A Switching Regression Approach to Spatial Patterns in Residential Water Demand

Théophile Azomahou

September 1999

BETA, Louis Pasteur University
Economics Department
GSP, Engees-Cemagref
61, Avenue de la Forêt Noire
F-67085 Strasbourg Cedex; France
phone: (+33) 3.90.41.40.71
fax: (+33) 3.90.41.40.50
e-mail: azomahou@cournot.u-strasbg.fr
url: http://cournot.u-strasbg.fr/azomahou/

Abstract

This study explores spatial regimes variation and the impact of public sector water pricing for municipality aggregate residential water demand including electricity price effects. I compare estimations from two cross-sections (1988.1 and 1993.1) on a French lattice sample, and propose a parametric spatial autoregressive regime switching model where water average price is approximated by a linear spline based on nonparametric regressions. I find evidence of spatial dependence. Consumers respond to both water and electricity average price. Changes in electricity price induce modifications of water consumption distribution according to patterns of water use. Public sector pricing results in shifts of role in regimes between periods.

Keywords: Linear spline, public water pricing, spatial regime, water demand

JEL Classification: C14, H41, Q25

Résumé

La présente étude s'intéresse à la modélisation statistique de la demande domestique agrégée d'eau potable, ainsi qu'aux effets de la tarification de l'eau dans le cadre d'une spécification spatiale avec changement de régime. Les effets d'agrégation spatiale sont en partie pris en compte au travers d'une régression non-paramétrique de la consommation sur le prix moyen de l'eau. Le résultat de cette estimation suggère une approximation de la variable «prix moyen» par un spline linéaire, qu'on introduit ensuite dans un modèle paramétrique spatial auto-régressif avec changement de régime. L'estimation est faite par la méthode du maximum de vraisemblance. Cette spécification est utilisée pour mesurer les effets d'une réforme de la tarification de l'eau en comparant les estimations issues d'un même échantillon de données à deux périodes différentes, 1988.1 et 1993.1. Les résultats d'estimation et les tests confirment l'existence d'effets spatiaux dans les comportements de consommation d'eau. La prise en compte d'effets croisés montre qu'une variation du prix moyen de l'électricité induit une modification de la distribution de la consommation d'eau selon différents usages. On observe également que les régimes changent de rôle entre 1988.1 et 1993.1.

1 Introduction

In many economic fields, it is often relevant to consider the spatial variation of some phenomenon. As noticed by Anselin (1988), several factors could lead to this. The first is a byproduct of measurement errors in observations for contiguous spatial units; for instance the arbitrary delimitation of spatial units, spatial aggregation, and most importantly, the problem of spatial externalities and spill-over effects across the boundaries of the spatial units. Indeed, in many practical situations, data are collected only at an aggregate level. Therefore, there may be little correspondence between the spatial scope of the phenomenon studied and the delineation of the spatial units. As a result, measurement errors are likely. Thus, the errors for one observation are likely to be related to the errors in the neighboring observations. Examples of this may rely on the use of data on population, employment and other economic activity collected for administrative units such as states, provinces, counties etc. So, any situation where the spatial organization and the spatial structure of phenomena matter may be considered.

More fundamentally, the second factor of spatial dependence follows from the importance of space as an element in structuring explanations of human behavior. That is, location and distance result in various interdependences. Thus, what is observed somewhere is partially determined by what happens elsewhere in the system. Another type of spatial effect is spatial heterogeneity. It is related to the lack of stability over space. As a result, in a regression framework, functional forms and regression parameters may vary with location. This source of spatial variation is likely to occur in econometric models estimated on cross-sectional data. In contrast to the spatial dependence case, the lack of uniformity over space is due to several factors such as unequal spatial units, different areas and shape. For example, urban built-up areas have unequal population or unequal income levels, and regions have various degrees of technological development.

If spatial heterogeneity can be handled through econometric specifications such as varying parameters, random coefficients or various forms of structural change like switching regressions, spatial dependence invalidates some methodological results and has made the development of a specialized set of methodologies necessary, Anselin (1988) and Cressie (1991). Estimation issues of spatial processes follow from the multi-directional aspect of spatial data, in contrary to the uni-directional variation of time series.

Various empirical studies are concerned with the spatial framework. Case (1991) applied this methodology to the demand for rice purchased in Indonesian markets using data on districts. It is shown that, when district-specific effects are uncorrelated with the right hand side variables of the regression function, there are clear benefits to spatial modelling. It is 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. LeSage (1995) has studied the spatial aspects of variation in housing values and found evidence of spatial correlation between housing values and home characteristics.

The purpose of this paper is two-fold. First, it is concerned with the spatial empirical analysis of residential water demand in the French "Department of Moselle", including electricity price effects.[1] As indicated by Hansen (1996), when estimating the determinant factors of residential water demand, we may expect to observe the indirect effects of energy factors, according to water consumption between different water-using tasks. Indeed, water is consumed by households in connection with different goods which involve use of water and in most cases sizeable amounts of energy. As noticed in a previous study, Azomahou (1999a), about 40% of the daily distribution of French residential water consumption between different water uses is concerned with water heating, mainly by electricity.

The second aspect of this work consists in evaluating the relevance of public water pricing. Indeed, in France, water is a public good. As a result, water pricing pertains to public sector policies. The data we used fell within an interesting time period: before and after the so-called "laws on water", Guellec (1995). These laws have strongly modified price structures, increasing water price in the same proportion for all residential consumers. It has also modified the working order of water agencies. As a result, several structural shifts may be expected.

Following the approach suggested by Quandt (1958) and Anselin (1990), I consider modelling these issues in a regime switching framework. Both spatial dependence

[1]The "Department of Moselle" is located in the north-east of France. The municipalities selected for this study are those for which we achived obtaining reliable informations.

and spatial heterogeneity are considered. In this context, the specification we will describe 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 paper is organized as follows. Section 2 presents the main features of the data. The aggregation issue is handled by separating data into proper regimes where structural shifts occur only between regimes. Within regimes, consumption behavior is assumed to be homogeneous according to the spatial specification. Section 3 is dedicated to the empirical modelling and estimation. We proceed in two stages. First, we use a nonparametric regression based on cubic splines to fit the functional form between water consumption and water average price. The estimates suggest an approximation of the average price variable by a linear spline. In the second stage, we specify a parametric regime switching spatial autoregressive model, incorporating the linear spline. Then, estimation is carried out by maximum likelihood.[2] The results are discussed in Section 4. Section 5 concludes the study.

2 Data

The "Department of Moselle" is made of about 730 municipalities out of which 115 neighbouring municipalities have been selected for the empirical study of households demand for drinking water. Households living in these municipalities are supplied with drinking water by a private operator. The data considered here constitute the first lattice collected from the French network of drinking water distribution, with a biannual frequency (from 1988.1 to 1993.2). A general and detailed data analysis (sample selection, descriptive statistics and several tests for spatial autocorrelation) can be found in Azomahou (1999a). We refer the reader to that study for a complete description and give here only the information necessary for interpretation of the results.

The present study is only concerned with the two cross-sections 1988.1 and 1993.1.

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

Table 1: Description of variables (115 observations)

Variable	period	mean	std.	min.	max.
Water consumption (in m3)	1988.1	69.68	27.75	1.11	153.15
Water average price (in FF)	1988.1	6.28	2.11	3.24	11.29
Electricity average price (in FF)	1988:1	0.94	0.02	0.93	1.00
Mean temperature (in Celsius)	1988.1	8.92	0.33	8	9.58
Disposable income (in 1000 of FF)	1988	57.51	8.28	33.38	75.24
Water consumption (in m3)	1993.1	72.14	26.51	0.81	157.33
Water average price (in FF)	1993.1	9.97	3.97	4.09	19.46
Electricity average price (in FF)	1993.1	0.98	0.01	0.97	1.01
Mean temperature (in Celsius)	1993.1	8.82	0.22	8.32	9.47
Disposable income (in 1000 of FF)	1993	66.97	11.76	34.33	97.68
Population:					
Proportion of persons $<$ 19 years	1990	0.28	0.04	0.13	0.31
Density of population	1990	1.10	2.60	0.0038	14.61
Employment:					
Proportion of workers	1990	29.96	4.39	11.92	37.82
Proportion of unemployed	1990	9.78	4.03	2.70	23.62

These cross-sections correspond to the beginning and to the end of the data collection period. The reasons for considering only these two time periods will become clear in the sequel. For the moment note that it allows to measure the effects of water pricing reform in France. This pricing reform pertains to public policies. Characteristics of the variables used are reported in Table 1 (see also the Appendix for the definition of variables and for data sources).

The organization and management of water utilities in France pertain to public service liability. The water price results from a negociation between local authorities and the water utilities which may be the local collectivity (municipality) itself or a private company. The Municipalities concerned by this study are supplied with drinking water by a private firm. These municipalities are organized into two sectors. We denote each sector by a dummy variable (dummy 1 for sector 1 and zero for sector 2). Out of 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 the network management issues (water transportation and various treatments to make water drinkable) and is closely linked to water price.[3] The water price is decomposed into two parts: the marginal price and the fixed charges. The marginal price of water is the same within a given sector but varies between sectors. Thus, we know that there is no intra-sectoral variation in the marginal price. The fixed charges vary both within and between sectors. Having incorporated the fixed charges, we obtain an average price which varies from one municipality to another (see Appendix for the computation of the average price). Then, the two sectors are used to define the regimes.

Moreover, by the so-called "M-49 directive", the laws on water have strongly modified the working orders of water agencies. Set up on November 1992, the "M-49 directive" imposes to water utilities (supply and cleaning up) the rule of budget balance. This forbids resorting to the general budget to support water related expenses (building up and maintenance of network, equipments, cleaning up...). This modification has been translated into a high increase of water price in the same proportion for all residential consumers. The aim is to let customers pay for the effective price of water. The use of two cross-sections on the same sample allows to measure the effect on residential water demand of this change in water pricing; that is the efficiency of the public sector water pricing.

3 The model

In this section we suggest a two-step empirical modelling of residential water demand based on data. A parametric regime switching regressive spatial autoregressive model is implemented using estimates from nonparametric regressions. The latter allow to fit the regression functional which is obtained by regressing water consumption (response variable) on water average price (predictor).

3.1 Nonparametric estimation

As pointed out in Section 2, the water average price is strongly spatially structured which suggests a spatial non-linearity for this variable. Given this, it does not seem

[3]We have computed the correlation coefficient between the average price variable and the sector dummy: (-0.33 for 1988.1 and -0.41 for 1993.1). There is evidence of correlation.

a priori tenable to assume a uniform slope for the average price variable. So, we study the monotonicity of the demand function with respect to the average price variable only, using nonparametric regressions based on cubic splines.[4] We assume that consumption and average price values are related by an unkown functional form:

$$y_i = f(x_i) + \varepsilon_i \qquad i = 1, \cdots, N \tag{1}$$

where y_i and x_i denote respectively consumption and average price variables. ε_i is the error term. We assume that $E(\varepsilon_i) = 0$, $V(\varepsilon_i) = \sigma^2$ and $E(\varepsilon_i \varepsilon_j) = 0$, $\forall i \neq j$.

Let us consider as an approximation of the true function $f(x_i)$, the cubic spline $S(\lambda)$ that minimizes the modified sum of squares residuals:

$$S(\lambda) = \lambda \sum_{i=1}^{N} (y_i - f(x_i))^2 + (1-\lambda) \int_{\mathbb{R}} f''(x)^2 \mathrm{dx}, \qquad 0 \leq \lambda \leq 1 \tag{2}$$

where $\int_{\mathbb{R}} f''(x)^2 \mathrm{dx}$ is the roughness penality which is used to quantify local variations. λ is the smoothing parameter. It defines the relative weighting of the residual sum of squares and the integrated squared second derivative. It is known that λ controls the trade off between reducing variance and reducing the possible bias in the predicted values. For $\lambda = 1$, the relation (2) is reduced to the function which minimizes the sum of squared residual errors (the linear regression of y on x). A λ value of 0 gives the function which minimizes the roughness penality (the interpolating cubic spline). Varying λ between 0 and 1 will allow for more variation in the fitted curve until it interpolates the data points. Here, we use the criterion of "generalized cross-validation" for the optimal choice of λ. Details on this methodology can be found in Silverman (1985) and Härdle (1990).

The results of smoothing for each period are reported in Figure 1. Scatterplots of data are represented. The curves indicate the cubic splines and the corresponding 95% confidence limits estimates. The straight line denotes the linear approximation of the regression function. We notice that the assumption of a decreasing residential water demand function with respect to the average price values can not be rejected.

[4] Another way to proceed consists in using $\sqrt{N}$ consistent semiparametric regression framework, Robinson (1988).

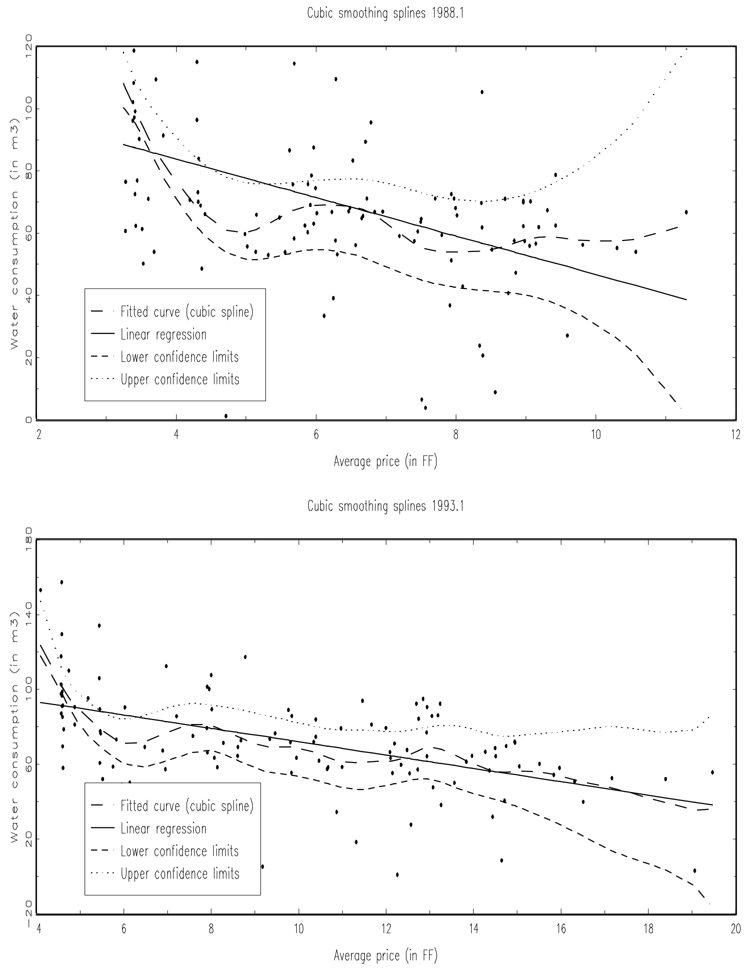

Figure 1: Nonparametric regressions. (Scatterplots of data, generalized cross-validation cubic splines, lower and upper 95% confidence limits and linear regressions fitted are reported. The fitted demand curves show non-linearity with respect to the water average price.)

However, what is not tenable is the hypothesis of a linear relation. Note also that these functions are not decreasing for some range of the predictor (around 6FF for 1988.1 and around 8FF for 1993.1). We obtain a significant and decreasing demand function for the average price values from 3FF to 4.5FF and from 4FF to 6FF respectively for 1988.1 and 1993.1, and a non-significant decreasing demand function otherwise with a weaker slope. Thus, we gained a new insight about water consumption behavior with respect to prices, that is to say, for the sample concerned, residential consumers are more responsive to lower price values. As a result, we can argue that a water pricing policy consisting in raising the water price in the same proportion for all households may not be optimal. This result cannot be observed from the linear fitting and suggests the use of a linear spline approximation for the water average price variable.[5] The proposed linear spline is specified as:

$$\overline{\beta}x_k + \underline{\beta}(x_k - \bar{x}_k)^+ \tag{3}$$

with,

$$(x_k - \bar{x}_k)^+ = \mathcal{M}ax(0, x_k - \bar{x}_k) \tag{4}$$

where x_k denotes the average price variable, and $\bar{x}_k$ denotes the break point of changing slope with $\bar{x}_k = 4.5$ for 1988.1 and $\bar{x}_k = 6$ for 1993.1. The relation (3) defines two slopes for the water average price predictor which will be included in a parametric model as described in the following section.

3.2 Parametric spatial regime model

To relate observations spatially, we used a binary contiguity matrix W constructed from the distance between these municipalities. First, a matrix of distances with elements d_{ij} based on latitude-longitude coordinates of the centroids from each municipality is computed using the Euclidean metric. Then, the information in the distance matrix is used to create a row-standardized spatial weights matrix W whose

[5]May be, this form results from the omission of other explanatory variables in the nonparametric regression. Note however that smoothing in higher dimension in our case is not relevant because of the limited number of observations.

elements are defined as follows:

$$\omega_{ij} = \begin{cases} 1 & \text{if } d_{ij} \in \mathcal{A} \\ 0 & \text{otherwise} \end{cases}$$

where $\mathcal{A}$ is a specified critical distance band. $\mathcal{A}$ is obtained from semivariograms estimation.[6]

We consider a specification of known transition point which yields for two observable regimes. Shifts between regimes are characterized by a dummy variable:

$$\mathcal{D}_i = \begin{cases} 1 & \text{if } i \in S_1 = \{i : i \in 1, \cdots, N_1\} & \text{Regime 1} \\ 0 & \text{if } i \in S_2 = \{i : i \in N_1 + 1, \cdots, N_2\} & \text{Regime 2} \end{cases}$$

where i denotes a given municipality. The first N_1 observations pertain to sector 1. They define the first regime. The remaining observations $N_1 + 1$ to N_2 pertain to sector 2, that is the second regime. The resulting spatial regime switching model has the following structure:

$$\begin{aligned} y_i = & \sum_{j \neq i} \rho \omega_{ij} y_j \\ & + \mathcal{D}_i \left(\sum_{k \in S_1} X_{i1}^{(k)} \beta_1^{(k)} + \epsilon_{i1} \right) \\ & + (1 - \mathcal{D}_i) \left(\sum_{k \in S_2} X_{i2}^{(k)} \beta_2^{(k)} + \epsilon_{i2} \right) \end{aligned} \tag{5}$$

$$|\rho| < 1; \qquad i = 1, \cdots, N, \quad j = 1, \cdots, N; \qquad i \neq j.$$

where y_i is the dependent variable and X_{i1}, X_{i2} are a set of explanatory variables, each of them conditioned by a regime. Recall that one element of X_i, say, $X_{i(1,2)}^{(k)}$

[6] Here we do not have any prior notion of which distance ranges are meaningful for the municipalities concerned. Moreover, these municipalities are not all contiguous. As a result, we cannot base the W matrix on municipalities "common border" criterion. Thereby, we use geostatistical tools (semivariograms estimation) to compute the elements of the spatial contiguity matrix by the so-called "Range". In variograms framework, the "range" is defined as the distance for which points are spatially dependent, Cressie (1985). Details on this construction are presented in Azomahou (1999b).

denoting the water average price variable is given by the linear spline in relation (3). In equation (5), each regime is characterized by a different value for the parameters β_1 and β_2. The structural change is then confined to $\beta^{(k)}_{(1,2)}$. The spatial factor ρ is assumed to be common to regimes. An element of the contiguity matrix W is given by ω_{ij}. Here, we work with a row-standardization version of W so that $\sum_j \omega_{ij} = 1$. As a result, $\sum_{j \neq i} \omega_{ij} y_j$ represents the mean consumption of neighboor municipalities for a given municipality i. The total number of observations N consists of N_1 in subset 1 and N_2 in subset 2 such that $N = N_1 + N_2$.

ϵ_{i1} and ϵ_{i2} are errors terms in the respective subsets S_1 and S_2. They are assumed to be i.i.d. and normally distributed with $E(\epsilon_{i1}) = E(\epsilon_{i2}) = 0$, $V(\epsilon_{i1}) = \sigma_1^2$, $V(\epsilon_{i2}) = \sigma_2^2$ and $\mathrm{cov}(\epsilon_{i1}, \epsilon_{i2}) = 0$. The resulting block variance-covariance matrix Ω of errors has the following structure:

$$\Omega = \begin{bmatrix} \sigma_1^2 I_{N_1} & O \\ O & \sigma_2^2 I_{N_2} \end{bmatrix} \tag{6}$$

with I_{N_1} and I_{N_2} as identity matrices of dimension corresponding to the number of observations in each subset. The relation (6) is usually referred to as a groupwise heteroskedastic specification. It means that the error variance is different according to regimes but is constant within regimes. As a result, the variance parameters σ_1^2 and σ_2^2 may be estimated directly from the residuals for each regime, provided that enough observations are available.

It is well known that spatial models of the form (5) cannot be estimated consistently by least squares. The spatial dependence in various spatial autoregressive models shows many similarities to the more familiar time-wise dependence. Therefore, one would expect the properties of least squares estimation for models with lagged dependent variables and/or serial correlation to hold with the spatial case. However, this is not so. The lack of a direct analogy results from the complex multi-directional characterization of dependence in space. This motivates the use of appropriate estimation methods among which the maximum likelihood.

Indeed, in space, there are many directions in which shifts can take place. On a regular lattice structure of one space lag, we obtain four directions for the "rook" and "bishop" criteria. The number of possible locations increases to eight for a "queen" type of contiguity. 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) and Kelejian and Prucha (1997b). Recently, LeSage (1997) has shown that Gibbs sampling provides a consistent alternative framework. For more details, see Anselin and Bera (1998) and Kelejian and Prucha (1997a).

Let $\theta = (\rho, \beta_1', \sigma_1^2, \beta_2', \sigma_2^2)'$ denote the vector of parameters to be estimated. The maximum likelihood estimation $\hat{\theta}_{ML}$ of θ is obtained as:

$$\hat{\theta}_{ML} = \arg\max_{\theta\in\Theta} \sum_{\substack{i=1\\ i\neq j}}^{N} \psi_i \left(y, Wy, X; \theta | \mathcal{D}_i\right) \tag{7}$$

with the usual assumptions on parameters space and parameters distribution. The log-likelihood function conditional on $\mathcal{D}_i$ is given by:

$$\begin{aligned}
\psi\left(\cdot, \theta | \mathcal{D}_i\right) = &-\frac{N}{2}\ln(2\pi) - \left[N_1 \ln\sigma_1^2 + (N-N_1)\ln\sigma_2^2\right] \\
&- \frac{1}{2\sigma_1^2}\sum_{\substack{i=1\\ i\neq j}}^{N_1}\left[\sum_{j\in J}\left(\mathbf{1}_{[i=j]} - \rho\omega_{ij}\right)y_j - \sum_{k\in S_1} X_{i1}^{(k)}\beta_1^{(k)}\right]^2 \\
&- \frac{1}{2\sigma_2^2}\sum_{\substack{i=N_1+1\\ i\neq j}}^{N_2}\left[\sum_{j\in J}\left(\mathbf{1}_{[i=j]} - \rho\omega_{ij}\right)y_j - \sum_{k\in S_2} X_{i2}^{(k)}\beta_2^{(k)}\right]^2 \\
&+ \frac{1}{N}\ln|I - \rho W|
\end{aligned} \tag{8}$$

where $\mathbf{1}_{[i=j]}$ denotes an indicator function, I is the identity matrix, W is the matrix of spatial influences and $|\cdot|$ denotes the determinant. This likelihood function is useful for our specification because the switching between regimes is determined by the dichotomous variable $\mathcal{D}_i$. Tests for structural stability are often associated with regime swithching specifications. Details can be found in Anselin (1990).

4 Estimation results

Three sets of estimates are presented. We start with a non-regime model where spatial dependence is included in the form of a spatial lag for the dependent variable. Then, we compare these estimates with those following from a spatial regime switching. So, in a first step, we do not take into account the regime switching, ignoring the issues related to structural stability. The results for the two cross-sections are reported in Table 2.

Table 2: Maximum likelihood estimates of spatial non-regime model (Explained variable: residential water consumption in m3 per municipality, per semester.)

	Cross section 1988.1			Cross section 1993.1		
Variable	coef.	std.err	t-value	coef.	std.err	t-value
Spatial lag	-0.449	0.299	-1.503	-0.754	0.333	-2.264
Water price	-1.763	0.405	-4.351	-2.678	0.740	-3.618
Spline price	1.875	0.495	3.786	2.466	0.742	3.321
Electricity price	0.296	0.052	5.592	0.249	0.076	3.257
Income	0.705	0.603	1.169	-0.0002	0.378	-0.0004
Temperature	0.955	5.292	0.180	11.227	7.997	1.403
Density of population	-1.850	0.757	-2.443	-1.845	0.709	-2.602
Persons $<$ 19 years	-0.843	0.495	-1.701	-0.264	0.460	-0.575
Workers	-2.552	0.606	-4.207	-1.384	0.563	-2.456
Unemployed	-2.570	0.560	-4.590	-2.209	0.546	-4.043
R^2	0.50			0.48		
σ^2 estimate	363.19			331.91		
Log likelihood	-502.45			-497.78		
Akaike	1024.91			1015.58		
Schwartz	1052.36			1043.60		

Both for 1988.1 and for 1993.1, most of the parameters for explanatory variables are significant, except for disposable income, mean temperature and the coefficient of the spatial lagged dependent variable. The latter is strongly significant for 1993.1. The average price variables (water and electricity) are of particular interest. Recall that our regressions include three average price variables. The first two follow from the linear spline specification for the water average price as indicated by the relation (3). Thus, for each cross-section, we have estimated two slopes for this predictor. Since we are interested in the cross-effects of energy price on residential water demand, we have also included electricity average price as additional regressor. The first slopes (for 1988.1 and for 1993.1) of water average price are strongly negatively significant.

The second slopes are also significant, but they are of the opposite sign. This is an expected result following from the linear spline characterization. Indeed, the slope of the water average price variable is the coefficient of this variable up to $\bar{x}_k$, and the sum of this predictor and the linear spline otherwise. Therefore, we obtained actually an estimated sum of slopes close to zero and non significant. The electricity average price influences water consumption positively. This is an unexpected and suprising result we will discuss.

In a second step, we have estimated a non-spatial model for groupwise heteroskedastic error under regime switching, using feasible generalized least squares (FGLS). The results are reported in Table 3. Several diagnostic tests are computed. We test for structural stability. We also computed Lagrange multiplier tests to determine the proper spatial specification (spatially lag dependent or/and spatial error). The groupwise heteroskedastic specification takes into account the possibility of not only having different regression coefficients in each regime, but different error variances as well. For the two cross-sections, the heteroskedasticity coefficients estimates are highly significant. These coefficients are the iterated FGLS estimates. The null hypothesis of the joint equality of coefficients is rejected by the Chow-Wald test (statistics are 24.75 and 15.97 respectively for 1988.1 and 1993.1).

The same indication is provided by a test on individual coefficients for electricity average price, disposable income and mean temperature for 1988.1, and for mean temperature and proportion of workers for 1993.1. For the sample concerned, these tests support the analysis of residential water demand in term of spatial regime switching. Tests for spatial lagged dependent variable and spatial error dependence are also carried out. For 1988.1 the Lagrange multiplier test provides strong evidence for spatial lagged dependent variable (statistic is 11.436 with a low p-value), whereas the spatial error dependence specification is rejected (statistic is 0.041 with a p-value of 0.83). The same indication is provided for 1993.1. As a result, the proper analysis of spatial regimes should include a spatial lag for the response variable, and not a spatial error. Based on these tests, we estimated a spatial lagged dependent model (relation 5) with regime switching. The results are reported in Table 4. Diagnostic tests support our modelling. The spatial lagged dependent variables coefficients are strongly negatively significant. These parameters are more significant than those obtained from the estimates of non-regimes specification.

Table 3: FGLS Estimates of groupwise heteroskedastic error regime model

	Cross section 1988.1			Cross section 1993.1		
Variable	coef.	std.err	t-value	coef.	std.err	t-value
REGIME 1:						
Water price	-2.141	0.574	-3.724	-3.087	2.614	-1.180
Spline price	2.194	0.681	3.221	2.938	2.668	1.101
Electricity price	0.340	0.076	4.455	0.396	0.206	1.922
Income	1.332	0.875	1.522	-0.200	0.752	-0.266
Temperature	-8.886	9.280	-0.957	-11.691	23.206	-0.503
Density of population	-1.728	0.831	-2.077	-2.391	0.904	-2.643
Persons $<$ 19 years	-0.619	0.646	-0.958	0.006	0.676	0.009
Workers	-2.327	0.758	-3.068	-0.792	0.871	-0.908
Unemployed	-2.387	0.759	-3.145	-2.057	0.868	-2.368
Skedastic coef.	345.178	56.367	6.123	428.189	69.923	6.123
REGIME 2:						
Water price	-0.864	0.498	-1.735	-2.972	0.645	-4.606
Spline price	1.260	0.698	1.805	2.590	0.623	4.152
Electricity price	0.080	0.074	1.086	0.051	0.079	0.645
Income	-1.081	0.774	-1.397	-0.194	0.376	-0.515
Temperature	22.002	8.148	2.700	37.615	8.218	4.576
Density of population	-2.104	1.732	-1.214	-1.531	1.180	-1.298
Persons $<$ 19 years	-1.706	0.775	-2.200	-1.163	0.537	-2.163
Workers	-1.882	0.965	-1.948	-2.807	0.676	-4.149
Unemployed	-3.760	0.802	-4.686	-2.598	0.637	-4.076
Skedastic coef.	247.658	55.378	4.472	120.561	26.958	4.472
R^2	0.591			0.538		
Log likelihood	-492.57			-486.25		
Akaike	1021.14			1022.64		
Schwartz	1070.55			1074.79		
Diagnostics tests:						
Chow-Wald (structural stability)	$\chi^2_{(9)}$: 24.75 p-value: 0.003			$\chi^2_{(9)}$: 15.971 p-value: 0.067		
Lagrange-M (spatial error model)	$\chi^2_{(1)}$: 0.041 p-value: 0.839			$\chi^2_{(1)}$: 0.000 p-value: 0.992		
Lagrange-M (spatial lag model)	$\chi^2_{(1)}$: 11.436 p-value: 0.000			$\chi^2_{(1)}$: 39.684 p-value: 0.000		

Table 4: Maximum likelihood estimates of spatial regime model

	Cross section 1988.1			Cross section 1993.1		
Variable	coef.	std.err	t-value	coef.	std.err	t-value
Spatial lag	-0.864	0.297	-2.908	-0.990	0.320	-3.088
REGIME 1:						
Water price	-2.407	0.542	-4.439	-3.604	2.219	-1.623
Spline price	2.362	0.631	3.740	3.442	2.262	1.521
Electricity price	0.433	0.076	5.691	0.481	0.175	2.740
Income	2.025	0.836	2.422	0.052	0.635	0.081
Temperature	-13.013	8.558	-1.520	-12.033	19.394	-0.620
Density of population	-1.450	0.772	-1.878	-2.173	0.763	-2.848
Persons $<$19 years	-0.898	0.599	-1.497	-0.036	0.565	-0.064
Workers	-2.126	0.697	-3.050	-0.584	0.730	-0.799
Unemployed	-1.916	0.715	-2.680	-1.757	0.736	-2.384
REGIME 2:						
Water price	-0.995	0.540	-1.842	-2.955	1.016	-2.908
Spline price	1.356	0.756	1.793	2.527	0.982	2.572
Electricity price	0.144	0.084	1.707	0.115	0.128	0.897
Income	-0.924	0.840	-1.100	-0.287	0.593	-0.484
Temperature	20.899	8.840	2.364	37.022	12.963	2.855
Density of population	-1.162	1.891	-0.614	-1.164	1.872	-0.622
Persons $<$ 19 years	-1.411	0.839	-1.680	-0.791	0.851	-0.929
Workers	-1.905	1.044	-1.824	-2.741	1.065	-2.572
Unemployed	-3.718	0.868	-4.280	-2.333	1.004	-2.322
R^2		0.58			0.51	
σ^2estimate		289.58			298.92	
Log likelihood		-490.18			-492.32	
Akaike		1018.37			1008.51	
Schwartz		1070.52			1057.92	
Diagnostics tests:						
Chow-Wald (structural stability)	$\chi^2_{(9)}$: 30.501 p-value: 0.000			$\chi^2_{(9)}$: 20.091 p-value: 0.017		

Although complementarity between the two goods (water and electricity) may be expected, the positive sign for the parameter of electricity average price variable indicates that, for the sample concerned, water and electriciy display substituability patterns. This means that an increase in electricity average price 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 electricity average price 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 electricity price, whereas the remainder (the share of residential water consumed without energy) does not. We have noticed in section 1 that about 40% of daily residential water consumption in France is concerned with heating. 60% would not be, which partly explains our result.

The spatial coefficient is also highly significant, which confirms the spatial modelling framework. Here, the spatial behavior may be viewed in two ways. First, we can argue that municipalities are actually influencing their neighbours. The water consumption behavior of other municipalities affects the consumption of a given municipality through social proximity. The estimated spatial parameter can be interpreted as a measure of externality. The significant spatial pattern may also be thought of as the reaction of households with respect to the availability of water resources. Our results indicate also a shift in role between regimes from one cross-section to another. This may be thought of partly as a consequence of the public sector water pricing policies through the "M-49 directive" (cf. Section 2).

5 Summary and conclusion

In this paper, we have taken advantage of data availability to specify a spatial regime switching approach for modelling residential water demand, including electricity price effects. Recall that in France, water is a public good. The comparison of two cross-sections allowed to evaluate the relevance and the structural instability induced by public water pricing.

So, what can we learn through this study? First, the estimated spatial lagged parameter is strongly significant, which means that households living in the same geographic area have approximately similar water consumption behaviors. The second point concerns the consumption tendencies between the two periods (1988.1 and 1993.1). Indeed, on a same sample, we have noticed a shift of role in regimes between 1988.1 and 1993.1. That is, in 1988.1, consumers in the first regime behaved in 1993.1 as if they were in the second regime. Thus, by charging all residential consumers in the same proportion, the "M49-directive" would induced structural instability. We can then argue that public water pricing as defined seems not to be efficient. Finally, some comments concerning price variables are in order. We find evidence that consumers respond jointly to water and electricity average price, not to water average price only. The use of electricity average price as additional regressor improved the model specification. Moreover, when constructing water average price, we have chosen a linear spline specification instead of a uniform slope. This raises the question of the proper price variable specification to be used.

A crucial assumption behind the maximum likelihood method used is that the disturbances are normally distributed. As pointed out by LeSage (1996), violation of this can arise from outliers or spatial enclave effects where a small cluster of observations display aberrant behavior. This aspect remains to be examined.

6 Appendix: variables used and related data source

The data were provided by "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), "l'Institut National de la Statistique et des Études Économiques" (National Institute of Statistics and Economic Studies) and "Electricité de Strasbourg" (Electricity of Strasbourg). The variables used in the estimation are the following.

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

Explanatory variables:

- Water average price (wap) in "FF" per m3; it is computed as follows: wap =

$\frac{(\text{cons}\times\text{mp})+\text{fc}}{\text{cons}}$, where (mp) denotes the marginal price, (fc) is the fixed charges and (cons) is water consumption,

- Electricity average price in "FF" per kwh,
- Disposable income per household paying taxes,
- Mean temperature in degree Celsius,
- Proportion of persons below 19 years,
- Proportion of workers,
- Proportion of unemployed,
- Density of population,
- Spatial lagged dependent variable.

Acknowledgements

Several results in this paper were originaly presented at the 34th International Conference of the American Water Resources Association (AWRA), November 1998, Point Clear, USA; under the title: "Spatial Patterns in Joint Residential Water and Electricity Demand: a French Case". Financial support of the "National School of Water and Environment Engineering of Strasbourg" (GSP, Engees-Cemagref) is gratefully acknowledged. I also wish to thank François Laisney for hepful comments and encouragement in developing this work. All errors remain my own.

References

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

——— (1990): "Spatial Dependence and Spatial Structural Instability in Applied Regression Analysis," *Journal of Regional Science*, (30,2), 185–207.

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

AZOMAHOU, T. (1999a): "Estimation of Spatial Panel Data Models Using a Minimum Distance Estimator: Application," *Working Paper BETA, no 9912, Louis Pasteur University, Strasbourg, Economics Department.*

——— (1999b): "Remarks on the Computation of Spatial Weights Matrices Using the Semivariogram Estimation," *forthcoming in Working Paper BETA, Louis Pasteur University, Strasbourg, Economics Department.*

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

CRESSIE, N. (1985): "Fitting Variogram Models by Weigthed Least-squares," *Mathematical Geology*, (17), 563–586.

CRESSIE, N. (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.

GUELLEC, A. (1995): "Le Prix de l'Eau de l'Exploitation à la Maîtrise," *Rapport d'Information de l'Assemblé Nationale par la Commission de la Production et des Echanges sur l'Eau.*

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

HÄRDLE, W. (1990): *Applied Nonparametric Regression.* Cambridge University Press.

KELEJIAN, H., AND I. PRUCHA (1997a): "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.

——— (1997b): "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model," *International Economic Review*, (Forthcoming).

LeSage, J. P. (1995): "A Multiprocess Mixture Model to Estimate Space-Time Dimensions of Weekly Pricing of Certificates Deposit," *Studies in Regional Sciences Series*, pp. 359–397.

——— (1996): "Spatial Modelling of Housing Values in Toledo," *Working Paper.*

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

Quandt, R. E. (1958): "The Estimation of the Parameters of a Linear Regression System Obeying Two Separate Regimes," *Journal of the American Statistical Association*, (53), 873–880.

Robinson, P. M. (1988): "Root-N-Consistent Semiparametric Regression," *Econometrica*, (56–4), 931–954.

Silverman, B. W. (1985): "Some Aspects of the Spline Smoothing Approach to Nonparametric Regression Curve," *Journal of the Royal Statistical Society*, (B, 47), 1–52.