PREAMBLE

The power of Thought — the magic of the Mind!

Byron, **Corsair**

Who was the innovator who first wrote of spatial econometrics? Probably hidden somewhere in the yellow-paged journals of yesteryear is a forgotten article, the first to break this barrier; if it exists, undoubtedly its discovery will occur at some future date. Certainly Paelinck was one of the first scholars to devote considerable thought to complications that lie dormant in traditional econometric analysis but become problematic when analyzing geo-referenced data. Over the past fifteen years he has repeatedly developed and/or modified econometric techniques in order to handle these complexities. As is characteristic of his earlier work, here he presents rationales, relevant properties, empirical examples, and possible extensions of estimators. In doing so, he highlights the difficulties of specification, interpretation, and computation. The purpose of this paper is to present the six new estimation techniques called simultaneous dynamic least squares, strictly positive conditional, linear logistic, least spheres, non-numerical regression, and distribution-free power. Acknowledging the innovativeness of Paelinck's work, Anselin emphasizes the focus of the six new estimators (i. e., problems of simultaneity in spatial modeling, data limitations, and complexities attributable to spatial interaction), as well as the technical issues of identifiability, distributional properties, and non-trivial implications associated with various approximations to non-linear estimators. All in all, this paper is as delightful an example of the spatial econometric viewpoint as can be found in the literature today.

The Editor

Some New Estimators in Spatial Econometrics

J. H. P. Paelinck

Department of Theoretical Spatial Economics, Erasmus University, P. O. B. 1738, 3000 DR Rotterdam, The Netherlands

Overview: The empirical study of spatial economic phenomena leads to a large number of specifically different problem settings. Reliable quantitative study of these problems often cannot proceed using standard econometric techniques, or approaches that initially were developed for other purposes. Moreover, better solutions can be obtained for these problems by modifying standard results in an appropriate way, or improving the properties of methods that already have been proposed. A number of new estimators are presented in this paper; it is believed that they will prove illuminating when applied to those spatial economic cases for which they have been developed. Without presenting an integrated body of econometric analysis—like k–class estimators, for instance—these new estimators represent a sample of spatial econometric estimation exercises that might usefully complement the body of knowledge already in existence.

1. Introduction

One important aspect of spatial econometrics is the development of estimators appropriate to given types of problems. In Paelinck and Klaassen (1979, Chapter 3) some special estimators, based on previous work, already have been presented. These include

- ⋆ estimators for a spatial income–generating model;
- ⋆ estimators for the interregional attraction model: MOLS (Multiregional Ordinary Least Squares), IOLS (Interregional Ordinary Least Squares), ISSML (Interregional Semi–Separable Maximum Likelihood); and,
- ⋆ estimation of threshold effects.

Estimators presented in Chapters 4 and 5 were original contributions at the time, and include

- ⋆ distribution–free testable spatial autocorrelation estimation;
- ⋆ component parameter estimation [for a recent application of this procedure, see Kuiper (1989)]; and,
- ⋆ fuzzy multiple regime estimation.

In this paper more recent materials that resulted from research undertaken since the publication of the aforementioned volume are presented. The organization of this presentation is as follows: rationale for the estimator, its presentation with relevant properties, an empirical example, and possible extensions. Additional aspects can be found in Ancot, Paelinck and Prins (1986), Ancot and Paelinck (1987) and Paelinck (1989).

2. Newcomers

2.1. Simultaneous Dynamic Least Squares (SDLS)

In Paelinck and Klaassen (1979, Chapter 7) it has been shown how SDLS estimation can simultaneously comply with synchronic, diachronic, sectoral and spatial interdependences. The model can indeed be written as

$$\mathbf{Ay} + \mathbf{Bx} = \boldsymbol{\xi}, \tag{2.1.1}$$

where matrix $\mathbf{A}$ represents the linkages between arbitrary endogenous variables $\mathbf{y}$ (spatialised or not, lagged or not), $\mathbf{Bx}$ the effects of exogenous shocks, and ξ stochastic elements. [1] The SDLS estimator is derived from the optimization problem of

$$\min_{\mathbf{A},\mathbf{B}} (\mathbf{y} - \mathbf{A}^{-1}\mathbf{Bx})'(\mathbf{y} - \mathbf{A}^{-1}\mathbf{Bx}) \tag{2.1.2}$$

and computed from the vector–matrix transformation of equation (2.1.1)

$$\mathbf{y} = \mathbf{X}\boldsymbol{\gamma} + \boldsymbol{\xi}, \tag{2.1.3}$$

with $\mathbf{y}$ and $\mathbf{X}$ respectively being a vector and a matrix of observed variables (one should note that vector $\mathbf{x}$ and matrix $\mathbf{X}$ are not the same), $\boldsymbol{\gamma}$ the vector of $\mathbf{A}$ and $\mathbf{B}$ coefficients, and (2.1.3) being in fact a so–called "normalised form" of equation (2.1.1). The estimator is given by

$$\boldsymbol{\gamma}^e = (\mathbf{X}'\mathbf{X}^e)^{-1}\mathbf{X}'\mathbf{y}, \tag{2.1.4}$$

where matrix $\mathbf{X}^e$ contains *estimated* endogenous variables. Numerical work conducted on this estimator has found that in practice convergence of a Gauss–Seidel nonlinear estimation procedure does occur (Prins, 1985).

Some properties of the equation (2.1.4) estimator are:

- ⋆ it is a generalized reduced form estimator;
- ⋆ if $\xi \sim \mathbf{N}(\mathbf{0}, \sigma^2\mathbf{I})$, then the estimator is a maximum likelihood (ML) one; and,
- ⋆ $\boldsymbol{\gamma}^e$ is consistent, with plim $\boldsymbol{\gamma}^e\boldsymbol{\gamma}^{e\prime} = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$.

Some early applications of this estimator can be found in Ancot, Kuiper and Paelinck (1981). It has been applied more recently to the estimation of discrete versions of the Lotka–Volterra model (Bagchus a.o., 1985; Budding a.o., 1985; Dickmann and Spoorendonk, 1987; for the model itself, one is referred to Peschel and Mende, 1986), the latter being then specified as

$$\Delta'\ln(x_t) = a + bx_{t-1} + cy_{t-1} \tag{2.1.5a}$$

$$\Delta'\ln(y_t) = d + ex_{t-1} + fy_{t-1} \tag{2.1.5b}$$

where the variable x represents population and y income per capita. Applied to the city of Rotterdam over the period 1946-1978, this estimator has given those results appearing in Table 1 (the computer program is discussed in Schueren, 1986). These tabulated values have acceptable interpretations. Loo (1987) also has studied other dutch cities, the principal problems encountered being that of the availability, the quality and the comparability over time of income figures. One could consider introducing distributed lags into this model, too.

TABLE 1
PARAMETERS OF A LOTKA–VOLTERRA MODEL
FOR THE CITY OF ROTTERDAM

Parameter	Estimated Value	Student's t–statistic
a	−0.8798	−7.68
b	0.0711	7.33
c	0.3988	4.22
d	1.0870	9.49
e	−0.8025	−8.51
f	−0.5355	−5.67
$a^{\star a}$	0.0362	1.64
$d^{\star a}$	0.0538	2.43

[a] Optimisation parameters for the starting point of an endogenous simulation.

Simultaneous dynamic least squares also can be useful in studying spatial autocorrelation. Consider the model

$$\mathbf{y} = \rho \mathbf{C}\mathbf{y} + \mu \mathbf{i} + \xi, \tag{2.1.6}$$

where $\mathbf{C}$ is a geographic contiguity matrix, and $\mathbf{i}$ is a vector of ones. The SDLS estimator is generated by minimising w.r.t. ρ and μ the expression

$$[\mathbf{y} - \mu(\mathbf{I} - \rho\mathbf{C})^{-1}\mathbf{i}]'[\mathbf{y} - \mu(\mathbf{I} - \rho\mathbf{C})^{-1}\mathbf{i}]. \tag{2.1.7}$$

Let us suppose that

$$|\rho\lambda(\mathbf{C})|_{\max} < 1, \tag{2.1.8}$$

so that one can consider an approximation supplied by only the linear terms of the spatial multiplier, $(\mathbf{I} + \rho\mathbf{C})$; equation (2.1.7) then can be rewritten as

$$[\mathbf{y} - \mu(\mathbf{I} + \rho\mathbf{C})\mathbf{i}]'[\mathbf{y} - \mu(\mathbf{I} + \rho\mathbf{C})\mathbf{i}], \text{ or} \tag{2.1.9}$$

$$[\mathbf{y} - \mu(\mathbf{i} + \rho\mathbf{n})]'[\mathbf{y} - \mu(\mathbf{i} + \rho\mathbf{n})], \tag{2.1.10}$$

where the vector $\mathbf{n}$ results from summing over the rows of matrix $\mathbf{C}$.

A first hypothesis can be that this summation gives a constant, ν such that

$$\mathbf{n} = \nu\mathbf{i}, \tag{2.1.11}$$

but differentiating equation (2.1.10) with respect to μ and ρ gives one and the same equation. This outcome is due to the non–discriminating effect of an infinite spatial structure.

If one defines

$$n = \mathbf{i}'\mathbf{n}, \text{ and} \tag{2.1.12a}$$

$$n^{\star} = \mathbf{n}'\mathbf{n}, \tag{2.1.12b}$$

then the two parameter estimates become

$$\rho = (\mathbf{y}'\mathbf{n} - \mu n)/(\mu n^{\star}), \text{ and} \tag{2.1.13a}$$

$$\mu = (\mathbf{i}'\mathbf{y} + \rho\mathbf{y}'\mathbf{n})/(r + 2\rho n + \rho^2 n^{\star}), \tag{2.1.13b}$$

where r is the number of spatial units. In equation (2.1.13), then, the non-spatial mean, $\mathbf{i}'\mathbf{y}/r$, is corrected for spatial autocorrelation.

From equations (2.1.13a) and (2.1.13b),

$$\rho = (n\mathbf{i}'\mathbf{y} - r\mathbf{y}'\mathbf{n})/(n\mathbf{y}'\mathbf{n} - n^{\star}\mathbf{i}'\mathbf{y}), \text{ and} \tag{2.1.14a}$$
$$\mu = (n\mathbf{y}'\mathbf{n} - n^{\star}\mathbf{i}'\mathbf{y})/(n^2 - n^{\star}r). \tag{2.1.14b}$$

These two expressions can be rewritten as

$$\rho = -(r/n)[(\mu^{s}/\mu^{\star}) - 1]/[(\mu^{s}/\mu^{\star}) - (n^{\star}/n)(n/r)], \text{ and} \tag{2.1.15a}$$
$$\mu = \mu^{\star}[(n/r)(\mu^{s}/\mu^{\star}) - n^{\star}/n)]/(n/r - n^{\star}/n), \tag{2.1.15b}$$

with

$$\mu^{\star} \triangleq \mathbf{i}'\mathbf{y}/r, \text{ and} \tag{2.1.16a}$$
$$\mu^{s} \triangleq \mathbf{n}'\mathbf{y}/n. \tag{2.1.16b}$$

From equation (2.1.15a) one finds that if $\mu^{s} = \mu^{\star}$, then ρ would be zero; the spatially corrected average does not add any information. Hence from equations (2.1.3b) or (2.1.15b) then, $\mu = \mu^{\star}$.

Study with respect to the critical values for ρ, namely 1 and -1, can proceed as follows. If $(n^{\star}/n)(n/r) - 1 = 1$, then $\rho = -r/n > -1$, so r/n is a damping factor. Positive autocorrelation is to be expected with "skew" spatial structures defined as

$$(n^{\star}/n)(n/r)^{-1} > \mu^{s}/\mu^{\star} > 1. \tag{2.1.17}$$

In the case of a constant number of first-order autoregressive links, ν, the expression for ρ becomes

$$\rho = (1/\nu)(\mu^{\star}/\mu - 1). \tag{2.1.18}$$

Supposing $\mu^{\star} > 0$ and $\mu > 0$, negative autocorrelation occurs for $\mu^{\star}/\mu < 1$, but will never be less than -1. Positive autocorrelation can only exceed $+1$ (*e. g.*, a non-stationary geographic process is operating) if $(\mu^{\star}/\mu - 1) > \nu$, which is a possibility for which the probability is unknown.

2.2. Strictly Positive Conditional Estimation (SPCE)

Ancot and Paelinck (1981) have drawn attention to the conceptual necessity of obtaining strictly positive values for certain parameters resulting from an a priori spatial theory. They have investigated the approach outlined here in the ensuing discussion. Let β be a parameter of the equation

$$\xi_i \triangleq y_i - \beta x_i, \tag{2.2.1}$$

the probability of observing jointly ξ_i and β being written as

$$p(\xi_i, \beta) = p(\xi_i|\beta)p(\beta), \tag{2.2.2}$$

where $p(\xi_i|\beta)$ is given by equation (2.2.1) and $p(\beta)$ is a prior density for parameter β. One estimates β under the hypothesis that over the observation period (or the observed regional

TABLE 2
FLEUR SECTOR NUMBER 28, PERIOD 1950–1960[a]

Variables	With Elimination of Parameters With the Wrong Sign		All Parameters		95% SPCE Bounds		
	β_{ols}	Student's t	β_{ols}	Student's t	β_{spce}	lower	upper
X_1	0.647	34.93	0.647	35.20	0.647	0.618	0.675
X_2	0.832	5.31	0.799	5.10	0.802	0.530	1.034
X_3	0.243	1.43	0.196	1.14	0.196	0.101	0.336
X_4	0.759	4.75	0.705	4.34	0.706	0.444	0.941
X_5	0.873	4.04	0.783	3.52	0.782	0.417	1.097
X_6	0.215	6.06	0.216	6.15	−0.100	−0.120	−0.084
X_7	0.388	11.42	0.403	11.50	0.403	0.360	0.452
X_8	0.434	20.52	0.424	19.41	0.424	0.396	0.452
X_9	*****	*****	−0.034	−1.52	0.100	0.091	0.110
X_{10}	0.119	1.76	0.110	1.64	0.110	0.083	0.144
X_{11}	0.063	0.81	0.688	0.89	0.079	0.049	0.095
X_{12}	−0.121	−1.16	−0.158	−1.49	−0.158	−0.293	−0.095
X_{13}	0.510	6.35	0.497	6.19	0.496	0.387	0.607
X_{14}	0.330	8.03	0.331	8.11	0.331	0.287	0.376
X_{15}	0.303	7.65	0.301	7.68	0.301	0.262	0.343
X_{16}	0.134	3.53	0.129	3.39	0.129	0.110	0.150
X_{17}	0.103	3.64	0.104	3.70	0.104	0.092	0.117

$R^2 = 0.974$ $R^2 = 0.975$

$\text{MSE}^b = 1.542$ $\text{MSE} = 1.493$ $\text{MSE} = 10.340$

"price" of SPCE = 6.927

[a] On the FLEUR model, see Ancot and Paelinck (1983).

[b] MSE is the residual variance.

system) β has been constant. This estimation has been investigated for $\xi \sim N(\mathbf{0}, \sigma^2\mathbf{I})$ and $\boldsymbol{\beta} \sim T(\boldsymbol{\beta}^\star)$, where T represents a Tanner distribution having estimates $\boldsymbol{\beta}^\star$. The estimated value, $\boldsymbol{\beta}^e$, is

$$\boldsymbol{\beta}^e = \tilde{\boldsymbol{\beta}}(\hat{\boldsymbol{\beta}}, \boldsymbol{\beta}^\star) + 2n\sigma^2(\mathbf{X}'\mathbf{X})(\hat{\boldsymbol{\beta}}^e)^{-1}\mathbf{i}, \tag{2.2.3}$$

where $\hat{\boldsymbol{\beta}}$ is the ordinary least squares (OLS) estimator, and n is the number of observations. One should note that

- ⋆ equation (2.2.3) has indeed a strictly positive (or strictly negative, if required) value in $\boldsymbol{\beta}^e$; and,
- ⋆ up to second-order σ^2, VAR($\boldsymbol{\beta}^e$) equals the OLS expression.

Table 2 summarizes an individual result borrowed from Enhus (1986) based upon this estimator.

TABLE 3
SYNTHESIS OF RESULTS OBTAINED BY ENHUS, 1986

FLEUR Sector	Period	Number of Replaced Coefficients	"Price"
19	1960–1970	2	6.337
28	1950–1960	2	6.927
37	1950–1960	2	7.559
19	1950–1960	2	8.912
53	1950–1960	2	14.643
28	1960–1970	1	17.168
37	1960–1970	2	54.603
7	1950–1960	3	229.399
7	1960–1970	3	566.485

Following Ancot and Paelinck (1981, p. 360, Property 3) the confidence intervals for the SPCEs have been assumed to be log–normal. The interpretation of the tabular results reported in Table 2 is obvious; the "price" to be paid for SPCE is the ratio of the residual variance SPCE/OLS (all parameters).

Table 3 presents an overview of those results obtained by Enhus (1986).

2.3. Linear Logistic Estimation (LLE)

In spatial analysis the presence of binary $0-1$ indicator variables that are to be predicted or statistically explained (presence or absence of certain elements) is frequent. Suppose the probability p_{ijk} for a firm of type i (I characteristics of a "plant profile") of exporting product j (J characteristics of a "product profile") to country k (K characteristics of an "export profile") to be logistic. The three profiles are represented by a vector $\mathbf{x}$, with "more" of a characteristic increasing the probability of exporting according to the function

$$p_{ijk} = [1 + \exp(-\mathbf{a}'\mathbf{x})]^{-1}, \tag{2.3.1}$$

with

$$\mathbf{a} \geq \mathbf{0}, \tag{2.3.2}$$

and with the observations being 0 (no exports) or 1 (exports). Let two variables be defined, one for exporters as

$$d_{1i} \triangleq 1 - [1 + \exp(-\mathbf{a}'\mathbf{x}_i)]^{-1}, \tag{2.3.3a}$$

and one for non–exporters as

$$d_{2i} \triangleq [1 + \exp(-\mathbf{a}'\mathbf{x}_i)]^{-1}, \tag{2.3.3b}$$

Thus one can easily compute

$$\mathbf{a}'\mathbf{x}_{1i} = \ln(d_{1i}^{-1} - 1) \triangleq \delta_{1i} > 0, \text{ and} \tag{2.3.4a}$$

$$-\mathbf{a}'\mathbf{x}_{2i} = \ln(d_{2i}^{-1} - 1) \triangleq \delta_{2i} > 0. \tag{2.3.4b}$$

Maximising $\sum_{i=1}^{n}(\delta_{1i}+\delta_{2i})$, under a norm restriction, one obtains

$$\mathbf{a}^e = \lambda^{-1}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{i}, \text{ and} \tag{2.3.5a}$$
$$\mathrm{VAR}(\mathbf{a}^e) = \lambda^{-2}(\mathbf{X}'\mathbf{X})^{-1}, \tag{2.3.5b}$$

which means that the ratio $a_k^e/\sigma(a_k^e)$ is independent of λ, so that the null hypothesis $\mathbf{a}=\mathbf{0}$ can be tested.

The numerical example presented in Table 4 has been explored here. Using those data, Table 5 compares the results from a classical "probit" analysis with that of LLE.

TABLE 4
DATA FOR L. L. E.

	Variables		
Values of i	Y	X_1	X_2
1	1.0	6.0	4.0
2	1.0	8.0	2.0
3	1.0	4.0	3.0
4	1.0	7.0	0.0
5	1.0	9.0	1.0
6	0.0	2.0	6.0
7	0.0	3.0	9.0
8	0.0	1.0	4.0
9	0.0	0.0	8.0
10	0.0	1.0	7.0

TABLE 5
PROBIT AND L. L. E.

Parameter	Probit	Student's t	L. L. E.	Student's t
of X_1	3.40	$0.20{\star}10^{-02}$	0.20	1.18
of X_2	−1.72	$-0.90{\star}10^{-03}$	−0.13	−0.70
Constant	−2.54	$-0.23{\star}10^{-03}$	−0.26	−0.18

One can see from these tabulated results that the signs as well as (for the parameters of $\mathbf{x}_1$ and $\mathbf{x}_2$) the ratios are consistent. The t statistics for the LLE estimators, however, are much less non–significant than are those for the probit estimators, the data obviously being ill–conditioned.

Recently this estimator has been extended to $0-z_i$ cases, where the z_i are possibly all different real numbers. For the latter observations the vector $\mathbf{i}$ in equation (2.3.5a) is extended by $\lambda\xi$, with vector

$$\boldsymbol{\xi} \triangleq [\ln(2-z_i k_i^{-1}) - \ln\{(z_i k_i - 1)\exp(\mathbf{a}'\mathbf{x}_i)+1\}], \tag{2.3.6}$$

which reduces to the binary $0-1$ case for a choice of $\lambda \to 0$ (very small "distances" required) or $z_i k_i^{-1} \to 1$, $\forall i$ (a perfect fit, which is the equivalent result). The k_is are variable asymptotes; for a specification of the form

$$k_i = \exp(\mathbf{b}'\mathbf{y}_i), \tag{2.3.7}$$

straightforward OLS (with an extra parameter for the $z_i = 0$ observations) allows the estimating of $\mathbf{b}$.

2.4. Least Spheres Estimation (LSE)

In some cases of spatial analysis, the presence of potentials (sums) can lead to multicollinearity; in such a case, another estimator, LSE, can bring relief, its objective function being

$$\Psi = [\sum_{i=1}^{n}\sum_{j=1}^{k}(a_j x_{ij} - y_i^\star)^2 + \sum_{i=1}^{n}\sum_{j=1}^{k}(x_{ij} - x_{ij}^\star)^2]/2, \tag{2.4.1}$$

where the a_j and x_{ij} are the endogenous variables to be estimated, and the starred variables $y_i^\star$ and $x_{ij}^\star$ are being observed (the number of degrees of freedom will remain $n-k$ in this case). More specifically, one minimises the sum of squares of the radii of the hyperspheres with centres $(y_i^\star, x_{ij}^\star)$, $\forall i$, that are tangent to the hyperplane $y_i = \sum_{j=1}^{k} a_j x_{ij}$, $\forall i$.

The estimators for $\mathbf{a} \triangleq [a_j]$ are

$$\tilde{\mathbf{a}} = (\mathbf{X}^{\star\prime}\mathbf{X}^\star - \varepsilon^{-2}I)X^{\star\prime}y^\star,$$

the stars indicating exogenous variables, which is a curious rejoinder to ridge regression. As the a_js are not inversely invariant with the measurement units of the x_{ij}s, ε should be maximised, and to guarantee positive definiteness, its sign reversed (the mathematical justification for this approach appears in Paelinck and Klaassen, 1979, pp. 54-55). Table 6 reproduces the results for this estimator applied to a tourist model of Swiss data for the "canton du Valais" (Bailly and Paelinck, 1988). Significance tests for the estimated parameters are available, and their results are reported in Table 6.

2.5. Non–numerical Regression (QUALIREG)

Suppose one wants to explain a phenomenon on which only qualitative observations are available (for example a vector of ranked items $\mathbf{y}' = [+++, -, 0, ++, \ldots]$), with the same situation prevailing for the matrix of explanatory variables, $\mathbf{X}$:

$$\mathbf{X} = \begin{pmatrix} + & - & \cdot & \cdot & \cdot \\ ++ & 0 & \cdot & \cdot & \cdot \\ - & ++ & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix} \tag{2.5.1}$$

Suppose matrix $\mathbf{X}$ is of order n–by–k (n observations on k explanatory variables). Such a situation is frequently encountered in spatial econometric analysis. The following programme gives a solution for the problem: find a vector of coefficients $\boldsymbol{\beta}^e$ maximising

$$\tau(\mathbf{y}, \mathbf{y}^e), \tag{2.5.2}$$

where τ is Kendall's rank correlation coefficient [2] and $\mathbf{y}^e$ the vector of estimated ranks of $\mathbf{y}$. A normalisation of $\boldsymbol{\beta}^e$ is necessary, which leads to the mathematical programme

$$\max_{\boldsymbol{\beta}^e} = \max_{\boldsymbol{\beta}^e} \boldsymbol{\beta}^{e\prime}\boldsymbol{\tau}^e \tag{2.5.3}$$

TABLE 6
PARAMETERS OF A TOURIST MODEL
APPLIED TO A "CANTON" IN SWITZERLAND

Parameters	Types of Villages			
	Valley Locations		Mountain Locations	
	French Speaking	German Speaking	French Speaking	German Speaking
Self–inducation/breaking	−0.074*	−0.760	−0.722	−0.236*
Locations Potential:				
French Speaking Valley	0.003*	0.109*	0.061*	0.078*
German Speaking Valley	−0.069*	0.111*	0.006*	−0.002*
French Speaking Mountain	−0.056*	0.111*	0.016*	0.013*
German Speaking Mountain	−0.010*	0.112*	−0.019*	−0.038*
Optimised starting point of endogenous dynamics	0.377*	−0.012*	0.290	0.421*
Autonomous growth/ decline rate	0.004*	0.109*	0.062*	0.080*
Pseudo–R^2	0.241*	0.563	0.460	0.416*

NOTE: * denotes significance at the 95% confidence level using chi square and F test statistics.

subject to:

$$-\mathbf{i} \leq \boldsymbol{\beta}^e \leq \mathbf{i} \tag{2.5.4}$$

where $\boldsymbol{\tau}^e$ is the vector of Kendall τ's corresponding to the permutations of columns of $\mathbf{X}$ producing $\mathbf{y}^e$. "Multiple correlation" and $\boldsymbol{\beta}^e$ tests are available, as Table 7 shows.

This method has been applied to an explanatory relation of water discharge per province in the Netherlands (see Davelaar a. o., 1983). Table 7 gives the results of a comparative exercise with OLS–estimation; more elaborate commentary on these results appears in Ancot and Paelinck, forthcoming.

2.6. Distribution–Free Power Estimator (DFPE)

Relations in spatial econometrics are often of a highly non–linear nature (see Paelinck and Klaassen, 1979, pp. 6-9). Power parameter specifications can be useful to model such behaviour; an early application of this perspective to a so-called "multiple gap" investment model, using other solution methods, is reported on in Ancot e. a. (1978).

Generalised Box–Cox transformations (see Box and Cox, 1964) will be discussed next, together with a proposed procedure for nonparametric estimation. This latter procedure—for generalised Box–Cox transformations—can proceed in the following manner. Suppose a non–linear relation exists and may be specified as

$$y_i^{\sigma} = \sum_{j=1}^{k} a_j x_{ij}^{\rho_j} + \xi_i \tag{2.6.1}$$

TABLE 7
COMPARISON OF THE RESULTS OBTAINED BY QUALIREG AND BY OLS

	Coefficients						
	b_1	Kendall's tau	b_2	Kendall's tau	b_3	Kendall's tau	Multiple Correlation
QUALIREG							
	0.5	0.24*	−0.4	0.31	−0.3	0.16	0.636[a]
	1.0	0.24	−0.8	0.31	−0.56	0.18	
	0.8	0.24	−0.6	0.27	−0.4	0.27	0.600
	1.0	0.20	−0.6	0.35	−0.6	0.09	

	Coefficients								
	b_0	S–t	b_1	S–t	b_2	S–t	b_3	S–t	Multiple Determination
OLS									
	3535.02	1.65	−1.50	−1.26	0.17	3.02	−103.35	−1.33	0.875

"S–t" is short for "Student's t"

[a] Significant at the 5% level, in accordance with Kendall's tau, implies a value of 0.385 or more.

Note: * denotes significance at the 5% level, using Kendall's tau (critical region is −0.385 or less).

one of the x_{ik}s being equal to unity if necessary (the regression constant). The typical "normal" equations for this situation are as follows (from minimising $\sum_{i=1}^{n} \xi_i{}^2 \triangleq \Psi$) : [3]

$$\frac{\partial \Psi}{\partial \sigma} = \sum_{i=1}^{n} [y_i^{\sigma} - \sum_{j=1}^{k} a_j x_{ij}^{\rho_j}] y_i^{\sigma} \ln(y_i) = 0 \tag{2.6.2a}$$

$$\frac{\partial \Psi}{\partial \rho_j} = \sum_{i=1}^{n} [y_i^{\sigma} - \sum_{j=1}^{k} a_j x_{ij}^{\rho_j}] x_{ij}^{\sigma_j} \ln(x_{ij}) = 0 \tag{2.6.2b}$$

$$\frac{\partial \Psi}{\partial a_j} = \sum_{i=1}^{n} [y_i^{\sigma} - \sum_{j=1}^{k} a_j x_{ij}^{\rho_j}] x_{ij} = 0 \tag{2.6.2c}$$

Equations (2.6.2a) and (2.6.2b) can be expressed in vector–matrix form as

$$\mathbf{y}_1(\boldsymbol{\rho}) = \mathbf{X}_1(\boldsymbol{\rho})\mathbf{a}, \tag{2.6.3a}$$

and similarly for equations (2.6.2c),

$$\mathbf{y}_2(\boldsymbol{\rho}) = \mathbf{X}_2(\boldsymbol{\rho})\mathbf{a}, \tag{2.6.3b}$$

where equation (2.6.3b) is a generalisation of the OLS normal equations.

This finding means that vector **a** can be eliminated by combining equations (2.6.3a) and (2.6.3b), yielding

$$\mathbf{y}_1(\boldsymbol{\rho}) = \mathbf{X}_1(\boldsymbol{\rho})\mathbf{X}_2^{-1}(\boldsymbol{\rho})\mathbf{y}_2(\boldsymbol{\rho}), \tag{2.6.4}$$

where the inverse term $\mathbf{X}_2^{-1}$ exists except in the presence of perfect multicollinearity or for $\boldsymbol{\rho} \equiv 0$. At least in principle, systems like equation (2.6.4) could be solved by Gauss–Seidel methods (see Hughes Hallett, 1984).

The specifications explored here are the following identical iterative equations, with l denoting the iteration step, derived from equation (2.6.4):

$$\boldsymbol{\rho}_l = [\hat{\boldsymbol{\rho}}\hat{\mathbf{y}}_1^{-1}(\boldsymbol{\rho})X_1(\boldsymbol{\rho})X_2^{-1}(\boldsymbol{\rho})y_2(\boldsymbol{\rho})]_{l-1}, \tag{2.6.5a}$$

a multiplicative identity, and

$$\boldsymbol{\rho}_l = \boldsymbol{\rho}_{l-1} + [\mathbf{X}_1(\boldsymbol{\rho})X_2{-1}(\boldsymbol{\rho})y_2(\boldsymbol{\rho}) - \mathbf{y}_1(\boldsymbol{\rho})]_{l-1}, \tag{2.6.5b}$$

an additive identity.

The first estimation results obtained were disappointing.[4] Either different answers were converged upon, or no convergence occurred at all. A next step was to proceed directly with equations (2.6.2a) and (2.6.2b) in their identical additive form, the advantage being that they are mono–parametrical. This strategy rendered poor results, too. Hence the differential form of the latter equations, *i. e.*

$$\sum_{i=1}^{n} e_i y_i^{\sigma}[\sigma \ln(y_i) + 1]dy_i = 0, \tag{2.6.2a*}$$

and

$$\sum_{i=1}^{n} e_i x_{ij}^{e_j}[\rho_j \ln(x_{ij}) + 1]dx_{ij} = 0, \tag{2.6.2b*}$$

has not been investigated.

Inspection of the objective function Ψ shows that, given $\forall i, j$, y_i and $x_{ij} > 1$, this function tends to zero for σ and $\rho_j \rightarrow -\infty$. At best a local minimum can be found within given bounds. Perhaps the unit hypercube $[-\mathbf{i}, +\mathbf{i}]$ should be used as these bounds, given that it covers the interesting extreme cases, to wit linear regression $(\sigma, \rho_j = 1)$, double logarithmic regression $(\sigma, \rho_j \rightarrow 0)$, and inverse linear regression $(\sigma, \rho_j = -1)$. Then a global local optimum could be searched for by appropriate methods, *e. g.* simulated annealing (Laarhoven, 1988); other methods presently are being investigated.

3. Conclusions

Econometric estimation remains a difficult exercise. The first difficulty resides in *specifying* an approach that may lead to operational results, starting with the stated aim of an exercise. Common aims include better simulation, more precise estimations, the obtaining of parameter estimates with specific properties, exploitation of poor quality data, or handling non–linear or interdependent relationships. Econometric wisdom is useful, for attaining some of these objectives, but should be complemented by ingenious inventiveness to generate the mathematical entities that correspond to given requirements.

A second point of difficulty worth mentioning is that parameters to be estimated are not to be assumed as being independently given, from outside of the problem, but rather have to be viewed *relative to the use* that is to be made of their estimated values. Typical uses include prediction (again in a relative sense, and in terms of a specific purpose), simulation (again as specified with regard to its aims, such as analyzing system stability and sensitivity, or exploring consequences of either policy measures or exogenous shocks), hypothesis testing (here sometimes absolute parameter values are irrelevant, as can be seen in the foregoing discussion of QUALIREG and LLE).

A third, and final, difficulty acknowledged here concerns the *computation* of parameter values. A desirable specification renders simple, robust estimation procedures that produce achievable results. This extremely useful property should be welcomed, as well as sought in a chosen procedure.

Consequently, the spatial econometric findings reported in this paper, especially when couched in terms of the three difficulties outlined in this conclusion, illustrate the enormous possibilities that still remain for contributing to this subfield. The most beneficial contributions will combine spatial economic theory, general econometric wisdom, and numerical analysis, with each of these three fields continuing to hold immense potential for research and discovery.

4. References

Ancot, J–P., Iwema, R. et Paelinck, J. (1980) Test d'une hypothèse d'investissement à écarts multiples, *L'Actualité Economique*, **1**, 40-59.

Ancot, J–P., Kuiper, J. et Paelinck, J. (1981) Réflexions sur la simulation de modèles dynamiques, *Working Paper* No. 1981/2, Foundations of Empirical Economic Research Series, Netherlands Economic Institute, Rotterdam.

Ancot, J–P. and Paelinck, J. (1981) Recent research in spatial econometrics, in D. Griffith and R. MacKinnon (eds.), *Dynamic Spatial Models*, Sythoff & Noordhoff, Alphen a/d Rijn, pp. 344-364.

Ancot, J–P. and Paelinck, J. (1983) The spatial econometrics of the European FLEUR–model, in D. Griffith and A. Lea (eds.), *Evolving Geographical Structures*, Martinus Nijhoff Publishers, The Hague, pp. 229-246.

Ancot, J–P., Paelinck, J. and Prins, J. (1986) Some new estimators in spatial econometrics, *Economics Letters*, **21**, 245-250.

Ancot, J–P. et Paelinck, J. (1987) Nouveaux estimateurs en économétrie spatiale, *Working Paper* No. 1987/6, Foundations of Empirical Economic Research Series, Netherlands Economic Institute, Rotterdam.

Ancot, J–P. and Paelinck, J. Non–parametric estimation: the QUALIREG–model, its estimation and testing with spatial cross–sectional data, *Cahiers du GAMA*, forthcoming.

Bagchus, R., Nes, P. van en Zaan, A. van der. (1985) Stedelijke dynamiek, een schatting van het prooi–roofdier model voor Rotterdam (Urban dynamics, an estimation of the prey–predator model for the city of Rotterdam), *Working Paper*, Faculty of Economics, Department of Theoretical Spatial Economics, Erasmus University, Rotterdam.

Bailly, A. et Paelinck, J. (1988) Un modèle économétrique du développement socio–spatial de régions touristiques, *Revue d'Economie Régionale et Urbaine*, **3**, 397-412.

Box, G. and Cox, D. (1964) An analysis of transformations, *Journal of the Royal Statistical Society*, **16B**, 211-243.

Budding, D., Cassa, H. en Hoek, P. van den. (1985) Het "prooi en roofdier"–model toegepast op de stedelijke bevolking (The prey–predator model applied to urban population), *Working Paper*, Faculty of Economics, Department of Theoretical Spatial Economics, Erasmus University, Rotterdam.

Davelaar, E., Jong, G. de en Koele, H. (1983) De methode QUALIREG (The QUALIREG method), *Working Paper*, Faculty of Economics, Department of Theoretical Spatial Economics, Erasmus University, Rotterdam.

Dickmann, A. en Spoorendonk, P. (1987) Een analyse van de stedelijke dynamiek van de gemeente Rotterdam (1946-1978) (An analysis of the urban dynamics of the city of Rotterdam), *Working Paper*, Faculty of Economics, Department of Theoretical Spatial Economics, Erasmus University, Rotterdam.

Enhus, J. (1986) Schatting van het FLEUR–model m.b.v. een Strictly Positive Conditional Estimator (Estimation of the FLEUR model by means of SPCE), unpublished Master's Thesis, Faculty of Economics, Department of Theoretical Spatial Economics, Erasmus University, Rotterdam.

Hughes Hallett, A. (1984) Simple and optimal extrapolations for first order iterations, *Journal of Computer Mathematics*, **15**, 309-318.

Kendall, M. (1962) *Rank Correlation Methods*, 3rd ed., Griffin, London.

Kuiper, J. (1989) Regional Analysis Using the Concept of Location Elasticities, *Revue d'Economie Régionale et Urbaine*, **3**, 363-391.

Laarhoven, P. van. (1988) *Theoretical and Computational Aspects of Simulated Annealing*, unpublished doctoral dissertation, Erasmus University, Rotterdam.

Loo, J. van. (1987) Schatten van een veralgemeend Lotka–Volterra model voor de stedelijke ontwikkeling in Nederland (Estimation of a generalised Lotka–Volterra model for the urban evolution in the Netherlands), unpublished Master's Thesis, Faculty of Economics, Department of Theoretical Spatial Economics, Erasmus University, Rotterdam.

Paelinck, J. and Klaassen, L. (1979) *Spatial Econometrics*, Saxon House, Farnborough; polish translation: *Ekonometria Przestrzenna*, Panstwowe Wydanictwo Naukowe, Warsawa, 1983.

Paelinck, J. (1989) Econométrie spatiale: contributions récentes après 20 ans d'histoire, *Cahiers Vilfredo Pareto*, forthcoming.

Peschel, M. and Mende, W. (1986) *The Predator–Prey Model*, Springer–Verlag, New York.

Prins, J. (1985) SDLS–schatting van een stedelijk model (SDLS–estimation of an urban model), unpublished Master's Thesis, Faculty of Economics, Department of Theoretical Spatial Economics, Erasmus University, Rotterdam.

Schueren, M. de van der (1986) Eindverslag SDLS schatting, stadsmodel en Lotka–Volterra (Final report SDLS estimation, urban model and Lotka–Volterra), *Working Paper*, Faculty of Economics, Department of Theoretical Spatial Economics, Erasmus University, Rotterdam.

NOTES

1. Column–vectors are systematically represented by lower case bold letters, matrices by upper case bold letters except for diagonal matrices identified by a cap, and $'$ denotes matrix transposition.

2. $\tau = 1 - 4Q/[n(n-1)]$, where Q is the number of so–called elementary permutations of $\mathbf{y}$ (see Kendall, 1962).

3. For small ξ_i, second order conditions for a minimum in σ and ρ_k will be satisfied for given a_js.

4. Computer computations by Niek Mares are gratefully acknowledged.

5. Computations have shown the existence of various singular points satisfying equations (2.6.2); also see Footnote 3.

DISCUSSION

"Some new estimators in spatial econometrics"

by J. H. P. Paelinck

In his chapter, Professor Paelinck presents six econometric estimators geared to problems encountered in spatial economics. They are the following: simultaneous dynamic least squares (SDLS), strictly positive conditional estimation (SPCE), linear logistic estimation (LLE), least spheres estimation (LSE), non–numerical regression (Qualireg), and distribution–free power estimation (DFPE). The latter is new, while the others are extensions and generalizations of earlier work, most notably in Ancot, Paelinck and Prins (1986), and Paelinck (1987).

Before I formulate some technical comments on the approaches suggested by Professor Paelinck, I would like to outline a more general organization of the six estimators based on three distinct perspectives or emphases, similar to the taxonomies suggested in my own chapter and in Anselin (1988a).

A first overall category is that of the particular perspective or paradigm represented in each method. Three distinct approaches are reflected in the six estimators. The classical perspective is taken for SDLS, LSE and LLE, which are examples of a maximum likelihood or pseudo–(quasi–)maximum likelihood (*e. g.*, Gourieroux *et al.*, 1984; White, 1984). The properties of these techniques are based upon asymptotic results, and hence they are alternative ways to solve estimation issues as problems of optimization (or best fit). A second approach is that reflected in the SPCE technique, where use is made of extraneous information to restrict the estimation, similar to Bayesian or Stein–like procedures (*e. g.*, Judge and Yancey, 1986). The crucial issue is how to derive the proper mix of data and prior information, which naturally leads to a concept of price associated with the imposed constraints. The third perspective is that of the nonparametric (robust) Qualireg and DFPE techniques, where limiting or otherwise unrealistic distributional assumptions are avoided.

A second important category of estimators involves the way in which specification issues particular to *spatial modeling* are taken into account. Foremost among these are the issues of spatial and space–time dependence as well as the various non–linearities associated with spatial interaction models (*e. g.*, distance decay functions, potentials). These topics are specifically addressed by SDLS, LSE and DFPE. A final category addresses the data limitations encountered in spatial analysis, namely problems of measurement (Qualireg) and positivity (SPCE).

In sum, the various new estimators suggested by Professor Paelinck focus on problems of simultaneity in spatial modeling, on data limitations encountered in spatial analysis, and on special complexities associated with spatial interaction. They compare to other recently advocated new directions, such as various shrinkage estimators (and the treatment of outliers), spatial adaptive filtering, and a spatial bootstrap estimator (for a review, see Anselin, 1988a).

From a technical standpoint, there are a number of issues raised by these new methods that merit closer scrutiny. First, the simultaneity expressed in most of the formulations (but especially for SDLS) raises questions of identifiability. In particular, the spatial structure inherent in the system under consideration somehow needs to be expressed in formal terms.

Whereas the SDLS approach is the most flexible one in this respect, as it avoids the familiar problem of assuming a weights matrix, it is conditional upon the availability of sufficient observations over space or space–time to allow for the identification of the structure of spatial interaction. In this respect, the re–introduction of a spatial weights matrix in the application of this technique to problems of spatial autocorrelation seems to be a step backward. Also, all six methods, to the extent that they deal with spatial dependence, presume spatial homogeneity, whereas spatial heterogeneity has been shown to be just as important a problem in empirical regional science (Anselin and Griffith, 1988).

The distributional properties of the various estimators are unclear. Based on asymptotic considerations, one may reasonably expect normality in most cases, but this is not necessarily reflected in the finite samples encountered in practice. The general issues involved in the trade–off between asymptotic normality and finite sample robustness are well–known, but the exact costs associated with each approach in practical empirical situations are less well understood. In addition, there is no unambiguous standard by which to compare the performance of the various new estimators in actual applications. Since most of the approaches are non–linear and do not necessarily result in residuals with a zero mean, the interpretation of the standard R^2 is not clear (Anselin, 1988b). The larger question is how to adequately summarize spatially differentiated or spatially dependent indicators of model accuracy. Unless this issue is addressed, there is really no standard by which to judge the superiority of these "new" approaches. Similarly, as yet there is no satisfactory way to assess the informational content of methods in which quantitative and qualitative measures are combined (as in Qualireg). Although the higher degree of realism expressed in the qualitative judgments in the data matrix is attractive, the degree of precision associated with the quantitative estimates remains unclear.

Finally, it may be interesting to compare some results for the spatial autoregressive model (2.1.6) between the SDLS approach and the more traditional regression approach. Using the notation of Paelinck, such a specification would be as follows:

$$\mathbf{y} = \rho \mathbf{C}\mathbf{y} + \mu \mathbf{i} + \xi$$

where $\mathbf{C}$ is the spatial weights matrix. With the simplifying condition (2.1.11), expression (2.1.18) finds the relation between ρ and μ as:

$$\rho = (1/\nu)(\mu^*/\mu - 1).$$

As shown in Anselin (1988a), a conditional least squares estimate for μ is:

$$\mu = (\mathbf{i}'\mathbf{i})^{-1}\mathbf{i}'(\mathbf{I} - \rho\mathbf{C})\mathbf{y}$$

or, with $\mathbf{i}'\mathbf{i} = r$, and a symmetric matrix $\mathbf{C}$, $(\mathbf{C}\mathbf{i})' = \mathbf{i}'\mathbf{C} = (\nu\mathbf{i})'$,

$$\mu = (1/r)(\mathbf{i}'\mathbf{y} - \rho\mathbf{i}'\mathbf{C}\mathbf{y}),$$
$$\mu = (1 - \rho\nu)(\mathbf{i}'\mathbf{y}/r)$$

or, in Paelinck's notation, with $\mathbf{i}'\mathbf{y}/r = \mu^*$:

$$\mu = (1 - \rho\nu)\mu^* \quad \text{and}$$
$$\rho = (1/\nu)(1 - \mu/\mu^*).$$

This result is not the same as expression (2.1.18), except for the uninteresting case where $\mu = \mu^*$ and thus $\rho = 0$. This simple derivation illustrates that the various approximations implied by the non-linear estimators are not trivial, and may have significant consequences for the ultimate results. The exact nature of these consequences needs to be investigated in further detail. The innovative methods suggested by Professor Paelinck provide a challenge to other researchers in spatial statistics to pursue this more extensively, both from a formal as well as from an empirical viewpoint.

References

Ancot, J-P., J. Paelinck, and J. Prins. (1986) Some new estimators in spatial econometrics. *Economics Letters*, **21**, 245-249.

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

Anselin, L. (1988b) Model validation in spatial econometrics: a review and evaluation of alternative approaches. *International Regional Science Review*, **11**, 279-316.

Anselin, L., and D. Griffith. (1988) Do spatial effects really matter in regression analysis? *Papers*, Regional Science Association, **65**, 11-34.

Gourieroux, C., A. Monfort, and A. Trognon. (1984) Pseudo maximum likelihood methods: theory. *Econometrica*, **52**, 681-700.

Judge, G., and T. Yancey. (1986) *Improved Methods of Inference in Econometrics.* Amsterdam: North Holland.

Paelinck, J. (1987) Empirical results with some new estimators in spatial econometrics, paper presented at the 27th Meeting of the Western Regional Science Association, Napa Valley, February.

White, H. (1984) *Asymptotic Theory for Econometricians.* New York: Academic Press.

Luc Anselin, University of California/Santa Barbara