(\theta), a_0)$. Then, the ALS estimator $\hat{a}(S_n)$ is given by the minimization program:

$$\hat{a} = \arg\min_{a\in\mathcal{A}}\Big(g(\hat{b}(\theta), a_0)\Big)' S_n\left(g(\hat{b}(\theta), a_0\right) \tag{4}$$

where $S_n \overset{\text{a.s.}}{\to} S$ a positive definite symmetric matrix, $\mathcal{A} = a(\mathcal{P}) \subset \mathbb{R}^K$ and $a_0 = a(P_0), \forall P \in \mathcal{P}$. The optimal choice for S is known to be the inverse of the covariance matrix of $g(\hat{b}(\theta), a_0)$ (Gouriéroux, Monfort, and Trognon (1985)) and (Kodde, Palm, and Pfann (1991)). Under suitable regularity conditions, the estimator $\hat{a}(S_n)$ exists and is consistent. Let Ω denote the approximation of S:

$$\Omega = \frac{1}{N}\left[J^{-1}IJ^{-1}\right] \tag{5}$$

where $J = \text{diag}\{J_1, \cdots, J_T\}$ is a block diagonal matrix with elements:

$$J_t = E\left(-\frac{\partial^2\psi(y_t, X_t; W, \theta_0)}{\partial\theta\partial\theta'}\right) \tag{6}$$

and elements of I are given by:

$$I_t = E\left(\frac{\partial\psi}{\partial\theta}(y_t, X_t; W, \theta_0)\frac{\partial\psi}{\partial\theta}(y_t, X_t; W, \theta_0)'\right) \tag{7}$$

We obtain a consistent estimator $\hat{\Omega}$ of Ω by replacing theoretical expectations by sample means as follows.

Let $\psi_i(y, X_t; W, \rho, \beta, \sigma^2)$ denote the log-likelihood for one observation:

$$\begin{aligned}\psi_i(y, X; W, \rho, \beta, \sigma^2) = &-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 + \frac{1}{N}\ln|A| \\ &- \frac{1}{2\sigma^2}\left[\sum_{j\in J}\left(\mathbf{1}_{[i=j]} - \rho\omega_{ij}\right)y_j - \sum_k X_{ik}\beta_k\right]^2\end{aligned} \tag{8}$$

where $j = 1, \cdots, J$ is the set of communities contiguous to a community i and $\mathbf{1}_{[i=j]}$ denotes an indicator function. Taking partial derivatives of (8) with respect to the parameters yields:

$$\frac{\partial \psi_i(\cdot)}{\partial \beta_k} = \frac{1}{\sigma^2}\left[\sum_{j\in J}\left(\mathbf{1}_{[i=j]} - \rho\omega_{ij}\right) y_j - \sum_h X_{ih}\beta_h\right] X_{ik} \tag{9}$$

$$\frac{\partial \psi_i(\cdot)}{\partial \rho} = \frac{1}{\sigma^2}\left[\sum_{j\in J}\left(\mathbf{1}_{[i=j]} - \rho\omega_{ij}\right) y_j - \sum_h X_{ih}\beta_h\right] \sum_{j\in J}\omega_{ij} y_j + N^{-1}\xi \tag{10}$$

$$\frac{\partial \psi_i(\cdot)}{\partial \sigma^2} = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}\left[\sum_{j\in J}\left(\mathbf{1}_{[i=j]} - \rho\omega_{ij}\right) y_j - \sum_k X_{ik}\beta_k\right]^2 \tag{11}$$

with $\xi = \frac{\partial \ln|A|}{\partial \rho} = -\left[\mathrm{tr}\left(A^{-1}W\right)\right]$.

Let $\hat{\mathrm{v}}_t = \left[\frac{\partial \psi_i(\cdot)}{\partial \beta_k} \mid \frac{\partial \psi_i(\cdot)}{\partial \rho} \mid \frac{\partial \psi_i(\cdot)}{\partial \sigma^2}\right]_{\theta=\hat{\theta}}$ be a block element of $\hat{I}$ of dimension $N \times (K+2)$ obtained by stacking the vector of derivatives evaluated at parameters estimates. The empirical variances matrix $\hat{I}$ of individual scores is given by the cross product of $\hat{\mathrm{v}}_{t,s}$ for $t \neq s$. The estimate $\hat{\Omega}$ of Ω is computed as $\hat{J}^{-1}\hat{I}\hat{J}^{-1}$.

From this general specification, we obtain various estimates by imposing two restrictions. The first restriction is that of fixed slopes expressed as:

$$g(\hat{b}(\theta), a) = \begin{pmatrix} \hat{\theta}_x^1 - \theta_x \\ \hat{\theta}_x^2 - \theta_x \\ \vdots \\ \hat{\theta}_x^T - \theta_x \end{pmatrix} = O \tag{12}$$

with $\hat{\theta}_t = (\hat{\theta}_t^0\ \hat{\theta}_t^x)'$, $t = 1, \cdots T$, where θ_t^0 and $\hat{\theta}_t^x$ denote respectively the parameters vector of varying intercept and the parameters vector of fixed slopes for the period t, and $a = \theta_x$. The second restriction is that of all identical parameters:

$$g(\hat{b}(\theta), a) = \begin{pmatrix} \hat{\theta}_1 - \theta \\ \hat{\theta}_2 - \theta \\ \vdots \\ \hat{\theta}_T - \theta \end{pmatrix} = O \tag{13}$$

with $\hat{b} = (\hat{\theta}_1, \cdots, \hat{\theta}_T)'$ and $a = \theta$. For each case, we obtain the ALS estimates by generalized least squares.

3.2 Estimation results

We use the model specified above to carry out empirical estimation on data described in the previous section.[14] Tables 6 and 7 present the unrestricted maximum likelihood estimates for the mixed regressive spatial autoregressive model for the twelve time periods. A Lagrange multiplier test shows rejection of the alternative spatial error specification for most cases except for 1988.2, 1991.1, 1992.1 and 1993.1.

For these cases, spatial dependence remains in the residuals and our specification is clearly rejected. Thus, a mixed autoregressive spatial moving average model, i.e., a model with a spatial lag dependent variable as well as a spatial moving average process in the error will be more appropriate. In the other cross-sections, the spatial dependence has been adequately dealt with. A spatial Breusch-Pagan test for spatial heteroskedasticity clearly indicates that heteroskedasticity patterns remain in the specification.

Characteristics variables: proportion of persons below 19 years, proportion of workers, proportion of unemployed, community equipments, density of population appear to be stable over time. Some of them (proportion of persons below 19 years, proportion of workers and proportion of unemployed) are highly significant in the unrestricted cross-sectional estimates. Note that the average price of water becomes significant only from 1990.1 on. The intercept varies widely but is not significant.

Table 8 reports the results from the asymptotic least squares for the two sets of restrictions. The minimum distance tests indicate no rejection for our restrictions (fixed slopes and all fixed parameters). Nevertheless, imposing additional restrictions may lead to rejection. For the first restriction, the estimated coefficients appeared to be significant except for the disposable income and the density of population variables. The other coefficients have the expected sign, except perhaps for the coefficient of the electricity price variable which is positive. This means that an increase in the electricity average price results in an increase in water consumption, which a priori appears to be surprising. Indeed, this result is in contradiction with the study of Hansen (1996) where the energy cross-price electricity is found to be negative. This result may indicate that consumers take into account the electricity block pricing structure where water consumption occurs effectively. For the second restriction, meteorological variables (rainfall and mean temperature) are no longer significant but are of the expected sign.

[14]GAUSS procedures to implement these calculations are available from the author on request.

Table 6: Unrestricted ML regressive spatial autoregressive estimates (continued)

	Cross-section estimates (and standard errors)					
Variable	1988.1	1988.2	1989.1	1989.2	1990.1	1990.2
Intercept	128.33	-157.20	-190.32	58.00	-653.15	-160.67
	(157.10)	(171.10)	(267.64)	(244.24)	(310.40)	(256.57)
Disposable Income	0.524	0.036	-0.564	-0.887	0.238	0.063
	(0.665)	(0.590)	(0.717)	(0.736)	(0.605)	(0.605)
Water price	-1.346	-0.860	-1.508	-1.702	-2.093	-1.391
	(1.275)	(1.195)	(1.082)	(1.209)	(1.116)	(1.137)
Electricity price	0.194	0.189	0.483	0.125	0.599	0.349
	(0.130)	(0.108)	(0.264)	(0.235)	(0.284)	(0.240)
Rainfall	-0.946	0.269	-1.220	-1.384	0.862	-0.552
	(0.380)	(0.406)	(0.373)	(0.426)	(0.376)	(0.318)
Mean temperature	-1.138	16.087	3.419	8.371	25.220	2.924
	(6.469)	(8.481)	(6.521)	(6.257)	(9.068)	(6.989)
Persons $<$ 19 years	-1.396	-1.694	-1.028	-1.239	-0.576	-0.882
	(0.575)	(0.530)	(0.596)	(0.607)	(0.600)	(0.582)
Workers	-2.662	-1.478	-2.194	-0.726	-2.724	-1.716
	(0.651)	(0.585)	(0.700)	(0.695)	(0.647)	(0.659)
Unemployed	-2.977	-2.624	-3.321	-2.740	-3.477	-3.124
	(0.603)	(0.561)	(0.603)	(0.614)	(0.621)	(0.616)
Equipments	-0.409	-0.678	-0.281	-0.211	-0.084	-0.078
	(0.352)	(0.330)	(0.359)	(0.366)	(0.359)	(0.359)
Density of population	0.166	0.123	0.225	0.004	0.316	0.081
	(0.400)	(0.374)	(0.405)	(0.415)	(0.412)	(0.410)
Spatial lagged variable	0.273	0.119	0.281	0.323	-0.004	0.310
	(0.282)	(0.325)	(0.288)	(0.292)	(0.335)	(0.292)
Diagnostics tests, (p-value)						
LM spatial error	0.704	3.911	0.219	0.826	0.269	0.374
	(0.401)	(0.047)	(0.639)	(0.363)	(0.603)	(0.540)
Spatial B-P heteroskedas.	13.753	7.778	19.826	13.538	16.634	25.224
	(0.131)	(0.556)	(0.019)	(0.139)	(0.054)	(0.002)
Number of obs	115					

Table 7: Unrestricted ML regressive spatial autoregressive estimates (end)

	Cross-section estimates (and standard errors)					
Variable	1991.1	1991.2	1992.1	1992.2	1993.1	1993.2
Intercept	-248.72	-423.66	-72.88	-21.60	-50.34	-320.07
	(348.13)	(330.08)	(346.51)	(317.15)	(322.76)	(350.23)
Disposable Income	0.067	0.100	0.127	-0.225	0.147	-0.269
	(0.612)	(0.518)	(0.457)	(0.443)	(0.439)	(0.483)
Water price	-1.212	-3.595	-2.177	-0.729	-1.949	-3.092
	(1.167)	(1.025)	(0.898)	(0.915)	(0.689)	(0.757)
Electricity price	0.443	0.500	0.166	0.215	0.147	0.482
	(0.346)	(0.297)	(0.334)	(0.276)	(0.306)	(0.281)
Rainfall	-1.384	-0.127	0.011	-0.911	0.002	-0.356
	(0.606)	(0.306)	(0.384)	(0.264)	(0.280)	(0.319)
Mean temperature	-9.589	10.831	10.025	3.240	13.723	7.660
	(8.177)	(8.561)	(7.646)	(12.058)	(9.354)	(8.216)
Persons $<$ 19 years	-0.829	-0.458	-0.697	-0.463	-1.138	-0.445
	(0.651)	(0.605)	(0.566)	(0.537)	(0.550)	(0.567)
Workers	-2.533	-2.088	-2.209	-1.613	-1.796	-2.517
	(0.719)	(0.685)	(0.689)	(0.638)	(0.646)	(0.652)
Unemployed	-3.088	-2.928	-3.067	-2.878	-2.479	-3.132
	(0.690)	(0.646)	(0.623)	(0.586)	(0.606)	(0.644)
Equipments	-0.022	0.135	-0.166	0.015	-0.555	0.204
	(0.402)	(0.379)	(0.364)	(0.335)	(0.345)	(0.365)
Density of population	-0.096	0.453	0.114	0.187	-0.004	0.523
	(0.457)	(0.435)	(0.416)	(0.392)	(0.406)	(0.421)
Spatial lagged variable	0.386	0.134	0.484	0.260	0.233	0.287
	(0.273)	(0.299)	(0.243)	(0.285)	(0.299)	(0.284)
Diagnostics tests, (p-value)						
LM spatial error	6.386	0.074	7.863	1.194	9.798	1.093
	(0.011)	(0.785)	(0.005)	(0.274)	(0.001)	(0.295)
Spatial B-P heteroskedas.	13.458	20.644	21.721	15.658	19.717	42.132
	(0.142)	(0.014)	(0.009)	(0.074)	(0.019)	(0.000)
Number of obs	115					

Table 8: Asymptotic least squares estimates

Variable	Restriction 1 (fixed slopes) coefficients	stderr	t-stat	Restriction 2 (all fixed parameters) coefficients	stderr	t-stat
Intercept	—	—	—	4.496	40.886	0.109
Disposable Income	0.092	0.145	0.637	0.205	0.170	1.201
Water price	-1.998	0.253	-7.878	-2.385	0.264	-9.017
Electricity price	0.240	0.055	4.324	0.201	0.038	5.226
Rainfall	-0.367	0.082	-4.474	-0.079	0.041	-1.901
Temperature	5.780	1.992	2.901	0.594	0.425	1.399
Persons $<$ 19 years	-0.991	0.155	-6.390	-0.959	0.183	-5.241
Workers	-2.104	0.175	-11.965	-2.131	0.201	-10.584
Unemployed	-2.931	0.167	-17.544	-2.800	0.196	-14.217
Equipments	-0.223	0.097	-2.292	-0.271	0.115	-2.347
Density of population	0.149	0.111	1.341	0.175	0.130	1.338
Spatial lagged variable	0.271	0.078	3.437	0.289	0.091	3.163
R^2	0.692			0.603		
$\bar{R}^2$	0.656			0.565		
$\chi^2(5\%)$	94.165			133.972		
d.o.f	143			121		
Number of obs ($N \times T$)	1380					

The spatial coefficient is also highly significant, which confirms the modelling framework. Here, the spatial behavior may be viewed in two ways. First, we can argue that households are actually influencing their neighbours. The water consumption behavior of other households affects the consumption of a given household through social proximity. In this sense, the estimated spatial coefficients represent a direct measure of externality. The significant spatial pattern may also be interpreted as the reaction of households with respect to the availability of water resources.

4 Conclusion

The aim of this paper was to specify panel data autoregressive spatial processes in the framework of a minimum distance method. For these models, the minimum-distance estimator is an attractive alternative to a direct maximum likelihood on the overall (T×N) observations. Indeed, in the context of random effects specification, our approach is easier to compute and may be consistent when the maximum likelihood would not be. Furthermore, specification tests can be run as well. We have presented the area of residential water demand including electricity aspects as a direct application. Investigating data, we show that spatial patterns actually exist. Estimation results indicate that our approach is feasible. Nevertheless, the specification remains to be improved. Future works may incorporate spatial heteroskedasticity in the error term. We can also take advantage of the Chamberlain approach in a random effects model (Chamberlain (1984)). These implementations may be done keeping in mind the asymptotic considerations induced by this modelling.

5 Appendix

5.1 Distribution of French daily residential water consumption between household tasks

Table 9: Water-using tasks (*Source:* "General Company of Waters")

Water consuming tasks	Proportion
Drink	1%
Cooking (heated)	6%
Dish washing (heated)	10%
Clothes washing (heated)	12%
Toilets	39%
Personal hygiene (heated)	20%
Outdoor use (including sprinkling)	6%
Other uses	6%

5.2 Definition of variables and related data sources

The data were provided by "Vivendi, La Compagnie Générale des Eaux (Direction Régionale Est" (General Company of Water(s)), "la Direction Générale des Impôts de la Moselle" (Regional Tax Center), "le Centre Départemental de la Météorologie de la Moselle" (Regional Center of Meteorological Studies) and "l'Institut National de la Statistique et des Études Économiques" (National Institute of Statistics and Economic Studies). The variables used in this study are the following.

Dependent variable: aggregate residential water consumption per community expressed in m3 per house.

Explanatory variables

• Water average price in "FF" per m3 (computed to include fixed charges),
• Electricity average price in "FF" per kwh,
• Disposable income per household paying taxes; available by year period, and then has been divided by 2 to obtain biannual values,
• Rainfall in m,
• Mean temperature in degree Celsius,
• Proportion of persons below 19 years,
• Proportion of workers,
• Proportion of unemployed,
• Index of equipments,
• Density of population,
• Spatial lagged dependent variable.

5.3 Computation of the contiguity matrix

Table 10: Characteristics of the distance matrix

Variables	statistics
Dimension (number of points)	115
Average distance between points	28.665
Distance range	85.135
Minimum distance between points	1
Maximum distance between points	86.135
Quartiles:	
First	13.317
Median	29.273
Third	41.641
Minimum allowable distance cutoff	5.362

The binary contiguity matrix W we use is created from information on the distance between communities. First, a matrix of distances D with elements d_{ij} based on latitude-longitude coordinates of the centroids from each community is computed using the Euclidean metric. In a second step, the information in the distance matrix is used to create a row-standardized spatial weights matrix W whose elements are defined as follows :

$$\omega_{ij} = \begin{cases} 1 & \text{if } d_{ij} \in [\mathrm{d_{inf}}; \mathrm{d_{sup}}] \\ 0 & \text{otherwise} \end{cases}$$

where $[\mathrm{d_{inf}}; \mathrm{d_{sup}}]$ is a specified critical distance. Here we do not have any prior notion of which distance ranges are meaningful. Then we arbitrarily choose a statistical meaningful one, that is the first and third quartiles. The reason for using such contruction in our study and not the usual "rook criterion" (common border based contiguity) is that the communities considered here do not constitute a block area and there are some isolated communities. Table 10 provides the main characteristics of the distance matrix.

Acknowledgements

This paper will be presented at the Fifth International Conference of the Society for Computational Economics (Computational Econometrics and statistics) in June 24-26, 1999, at Boston College. I am grateful to my advisor, Professor François Laisney, for his detailed constructive criticisms and continued encouragement. I also thank Olivier Alexandre for his suggestions. All errors remain my own.

References

Anselin, L. (1988): *Spatial Econometrics Methods and Models.* Dordrecht Kluwer Academic Publishers, Regional Research Institute, West Virginia University, first edn.

Anselin, L., and A. Bera (1998): "Spatial Dependence in Linear Regression Model with an Introduction to Spatial Econometrics," *Handbook of Applied Economic Statistics*, pp. 237–289.

Case, A. (1991): "Spatial Patterns in Household Demand," *Econometrica*, (59), 953 –965.

Chamberlain, G. (1984): "Panel Data," *Handbook of Econometrics*, (2), 1248–1313.

Cliff, A. D., and J. K. Ord (1981): *Spatial Processes Models & Applications.* London Pion.

Cressie, N. A. C. (1991): *Statistics for Spatial Data.* Willey-Interscience.

DRISCOLL, J. C., AND A. C. KRAAY (1998): "Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data," *The Review of Economics and Statistics*, pp. 549–560.

GOURIÉROUX, C., A. MONFORT, AND A. TROGNON (1985): "Moindres Carrés Asymptotiques," *Annale de l'INSEE*, (58), 91–122.

HANSEN, L. G. (1996): "Water and Energy Price Impacts on Residential Water Demand in Copenhagen," *Land Economics*, (72–1), 66–79.

INSEE (1998): *Tableaux de l' Economie Lorraine 96/97*. INSEE Lorraine.

KELEJIAN, H., AND I. PRUCHA (1997): "Estimation of Spatial Regression Models with Autoregressive Errors by Two-Stage Least Squares Procedures: A Serious Problem," *International Regional Science Review*, (20–1), 103–111.

——— (1998): "A Generalized moments Estimator for the Autoregressive Parameter in a Spatial Model," *International Economic Review*, (Forthcoming).

KODDE, D. A., F. C. PALM, AND G. A. PFANN (1991): "Asymptotic Least-Squares Estimation Efficiency Considerations and Applications," *Journal of Applied Econometrics*, (5), 229–243.

LESAGE, J. P. (1997): "Bayesian Estimation of Spatial Autoregressive Models," *International Regional Science Review*, (1–2), 113–129.

LESAGE, J. P., AND M. R. DOWD (1997): "Analysis of Spatial Contiguity Influences on State Price Level Formation," *International Journal of Forcasting*, (2), 245–253.

SILVERMAN, B. W. (1986): *Density Estimation for Statistics and Data Analysis*. Chapman & Hall, first edn.

WAND, M. P., AND M. JONES (1995): *Kernel Smoothing*. Chapman and Hall.