Estimation of spatial panel data models using a minimum distance estimator: application

Théophile Azomahou*

BETA-Theme, Université Louis Pasteur, Strasbourg,
Department of Economics
61, avenue de la Forêt Noire
F-67085 Strasbourg Cedex; France

Abstract

This paper considers estimating panel data spatial autoregressive models in the framework of minimum distance estimators. A spatial weighting matrix based on the distance between points is constructed to relate observations spatially. To overcome the computational difficulties that beset spatial processes, the model is estimated in two stages. First, the data are treated as T cross-sections, the parameters of which are consistently estimated by pseudo-maximum likelihood. A consistent asymptotic covariance matrix is computed as the norm of a quadratic form for the second stage. Minimum distance estimates are then derived under the restrictions of common slopes and complete equality of parameters. Finally, spatial elasticities are investigated. This framework is applied to estimating empirically spatial patterns in the residential demand for water using a lattice sample.

Key Words and Phrases: Minimum distance estimator, panel data, spatial dependence, water demand.

JEL Classification: C13; C23; D12

*phone: (33) 390 414 071; fax: (33) 390 414 050; e-mail: azomahou@cournot.u-strasbg.fr

1 Introduction

It is often of interest 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. There are many varied applied econometric applications. LeSage and Dowd (1997) use this methodology to examine the influences of spatial contiguity on state price level formation. A similar framework has been used by Case (1991) to describe spatial patterns in household demand for rice in some Indonesian districts. Recent examples of empirical work that explicitly incorporates spatial dependence includes, among others, the analysis of innovation decisions, Hautsch and Klotz (1999), the forecasting of cigarette demand using panel data, Baltagi and Li (1999), real wages variation to local and aggregate unemployment rates over time, Ziliak *et al.* (1999), and the estimation of a hedonic model for residential sales transactions, Bell and Bockstael (2000).

In a regression framework, spatial autocorrelation (more generally, spatial dependence) occurs when the dependent variable and/or the error term of a regression function is correlated at each location with observations of the dependent variable and/or values of the error term at other locations. As pointed out by Anselin (1988), ignoring this structure when it actual exists results in mis-specification and in estimation bias. While most studies focus on cross-sectional specifications, spatial models for panel data have not received much attention.

As outlined by Case (1991), fixed effect specifications can be used to control for spatial components in panel data. In some cases, when there is no intra-regional variation in variables of interest, a spatial modeling approach may be however more appropriate. This would be the case when the variation in the variable depends upon the 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 modeling spatial random effects, and shows that when specific effects are uncorrelated with the right hand side variables, there are clear benefits to a spatial specification. More generally, it can be argued that the equicorrelated structure of individual dependence that is typically specified in error-component models for panel data does not allow for distance decay effects. Moreover, this equicorrelation is associated with the time dimension and not the individual dimension of the data set. As a result, such a structure is not adequate for estimating spatial patterns in panel data. This study provides both theoretical and empirical advances on this topic.

We consider estimating panel data spatial autoregressive models within the framework of minimum distance estimation. We specify a mixed regressive spatial autoregressive model which defines a class of random fields, i.e. models derived from processes indexed by space, time and cross-sectional dimensions. We work with a row-standardized spatial weighting matrix, i.e. the spatial weighting 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 neighboring observations. The specification is assumed to be the true data generating process which relates observations with reference to points in space and time. The model is then estimated in two stages.

In order to overcome the computational difficulties that beset spatial processes (and assuming the errors to be normally distributed) the data are treated as T cross-sections in the first stage, the parameters of which are estimated by a pseudo-maximum likelihood procedure. Under suitable regularity conditions this stage provides both unrestricted consistent parameter estimates, including the spatial coefficient, and elements of scores which are used to compute the consistent asymptotic covariance matrix as a norm for a quadratic form for the second stage. The Minimum distance method is then applied under the assumption of a linear relationship between the auxiliary parameters and the parameters of interest in the estimating equations. We consider two different sets of restrictions to relate the time dimension of the panel: the common slopes (or fixed slopes) and the complete equality of parameters (or all identical parameters). The minimum distance estimates are computed for each case and are consistent and asymptotically efficient. Finally, time varying spatial elasticities based on the minimum distance estimates are investigated.

This specification is used for an empirical analysis of the spatial variations of the residential demand for water for the French department of "Moselle", including the effects of energy (electricity) price. At this stage it is important to explain why the price of electricity can be used as an additional regressor in the specification of the model and why the data at hand are appropriate to the spatial context.

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. Water is consumed by households jointly with different tasks involving in most cases sizable amounts of energy and other goods (appliances, etc.). Table 10 in Appendix 6.1 reports the daily distribution of French residential consumption of water between household tasks. We observe that about 50% of this distribution is concerned with water heating (mainly by electricity). We combine this consideration with spatial aspects for two reasons.

The first reason is an empirical concern. Several studies have pointed out the existence of a regionalized behavior for the consumption of water by households living the concerned municipalities. Such a behavior may also be linked to the availability of water resources. See for example INSEE (1998) for more details on this purpose.[1] Furthermore, as will be seen later, these municipalities have been split into spatial sectors for the purpose of water network management. In this context, the specification used 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. The second reason is attached to the theoretical framework. As outlined earlier and as will be seen in description of the data, there is no intra-regional variation in water prices, variations in this variable depend on the distance between municipalities. As a result, a spatial approach seems more appropriate and should be preferred to the pure fixed effect modeling. Furthermore, in our panel the number of cross-sections is larger than the number of waves. In this case, the framework such as the one suggested by Whittle (1954) cannot be applied.[2] All these reasons motivate the use of the spatial approach adopted here.

Section 2 presents the model. The proposed specification combines elements of spatial modeling and the panel data framework using a minimum distance approach.

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

[2]Whittle (1954) suggests that if panel data are available and if the time dimension is sufficiently large, $T > N$, one can consider e.g. a seemingly unrelated regression specification, or an error component model to permit for cross-sectional correlation, and estimate the cross-sectional correlations through the time dimension of the panel.

Section 3 is dedicated to the data. We describe the sample and basic descriptive statistics. Spatial correlograms are computed to check for spatial patterns. We also use a nonparametric density estimation to identify "spatial sector tendencies" in the distribution of the average price of water. Estimation results involving both parameter estimates and spatial elasticities are presented in Section 4. Concluding remarks are given in Section 5.

2 Spatial model for panel data

Let us consider a spatial autoregressive model for panel data containing a spatial lag of the response variable as an additional regressor. Such a model has the following structure:

$$y_{it} = \sum_{j \neq i} \rho \omega_{ij} y_{jt} + \sum_{k=1}^{K-1} x_{it}^{(k)} \beta_k + \varepsilon_{it} \quad |\rho| < 1, \tag{1}$$

$$i = 1, \cdots, N; \; j = 1, \cdots, N; \; t = 1, \cdots, T.$$

where y_{it} is the i-th observation on the dependent variable at period t, $x_{it}^{(k)}$ is the i-th observation for the k-th explanatory variable, y_{jt} is the j-th observation on the dependent variable contiguous to i. ρ is a scalar, the spatial coefficient, and the β's are $k-1$ parameters of the remaining explanatory variables. ρ and β are the parameters of primary interest to be estimated. ω_{ij} is an element of the spatial weighting matrix, the computation of which is given in Appendix 6.3. Relation (1) can be rewritten in a more convenient stacked vectors and matrices form.

For each time period, let $y = (y_1, \cdots, y_i, \cdots, y_N)'$, $X = (X_1, \cdots, X_i, \cdots, X_N)'$ and $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_i, \cdots, \varepsilon_N)'$. We organize the data as such due to the introduction in the sequel of a spatial weighting matrix W which remains fixed over time. Furthermore, this allows to take observations by cross-section. Hence, for each period, y and ε are of dimension $(N \times 1)$ and X is $N \times (K-1)$. The structure of the model implies that each cross-section follows a spatial autoregressive process. Then, in stacked form the model is

$$y = [Wy, X]\,\theta + \varepsilon, \quad \theta = \left(\rho, \beta'\right)', \tag{2}$$

where W is a known $(N \times N)$ spatial weighting matrix, usually containing first-order contiguity relations or functions of the distance between spatial units; again, see

Appendix 6.3 for the computation of W. Here, we work with a row-standardized version of W, i.e. W is normalized so that its rows sum to unity. This standardization produces a spatial lagged variable Wy, (also termed "regionalized variable") that represents an average of values from the neighbouring y. X is the matrix of explanatory variables, θ is a $(K \times 1)$ vector of unknown parameters. It is composed of the spatial coefficient ρ and the vector β of dimension $(K-1)$ of the other explanatory variables. Relation (2) is analogous to the multivariate lagged dependent variable model for time series regressions, with a spatial parameter ρ indicating the extent to which variations in y are explained by the average of the values of its neighbouring observations.

To simplify understanding of how the estimation procedure works, let us consider G-variates y_t for $t = 1, \cdots, T$, generated from $y_t = f(y_t, X_t, W; \theta_0) + \varepsilon_t$ where $\theta_0 \in \Theta \subset \mathbb{R}^K$, $y_t \in \mathbb{R}^G$, $X \in \mathbb{R}^P$, $\varepsilon_t \in \mathbb{R}^G$. Each wave t contains N cross-sections. W is a given time invariant square matrix of dimension $N \times N$ which relates observations spatially. Assume that the conditional distribution of ε_t given X_t is equal to the product of the conditional distributions for $t \neq s$. Furthermore, assume this distribution to be Gaussian with $E(\varepsilon_t) = 0$ and $E(\varepsilon_t \varepsilon_s') = \sigma_t^2 I$, for $t \neq s$.

The assumption of normally distributed errors allows us to estimate, in a first stage, the parameters of each cross-section separately by pseudo-maximum likelihood. This first stage estimation procedure is termed "pseudo" as it does not account for the time dimension of the panel. That is, the conditional distributions in this stage are assumed to be time invariant. In a second stage the minimum distance method may be applied and the time dimension of the panel is taken into account. Moreover this second stage provides efficient estimates, since the quadratic form to be minimized is optimally specified in the sense of Hansen (1982).

Formally, let $\hat{\theta} = (\hat{\rho}, \hat{\beta}', \hat{\sigma}^2)'$ denote the unrestricted pseudo-maximum likelihood estimates for parameters $\theta = (\rho, \beta', \sigma^2)'$ for each cross-section. That is

$$\hat{\theta} = \arg\max_{\theta \in \Theta} \sum_{i=1}^{N} \sum_{t=1}^{T} \psi\left(y_{it}, X_{it}, W; \theta\right), \tag{3}$$

where $\psi(\cdot)$ denotes the log likelihood function computed as

$$\psi(y, X, W; \rho, \beta, \sigma^2) = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln \sigma^2 + \ln |\Phi| - \frac{1}{2\sigma^2} \eta' \eta, \tag{4}$$

with $\eta = \Phi y - X\beta$, $\Phi = I_N - \rho W$ and $|\cdot|$ denotes the determinant. In a second stage one may use the unrestricted maximum likelihood estimator of the first stage to form the restricted minimum distance estimates by imposing several restrictions of the form $g(\hat{b}(\theta), a) = 0$. These restrictions link the set of parameters of interest $\mathcal{A} = a(\mathcal{P}) \subset \mathbb{R}^K$, $a_0 = a(P_0), \forall P \in \mathcal{P}$, where a_0 denotes the true value of a, and the set of auxiliary parameters: $\mathcal{B} = b(\mathcal{P}) \subset \mathbb{R}^H$ et $b_0 = b(P_0), \forall P \in \mathcal{P}$, where $b_0 = b(P_0)$ denotes the true value of b. The estimating equations are such that

$$g(b, a) = 0, \quad \text{with} \quad g(b(P), a) = 0 \Rightarrow a = a(P), \quad \forall P \in \mathbb{P}. \tag{5}$$

Expression (5) means that there exists a sequence $\hat{b}_n = \hat{b}_n(y_1, \cdots, y_i, \cdots, y_n)$ of estimators for b such that: (i) $\hat{b}$ converges towards $b_0 = b(P_0)$, P_0 a.s., (ii) the asymptotic distribution of $\sqrt{N}(\hat{b}_n - b_0)$, with a covariance matrix $\Sigma_0 = \Sigma(P_0)$, is

$$\sqrt{N}(\hat{b}_n - b_0) \underset{N\to\infty}{\overset{L}{\longrightarrow}} \mathcal{N}(0, \Sigma_0).$$

Under these conditions, the minimum distance 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 $g(\hat{b}(\theta), a_0)$. This leads to the minimization program

$$\hat{a} = \arg\min_{a\in\mathcal{A}} \Big[g(\hat{b}(\theta), a_0)\Big]' S_n \Big[g(\hat{b}(\theta), a_0\Big], \tag{6}$$

where $S_n \overset{\text{a.s.}}{\to} S_0$ is a positive definite symmetric matrix. The optimal choice for S is known to be the inverse of the covariance matrix of $g(\hat{b}(\theta), a_0)$. See e.g. Gouriéroux *et al.* (1985) and Kodde *et al.*(1991) for further details. Under usual regularity conditions, the estimator $\hat{a}(S_n)$ exists and is consistent. Furthermore, the following asymptotic distribution holds

$$\sqrt{N}\left(a_n(S_n) - a_0\right) \underset{N\to\infty}{\overset{L}{\longrightarrow}} \mathcal{N}\left(0, \Omega_0 = \Omega(S_0)\right). \tag{7}$$

In the case of spatial stochastic process, consistent estimation of S_0 is not trivial. Indeed, such a computation makes use of the spatial weighting matrix in the likelihood function. In order to form a consistent estimation of S_0, we propose the following approximation.

Let Ω_0 denote a consistent approximation of S_0 such that

$$\Omega_0 = \frac{1}{N}\left[J_0^{-1} I_0 J_0^{-1}\right], \tag{8}$$

where $J_0 = \text{diag}\{J_1, \cdots, J_T\}$ is a block diagonal matrix with elements

$$J = E\left(-\frac{\partial^2 \psi(y, X; W, \theta_0)}{\partial\theta\partial\theta'}\right), \tag{9}$$

and the elements of I_0 are given by

$$I = E\left(\frac{\partial\psi}{\partial\theta}(y, X; W, \theta_0)\frac{\partial\psi}{\partial\theta}(y, X; W, \theta_0)'\right). \tag{10}$$

A consistent estimator $\hat{\Omega}$ of Ω is obtained as follows. Let $\psi_i(y, X; W, \rho, \beta, \sigma^2)$ denote the log-likelihood for one observation. More formally we have

$$\begin{aligned}\psi_i(y, X, W; \rho, \beta, \sigma^2) = & -\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln\sigma^2 + \frac{1}{N}\ln|\Phi| \\ & -\frac{1}{2\sigma^2}\left[\sum_{j\in J}\left(\mathbf{1}_{[i=j]} - \rho\omega_{ij}\right)y_j - \sum_k X_i^{(k)}\beta_k\right]^2,\end{aligned} \tag{11}$$

where $j = 1, \cdots, J$ is the set of spatial units contiguous to a unit i and $\mathbf{1}_{[i=j]}$ denotes an indicator function. Taking partial derivatives of (11) with respect to the parameters yields

$$\frac{\partial}{\partial\beta_k}\psi_i(\cdot) = \frac{1}{\sigma^2}\left[\sum_{j\in J}\left(\mathbf{1}_{[i=j]} - \rho\omega_{ij}\right)y_j - \sum_h X_i^{(h)}\beta_h\right]X_i^{(k)}, \tag{12}$$

$$\frac{\partial}{\partial\rho}\psi_i(\cdot) = \frac{1}{\sigma^2}\left[\sum_{j\in J}\left(\mathbf{1}_{[i=j]} - \rho\omega_{ij}\right)y_j - \sum_h X_i^{(h)}\beta_h\right]\sum_{j\in J}\omega_{ij}y_j + N^{-1}\xi, \tag{13}$$

$$\frac{\partial}{\partial\sigma^2}\psi_i(\cdot) = -\frac{1}{2\sigma^2} + \frac{1}{2\sigma^4}\left[\sum_{j\in J}\left(\mathbf{1}_{[i=j]} - \rho\omega_{ij}\right)y_j - \sum_k X_i^{(k)}\beta_k\right]^2, \tag{14}$$

with

$$\xi = \frac{\partial}{\partial\rho}\ln(I - \rho W) = -\text{tr}\left([I - \rho W]^{-1}W\right).$$

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

3 Data

The department of "Moselle" consists of about 730 municipalities out of which 115 have been selected for the empirical study of households' demand for drinking water.[3] Households living in these municipalities 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 is collected biannually from 1988.1 to 1993.2, a balanced panel of 1380 spatial observations. Some variables do not require important changes before being used. Others have been constituted from information available in the last municipal inventory.[4] This section describes the sampling and relevant features of the variables. See Appendix 6.2 for an overview on data sources.

3.1 Sampling and descriptive statistics

The first step of this study was the collection of data. Since this kind of data had never been collected before, two important issues arose from a closer look of the consumption values. The first was the identification of households' consumption. The network manager (a private operator) provides water services to the subscribers, i.e. citizens living in individual houses or in collective blocks of flats (for instance council flats), as well as industrial consumers and businesses. The households' demand gathers together individual user consumption and collective user consumption. Most of the households living in collective lodging do not yet have meters that indicate accurately the amount of their consumption. Also, for these consumers the charge for water is included in the rent. We can then suppose that the households concerned are not aware of the necessity to control their budget with respect to water expenses.

Moreover, there are also blocks of flats sheltering small businesses. In the case when a household living in a collective lodging gets a business linked to his subscriber regime, former's consumption of water cannot be distinguished from the latter's. A similar issue occurs for some households living in individual houses. Indeed, for those

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

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

among them who possess e.g. farms the identification of purely domestic volumes is difficult. For all of these reasons, and in order to reduce the evaluation errors as well as to be sure that the target sector corresponds to the residential one, we have selected subscribers connected to the network of drinking water with a main water capacity of 15mm in diameter, when this information was available. Despite this choice, we cannot exclude that some marginal consumption values coming from small businesses or other consumption different from domestic consumption is still in the 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 because they existed under a high level of aggregation. This concerns only very a few unionized municipalities. The non-unionized municipalities display a half year water volume. Unions result from the gathering of municipalities; we use union data to estimate the volume consumed when municipalities' data is missing. 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, we face two possible cases: either only some municipalities composing the union are considered or the details of the volume consumed are not available. In the former case we suppose that the consumption in the other municipalities varied in the same proportion. In the latter case, the average weight of each municipality in the union is computed. As a result, the data possess two characteristics which make their biannual use delicate.

On the one hand, the water reading frequency ran from at least a quarterly period to an annual one whilst the pricing remains biannual. The accurate biannual readings are available for 1988.1 for all the municipalities, as well as the readings of 1993.1. From 1990.1 to 1992.2 some municipalities adopted an annual reading. In this case, and to reduce the cost induced by meter readings, the volumes for one semester are estimated from the consumption of a preceding year, where the duration between two readings does not always equal 52 weeks. Moreover, the calendar year is no longer taken into account, instead the period stretching from June to June is considered. On the other hand, we face differences in the frequency of data collection. The consumption reading frequency may vary from one year to another because of climatic hazards or other unforseeable parameters. To correct these biases, the

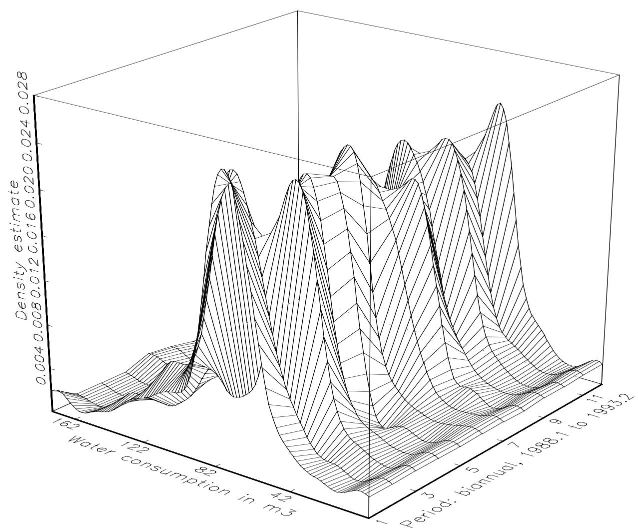

Figure 1: Distribution of the residential consumption for water, kernel density estimates per time period. The figure shows a mainly uni-modal distribution around 60 and 80 m3 for each period.

consumption values presented in this study are corrected to lead to a frequency of 52 weeks.

These two characteristics, estimated values and differences in the reading frequency, are possible sources of measurement errors. Despite this, it is important to note that the percentage of the initial sample that cannot be used because of identification problems is about 3.6% and the percentage of the data that had to be reconstructed in order to obtain consumption values is about 5.8% . In light of these issues, we compute a nonparametric density estimation to have a closer look at the distribution of consumption values. Figure 1 shows the results. We notice mainly a uni-modal distribution between 60 and 80 m3 for each period. These estimates reinforce our recording target, i.e. the consumption of residential subscribers. This means that the identification issue we have faced for consumption values seems to be well handled.

We have recorded the aggregate water consumption per municipality in cubic

Table 1: Descriptive statistics of water consumption and the average price.

Period	Consumption in m3 mean	std.	min.	max.	Average price in FF 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

meters per house. Since urban municipalities are larger than rural ones, each consumption value has been divided by the total number of households per community in 1990, the year of the last available inventory. This is also when the last general population 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. See Appendix 6.2 for an overview of the list of variables.

National statistics indicate an average water consumption tendency of about 120 m3 per house per year. These figures vary somewhat: old houses consume less water 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 averages of recorded consumption are of the same magnitude. Minimum values can be considered as the consumption of rural municipalities. These tendencies are also indicative of the standard of living of the population considered. As a whole, there are no outliers in consumption values. Note however some high values exist 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 consumption in addition to purely domestic consumption.

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(*).

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

(*): Values are expressed in thousands of FF.

Table 4: Descriptive statistics of municipalities's characteristics(*).

Variable	mean	std.	min.	max.
Proportion of persons <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

(*): Statistics are computed for 1990, the year of reference.

Again these statistics support, on average, our recording target sector: the water consumption of residential subscribers.

Disposable income statistics are characterized by very low values. Consider for example the year 1990 where the minimum values are the lowest, i.e. 31,820 FF per taxed household. This gives a monthly disposable income of 2,651.66 FF. Supposing that this household is made up of a single member, the latter roughly earns the so-called "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 actual household incomes by 30% on average.[5] This under-estimation is extremely high for the self-employed (43%), even more so 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 in social benefits and only a slight increase in taxes, the "Moselle" departement is below national indicators.

Average price values clearly indicate relevant patterns. The average price continuously increases over the twelve biannual periods. 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 onwards the tendency is even higher. This last tendency indicates an important modification in the structure of water price. As a result, the price variable suggests a clustering pattern. It also presents an increasing dispersion within clusters with stable mini-

[5]These figures can be found e.g. in Tableaux de l'Economie Lorraine 1997/1998 (Tables of Lorraine Economics), INSEE (1998).

mum values (around 3,5 FF). All these figures are examined more carefully in the next section.

Finally, note that the meteorological variables (rainfall and temperature) presented here are not dummy variables as is usually the case in the literature. The values on these variables were recorded by the Regional Center of Meteorological Studies. As we may expect, the first semester values are less than those of the second semester.

3.2 Distribution of the average price of water

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

According to the water supplier, the municipalities are split into two sectors, however 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 the 115 municipalities, 65.2% belong to sector 1. The sectors correspond to two distinct areas of water management. This spatial arrangement is mainly due to network management issues (water transportation, treatment to make water drinkable, etc.) and is closely linked to elements of water prices.[6] 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. However, the average price of water varies from one community to another when the fixed charges of water are included. Moreover, the laws on water of November 1992, the so-called "M-49 directive", have strongly modified the working orders of water agencies.[7] This modification translated

[6]To make ideas clear, we computed the correlation coefficient between the average price of water and the sector dummy for the twelve time periods: -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.

[7]Set up on November 10th, 1992 (its implementation date) the "M-49 directive" imposes the rule of budget balance to water services (supply and cleaning up). They are no longer allowed to include expenditures on water spending (building up and maintenance of network, equipment,

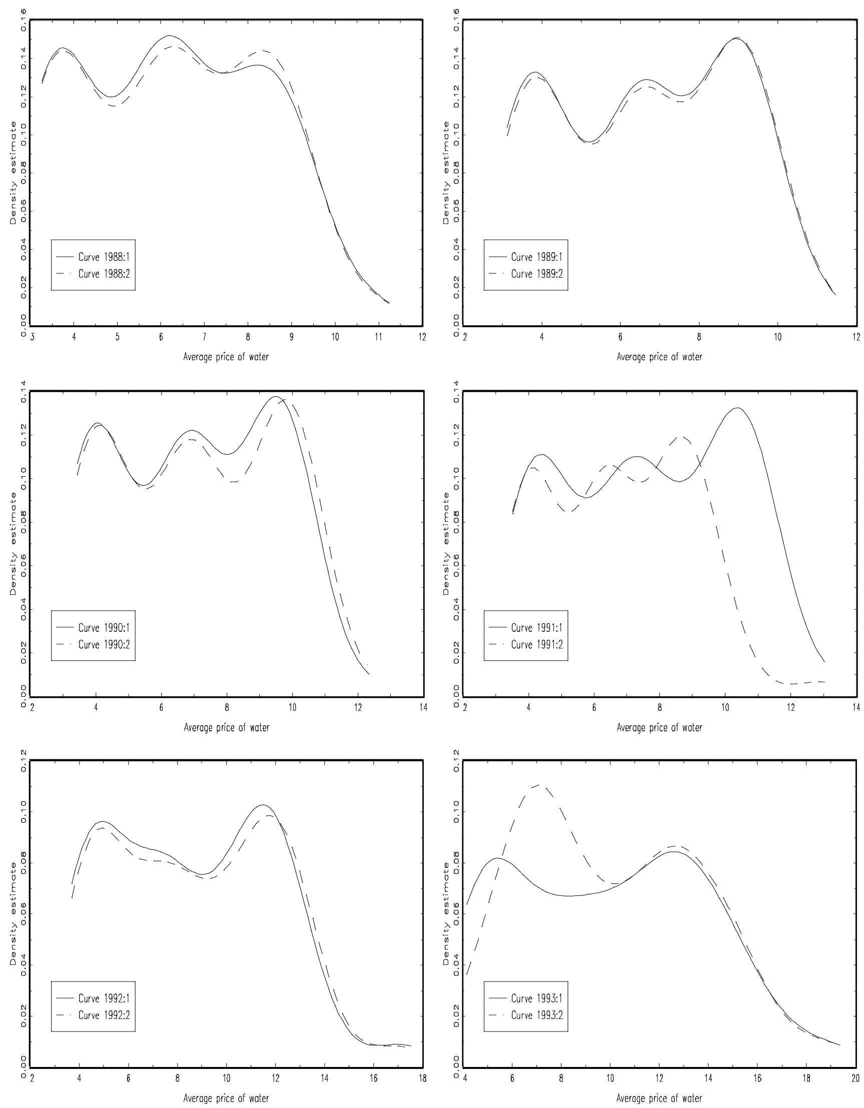

Figure 2: Distribution of the average price of water, kernel density estimates. We observe three modes from 1988.1 to 1991.2. In 1991.1 and 1992.2, the central mode starts disappearing. From 1992.1 on, only two modes remain.

into high increases in water prices. The aim is to let customers pay for the effective price of water, instead of 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. See e.g. Silverman (1986) and Wand and Jones (1995) for details. Figure 2 shows the kernel density estimate of the average price of water for each time period. In the estimation procedure, we use the Epanechnikov kernel and the cross-validation method for the choice of the bandwidth. Two main conclusions can be stated. First, 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 mainly be explained by the modifications that occurred in water pricing in 1992. These modifications may be due to the "M-49 directive" which resulted in a change in water pricing. Not only did the price increase continuously as indicated by descriptive statistics, but now, two sectors appear clearly from 1992. Second, the distribution reveals also that there may be three sectors up to 1992.1. Thus, sector design effects clearly appear in the average price of water. As a result, we may expect a within-sector behavior regarding water consumption as well as a spatial effect. However, the lack of information on the exact number of sectors at the beginning of data collection period precludes the use of a switching regression framework.

3.3 Testing spatial autocorrelation

We introduce various analytic methods which are of value in assessing the spatial scale of a process. The variables of interest are: the water consumption, the average price of water and the disposable income. We use Gs-statistics which provide a measure of overall spatial association as well as observation-specific spatial association. See Appendix 6.4 for a definition. These statistics are computed by defining a set of neighbouring municipalities. For each location, neighbouring municipalities are considered as those which fall within a distance band.

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. The results of the tests are reported in Table 5. We observe a significant

cleaning up...) in their general budget.

Table 5: $G^*(\delta)$-test for specific spatial autocorrelation.

	Consumption		Water 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

value for consumption (except for 1992.1 and for 1993.1) which is indicative of a spatial clustering of low values. The G-statistics for the average price of water are all highly significant. Then, a spatial dependence for high values occurs. Except for 1993, spatial autocorrelation for income values is rejected. The G-statistic test has a "static aspect" and does not provide information on the spatial dynamics of the process. This issue is handled using spatial correlograms. See Appendix 6.4 for a definition.

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 weighthing matrices using spatial correlograms. 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 the higher order contiguity results in unconnected spatial units, i.e. spatial units for which the corresponding row in the contiguity matrix consists only in zeros. For further technical details and discussions on spatial correlograms see Cliff and Ord (1981) and Cressie (1991).

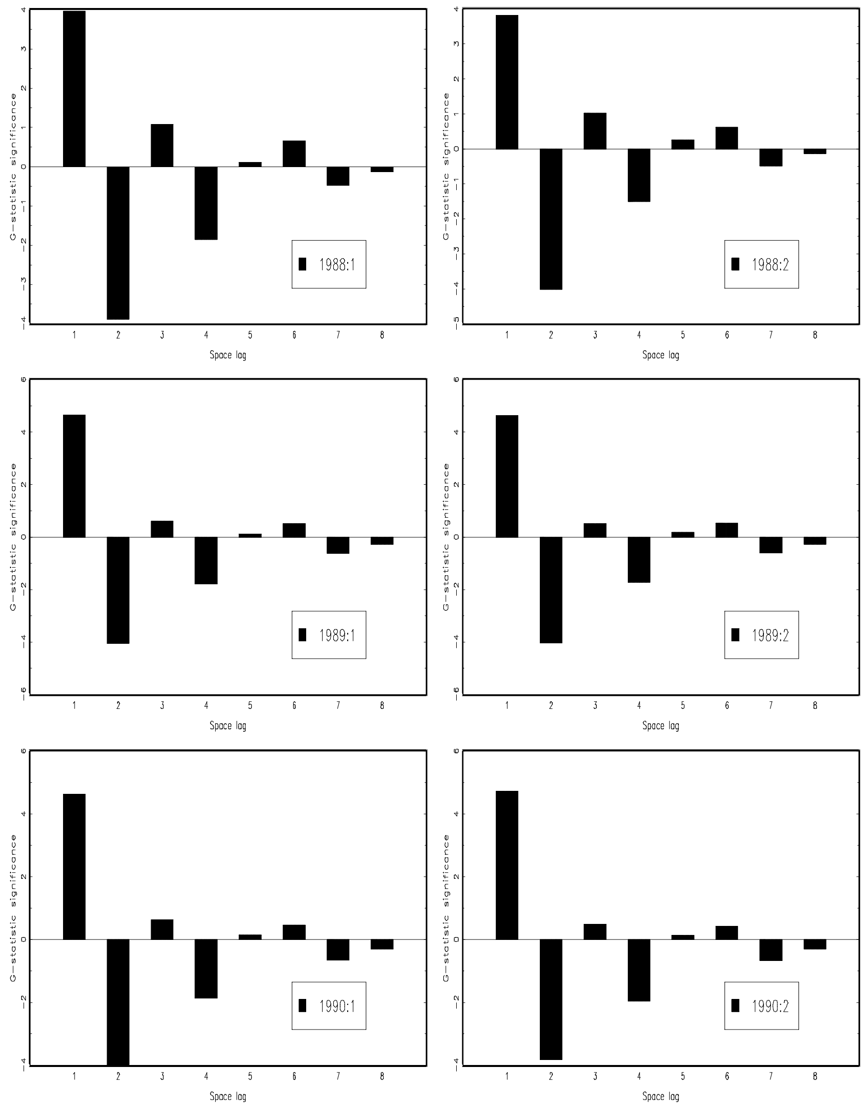

Figure 3: Estimation of spatial correlograms for the residential consumption of water from 1988.1 to 1990.2. Up to eight spatial lags on the X-axis and the t-value of the G-statistics on the Y-axis. The first two lags of each correlogram are highly significant, indicating spatial dependence which decreases with spatial lags.

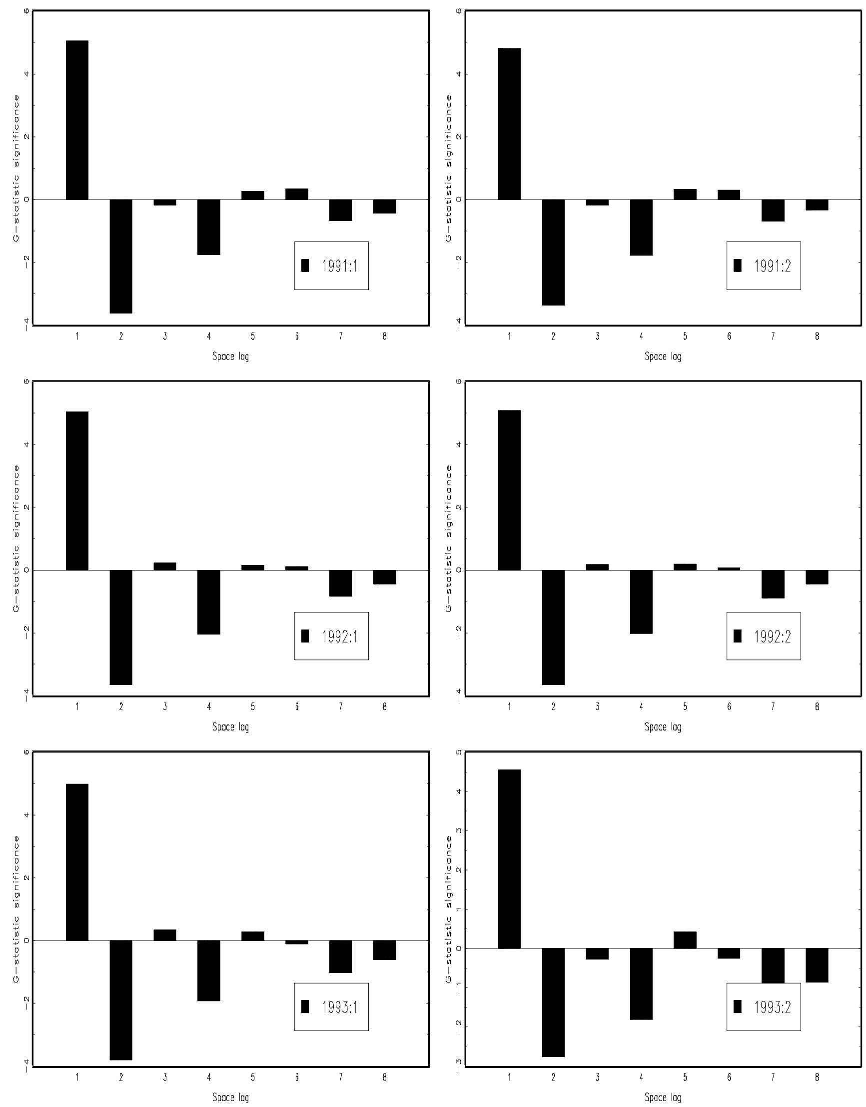

Figure 4: Estimation of spatial correlograms for the residential consumption of water from 1991.1 to 1993.2. Up to eight spatial lags on the X-axis and the t-value of the G-statistics on the Y-axis. The first two lags of each correlogram are highly significant, indicating spatial dependence which decreases with spatial lags.

The results of the estimated spatial correlograms for each time period are reported in Figures 3 and 4. Spatial lags are reported on the X-axis (up to eight lags are computed), and the t-statistics associated to the G-values are indicated on the Y-axis. To ease presentation, other statistics related to spatial correlograms (expectation, standard deviation etc.) are not reported here. 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.

4 Estimation results

We use the theoretical framework sketched above to carry out empirical estimation on data described in the previous section.[8] Tables 6 and 7 present the unrestricted pseudo-maximum likelihood estimates for the twelve time periods. A Lagrange multiplier test rejects the alternative spatial error specification for most cases except for 1988.2, 1991.1, 1992.1 and 1993.1. See e.g. Anselin (1988, p. 66-72) for a definition of these tests.

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.

We use the following characteristics variables: proportion of persons below 19 years, proportion of workers, proportion of unemployed, municipalities's equipment

[8] GAUSS procedures to implement the calculations in this paper are available from the author on request.

Table 6: Unrestricted pseudo-maximum likelihood 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)
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)
Equipment	-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 dep. var.(i)	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(ii)						
LM spatial error.(iii)	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.(iv)	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		

(i): "Spatial dep. var." means the spatial lagged dependent variable, computed as Wy.

(ii): p-values are in parenthesis.

(iii): Lagrange multiplier test for the spatial model.

(iv): Spatial Breusch-Pagan test for spatial heteroskedasticity.

Table 7: Unrestricted pseudo-maximum likelihood 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)
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)
Equipment	-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 dep. var.(i)	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(ii)						
LM spatial error.(iii)	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.(iv)	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		

(i): "Spatial dep. var." means the spatial lagged dependent variable, computed as Wy.

(ii): p-values are in parenthesis.

(iii): Lagrange multiplier test for the spatial model.

(iv): Spatial Breusch-Pagan test for spatial heteroskedasticity.

Table 8: Minimum distance estimates.

Variable	Restriction 1 (common slopes) coef.	std.err	t-stat.	Restriction 2 (equality of parameters) coef.	std.err	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
Equipment	-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 dep. var.(*)	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	
degree of freedom		143			121	
Number of obs. $(N \times T)$			1380			

(*): "Spatial dep. var." means the spatial lagged dependent variable, computed as Wy.

All estimates are carried out with a significance level of 5% .

and the density of population. 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.

The minimum distance procedure implemented in the second stage is based on the hypothesis that

$$H_0 : \{\theta / \exists a \in \mathcal{A} \subset \mathbb{R}^q : g(b(\theta, a) = 0\}, \tag{15}$$

where the function $g(.)$ is valued in $\mathbb{R}^r$. It is usually assumed that $\partial g / \partial \theta'$ and $\partial g / \partial a'$ are respectively of rank r and q. For empirical estimation one needs to assume a functional form for the estimating equations $g(b(\theta, a)$. Here, we assume that $g(b(\theta, a)$ is linear with respect to the parameters of interest, a. From this specification, minimum distance estimates are obtained by imposing two restrictions. The first restriction is that of common slopes or fixed slopes, expressed as

$$g(\hat{b}(\theta), a) = \begin{pmatrix} \hat{\theta}_x^1 - \theta_x \\ \hat{\theta}_x^2 - \theta_x \\ \vdots \\ \hat{\theta}_x^T - \theta_x \end{pmatrix}, \tag{16}$$

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 complete equality of parameters or all identical parameters, i.e.

$$g(\hat{b}(\theta), a) = \begin{pmatrix} \hat{\theta}_1 - \theta \\ \hat{\theta}_2 - \theta \\ \vdots \\ \hat{\theta}_T - \theta \end{pmatrix}, \tag{17}$$

with $\hat{b} = (\hat{\theta}_1, \cdots, \hat{\theta}_T)'$ and $a = \theta$. For each case, the minimum distance estimates are computed by generalized least squares procedures associated with the estimation of Ω_0 in relation (8) as $\hat{\Omega}^{-1} = [\hat{J}^{-1}\hat{I}\hat{J}^{-1}]^{-1}$. Table 8 reports the results from the minimum distance estimates. To check for the validity of the specification $g(.) = 0$ underlying the parameters of interests, we also computed a specification test based on the minimum distance estimates. Since we assumed that $g(.)$ is linear in a, relation (18) turns out to be

$$H_0 : \{\exists a : b(\theta) = H(\theta)a\}, \tag{18}$$

where $b(\theta)$ is a vector of dimension r and $H(\theta)$ is a matrix of one and zeros of dimension $r \times q$. The statistic test is written as $T_n = NT[b(\hat{\theta}_n - H(\hat{\theta}_n)\hat{a}_n]'\hat{\Omega}_n[b(\hat{\theta}_n - H(\hat{\theta}_n)\hat{a}_n]$ and the distribution of T_n is such that $T_n \geq \chi^2_{1-\alpha}(r-q)$.

For each restriction, the minimum distance tests computed in Table 8 are 94.165 and 133.972 respectively. Given the associated degree of freedom, there is no rejection of the adopted specification both under fixed slopes and all identical parameters restrictions. For the first restriction, the estimated coefficients appear to be significant except for 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 seems a priori, surprising.

Indeed, although complementarity between the two goods (water and electricity) may be expected, the positive sign for the parameter of the average price of electricity indicates that, for the sample concerned, water and electricity display substitutability patterns. This means that an increase in the average price of electricity may result in more water consumption by residential consumers. This a priori surprising result is in contradiction with the study of Hansen (1996) where the energy cross-price parameter is found to be negative. Our cross-effects estimates suggest that changes in the average price of electricity may induce modifications in the distribution of residential water consumption for different uses. That is to say, the share of residential water consumed in connection with electricity may decrease with the price of electricity, whereas the remainder (the share of residential water consumed without energy) does not. We noticed in Section 1 that about 50% of daily residential water consumption in France is concerned with heating. Hence, the remaining 50% may explain our result partly. This may also indicate that consumers take into account the electricity block pricing structure where water consumption occurs effectively. For the second restriction, meteorological variables (rainfall and temperature) are no longer significant but are of the expected sign.

The spatial coefficient is also highly significant, which confirms the modeling 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 an externality. The significant spatial pattern may also be interpreted as the reaction of households with respect to the availability of water resources.

Time varying spatial elasticity at means is computed as

$$\mathcal{E}_{kt} = \hat{\beta}_k \left(\frac{\bar{X}_{kt}}{\bar{Y}_t} \right), \qquad k = 1, \cdots, K, \;\; t = 1, \cdots, T.$$

The above relation is termed "spatial elasticity" as it makes use of the variation in variables of interest between municipalities. Results of the elasticities for the dispos-

Table 9: Spatial elasticities at means.

	Restriction 1 (common slopes)			Restriction 2 (equality of parameters)		
Period	Income	Price	Spatial(*)	Income	Price	Spatial(*)
1988.1	0.038	-0.180	0.265	0.085	-0.215	0.282
1988.2	0.037	-0.181	0.264	0.084	-0.217	0.281
1989.1	0.036	-0.185	0.264	0.083	-0.221	0.282
1989.2	0.036	-0.181	0.265	0.081	-0.217	0.282
1990.1	0.038	-0.192	0.265	0.086	-0.229	0.283
1990.2	0.039	-0.198	0.264	0.087	-0.236	0.281
1991.1	0.038	-0.204	0.265	0.086	-0.243	0.282
1991.2	0.039	-0.212	0.266	0.087	-0.253	0.284
1992.1	0.042	-0.241	0.266	0.093	-0.287	0.284
1992.2	0.041	-0.247	0.267	0.092	-0.295	0.283
1993.1	0.042	-0.276	0.266	0.095	-0.330	0.284
1993.2	0.043	-0.297	0.267	0.096	-0.354	0.285

(*): "Spatial" means the spatial lagged dependent variable.

able income, the average price of water and the spatial lagged dependent variable are reported in Table 9.

Although the coefficients (from Table 8) used in computing the elasticities are highly significant, these elasticities are very weak. They do not exceed 1% in absolute value. Elasticities related to the restriction of complete equality of parameters are higher than those associated with the common slopes restriction. We also observe that the values vary weakly over time.

5 Conclusion

The aim of this study was to theoretically and empirically specify panel data spatial autoregressive models in the framework of a minimum distance. The methodological twist of the paper is to separately estimate parameters of cross-sections for different time periods using pseudo-maximum likelihood procedures, and to use a minimum distance procedure to impose restrictions across waves, such as common slopes and complete equality of parameters. For the proposed specification we argue why the minimum distance approach is attractive and may be preferable to pure fixed effects framework or a restricted maximum likelihood method.

Empirically, we examine whether households look to a reference group when making water consumption decisions. In order to answer this question, we make use of spatial modeling in a panel data context. Moreover, the paper presents a consistent and asymptotically efficient minimum distance estimator that is applied in the empirical investigation. In the end, from an empirical viewpoint, what do we learn about the residential demand for water? The estimated spatial lagged parameter is strongly significant, which means that households living in the same geographic area have approximately similar water consumption behaviors. This provides us with a measure of exernality which is not usually observable. We also find evidence that consumers respond jointly to the average price of water and electricity, not only to the average price of water.

Finally, panel data for lattice samples provides the opportunity of improving estimation efficiency in spatial models. Future effort should be directed to spatial fixed effects specifications where the error component term may also be spatially autoregressive.

6 Appendix

6.1 Water using tasks

Table 10: Water-using tasks.

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

Source: "General Company of Waters"

Table 10 provides the distribution of French daily residential water consumption between household's tasks. It is observed that about 50% is concerned with heating.

6.2 Data sources and list of variables

The data comes from different sources. They were provided by: "la Compagnie Générale des Eaux, Direction Régionale Est" (General Company of Water), "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). Table 11 gives a complete list of the variables used in the empirical analysis.

Table 11: Variables used in the empirical analysis.

Variables designation	
* Residential water consumption,	aggregate values by municipality.
* Average price of water,	computed to include fixed charges.
* Average price of electricity,	in FF per kwh .
* Disposable income,	evaluated on households paying taxes.
* Mean rainfall,	in m.
* Mean temperature,	in degree Celsius.
Characteristics of municipalities:	
* Proportion of persons $<$ 19 years.	
* Proportion of unemployed.	
* Proportion of workers.	
* Density of population.	
* Index of equipment.	

6.3 Computation of the spatial weighting matrix

The binary spatial weighting matrix W we have used is created from information on the distance between municipalities. First, a matrix of distances D with elements d_{ij} based upon latitude-longitude coordinates of the centroids from each municipality is computed using the Euclidean metric. Characteristics of the distance matrix are summarized in Table 12. In a second step, the information in the distance matrix is used to create a row-standardized spatial weighting matrix W whose elements ω_{ij} are defined as follows.

$$\omega_{ij} = \begin{cases} 1 & \text{if } d_{ij} \in [\delta_1; \delta_2] \\ 0 & \text{otherwise} \end{cases}$$

where $[\delta_1; \delta_2]$ is a specified critical distance band. Here we do not have any prior notion of which distance ranges are meaningful. Hence, we choose a statistically meaningful one, i.e. the first and third quartiles: $\delta_1 = 13.317$ km and $\delta_2 = 41.641$ km. The reason for using such construction and not the usual common border criterion is that the municipalities considered here are not all contiguous.

Table 12: 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

6.4 Definition of spatial statistics

For further technical details and discussions on the Gs-statistics and spatial correlograms, see e.g. Cliff and Ord (1981), Cressie (1991) and Getis and Ord (1992).

6.4.1 The Gs-statistics

Formally, for a cutoff distance δ, the G-statistic denoted $G(\delta)$ is defined as

$$G(\delta) = \frac{\sum_i \sum_j w_{ij}(\delta) z_i z_j}{\sum_i \sum_j z_i z_j}, \qquad i = 1, \cdots, N, \;\; j \in J,$$

where z_i is the value observed at location i, $w_{ij}(\delta)$ stands for an element of the symmetric (unstandardized) spatial weighting matrix for distance δ and J is the set of j neighbours to i. Inferences on $G(\delta)$ are typically based on a standardized t-value. This is computed by substracting the theoretical expectation before dividing the result by the theoretical standard deviation: $t_G = \{G(\delta) - E[G(\delta)]\}/SD[G(\delta)]$, with the notation SD denoting the standard deviation. Based on asymptotic considerations, the t-value follows a standard normal distribution and the significance of $G(\delta)$ can be stated by comparing the computed t-value to it probability in the usual normal Table.

For each observation i, the $G^*(\delta)$ statistic for a specific spatial association indicates the extent to which that location is surrounded by high values or low values of the variable of interest. Formally, for a given distance δ

$$G^*(\delta) = \frac{\sum_j w_{ij}(\delta) z_j}{\sum_j z_j}.$$

For this statistic, $j = i$ is included in the sum symbol. This means that $G^*(\delta)$ provides a measure of spatial clustering that includes the observation under consideration. Inference about the significance of $G^*(\delta)$ is derived as for $G(\delta)$.

6.4.2 Statistic for spatial correlograms

Consider a system of N sites with random variables $x_1, \cdots, x_N$ and let the sites i and j be δth-order neighbours. Then, the δth-order sample spatial autocorrelation is given by

$$C(\delta) = \frac{N}{\Delta(\delta)} \frac{z'Wz}{z'z},$$

where $z' = (z_1, \cdots, z_N)$, $z_i = x_i - \bar{x}$, $i = 1, \cdots, N$, and $\Delta(\delta) = \sum_i \sum_j w_{ij}(\delta)$. Alternatively, this statistic may be rewritten as

$$C(\delta) = \frac{N}{\Delta(\delta)} \frac{\sum_{(i,j)} w_{ij}(\delta) z_i z_j}{\sum_i z_i^2}.$$

Observe that the symmetric form of W in the statistic means that each term appears twice in the summation. Readers are referred to the literature mentioned above for the computation of the means and the variances of these measures. The plot of $C(\delta)$ against δ yields the spatial correlogram.

Acknowledgements

This paper has been presented at the Ninth International Conference on Panel Data, at the University of Geneva, June 22-23, 2000. Suggestions from participants of the conference are acknowledged. I am grateful to François Laisney for his detailed constructive criticisms and continued encouragement in developing this work. I also thank David Belsley for helpful comments. Financial support of the "National School of Water and Environment Engineering of Strasbourg" (GSP, Engees-Cemagref) is gratefully acknowledged. All errors and shortcomings remain my own.

References

Anselin, L. (1988): *Spatial Econometrics Methods and Models.* Dordrecht: Kluwer.

Baltagi, B. H., and D. Li (1999): "Prediction in the Panel Data Model with Spatial Correlation," *Texas A&M University, Department of Economics*, p. 17 pages.

Bell, K. P., and N. E. Bockstael (2000): "Applying the Generalized-Moments Estimation Approach to Spatial Problems Involving Microlevel Data," *The Review of Economics and Statistics*, (82–1), 72–82.

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

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

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

Getis, A., and J. K. Ord (1992): "The Analysis of Spatial Association by Use of Distance Statistics," *Geographical Analysis*, (24), 189–206.

Gouriéroux, C., A. Monfort, and A. Trognon (1985): "Moindres Carrés Asymptotiques," *Annale de l'INSEE*, (58), 91–122.

Hansen, L. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, (50), 1029–1054.

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

Hautsch, N., and S. Klotz (1999): "Estimating the Neighborhood Influence on Decision Makers: Theory and an Application on the Analysis of Innovation Decisions," *Center of Finance and Econometrics, University of Konstanz*, p. 26 pages.

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

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., 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.

Whittle, P. (1954): "On Stationary Processes in the Plane," *Biometrica*, (41), 434–449.

Ziliak, J. P., B. A. Wilson, and J. A. Stones (1999): "Spatial Dynamics and Heterogeneity in The Cyclicality of Real Wages," *The Review of Economics and Statistics*, (81–2), 227–236.