Appeared in: Economics Letters, 2006, 91, 138-145.

ARTICLE IN PRESS

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Economics Letters xx (2006) xxx–xxx

economics letters

www.elsevier.com/locate/econbase

Functional form and spatial dependence in dynamic panels

Zhenlin Yang *, Chenwei Li, Y.K. Tse

School of Economics and Social Sciences, Singapore Management University, Singapore 178903, Singapore

Received 19 April 2005; received in revised form 11 October 2005; accepted 11 November 2005

Abstract

This paper proposes a generalized dynamic error component model that simultaneously accounts for the effects of functional form and spatial dependence. Maximum likelihood method is used for model estimation and inference. An empirical illustration using the demand for cigarettes data is given.

Keywords: Dynamics; Functional form; Spatial correlation; Maximum likelihood estimation

JEL classification: C23; C51; C53

1. Introduction

Functional form, spatial dependence and dynamics are three important issues in modelling economic panel data. While the spatial panel model and dynamic panel model have both become popular in the econometrics literature, most of econometric analysis and empirical studies using panel model ignore one or more of these issues, in particular the issue of functional form.[1] When a priori information and theoretical foundation are lacking, it is generally advisable to start an empirical investigation with a general model so that certain effects can be formally tested.

* Corresponding author. Tel.: +65 6828 0852; fax: +65 6828 0833.
E-mail address: zlyang@smu.edu.sg (Z. Yang).

[1] See Elhorst (2003) for a review on spatial panel models, and Hsiao (2003, Ch. 4) and Baltagi (2001, Ch. 8) for recent developments on dynamic panel models.

0165-1765/$ - see front matter
doi:10.1016/j.econlet.2005.11.011

In this paper, we propose a generalized error component model that simultaneously takes into account the issues of functional form, spatial dependence and dynamic effects. We adopt the maximum likelihood estimation (MLE) technique for model estimation and inference. The reason for using MLE method is that the standard generalized least squares (GLS) method and instrumental variable (IV) method (or GMM in general) are not applicable to our model as the response involves an unknown parametric transformation (Davidson and MacKinnon, 1993, p. 243). While it is widely viewed that MLE method suffers from the problems of computational complexity and lack of robustness against distributional misspecification, we show that the amount of computation involved in the MLE method is feasible for a desktop computer for data sets of moderate sizes. Furthermore, introducing a parametric transformation on the response variable has provided protection, at least to a certain degree, against some forms of distributional misspecification since one relies on data to choose a transformation to make the transformed data conformable with normality.

Section 2 introduces the model and presents the maximum likelihood procedures for model estimation and inference. Section 3 presents an empirical application using the popular cigarettes demand data. Section 4 concludes the paper.

2. The model and model estimation

We propose a generalized error component model with functional form transformation, spatial error correlation and dynamic effect. The model takes the following general form,

$$h(\boldsymbol{Y}_{ti}, \lambda) = \rho h\left(\boldsymbol{Y}_{t-1,i}, \lambda\right) + \sum_{j=1}^{k_1} \beta_j Z_{tij} + \sum_{j=k_1+1}^{k} \beta_j h\left(\mathbf{X}_{tij}, \lambda\right) + \boldsymbol{u}_{ti}, \tag{1}$$

where $t=1, 2, \ldots, T$, $i=1, 2, \ldots, N$, $h(\cdot, \lambda)$ is a monotonic transformation, known except the indexing parameter λ, called the *transformation parameter*. The parameter ρ with $|\rho|<1$ characterizes the dynamic effect. The Z variables may contain the column of ones, dummy variables such as the fixed time effects, as well as other variables that do not need to be transformed. The X variables are continuous covariates that need to be transformed. The disturbance vector $\boldsymbol{u}_t=(\boldsymbol{u}_{t1}, \boldsymbol{u}_{t2}, \ldots, \boldsymbol{u}_{tN})$ is assumed to have random regional effects and spatially autocorrelated residual disturbances, i.e.,

$$\boldsymbol{u}_t = \mu + \boldsymbol{\epsilon}_t, \tag{2}$$

$$\boldsymbol{\epsilon}_t = \delta \mathbf{W} \boldsymbol{\epsilon}_t + v_t, \tag{3}$$

where $\mu_i \sim N(0, \sigma_\mu^2)$ and $v_{ti} \sim N(0, \sigma_v^2)$ for all t and i. The μ_is and $\boldsymbol{\epsilon}_{ti}$s are independent of each other and among themselves. The parameter δ is the spatial autoregressive coefficient and $\mathbf{W}$ is a known $N \times N$ spatial weight matrix whose diagonal elements are zero. It is assumed that $(\mathbf{I}_N - \delta \mathbf{W})$ is non-singular for all $|\delta|<1$, where $\mathbf{I}_N$ is an $N \times N$ identity matrix.

Let $B=\mathbf{I}_N - \delta \mathbf{W}$. We have $\boldsymbol{\epsilon}_t = B^{-1} v_t$. Let $\boldsymbol{Y}$ be the $TN \times 1$ vector of original observations arranged according to $i=1, \ldots, N$ for each of $t=1, \ldots, T$. Let $\boldsymbol{Y}_{-1}$ be the lagged $\boldsymbol{Y}$, and $h(\boldsymbol{Y}, \lambda)$ and $h(\boldsymbol{Y}_{-1}, \lambda)$ be the corresponding transformed $\boldsymbol{Y}$ and $\boldsymbol{Y}_{-1}$. Define $\boldsymbol{Y}(\lambda,\rho)=h(\boldsymbol{Y},\lambda)-\rho h(\boldsymbol{Y}_{-1},\lambda)$. The model specified by Eqs. (1)–(3) can be compactly written in matrix notation as

$$\boldsymbol{Y}(\lambda,\rho)=\mathbf{X}(\lambda)\beta+\boldsymbol{u},\qquad \text{with } \boldsymbol{u}=(\boldsymbol{1}_T\otimes\mathbf{I}_N)\mu+\left(\mathbf{I}_T\otimes B^{-1}\right)v \tag{4}$$

where $\mathbf{X}(\lambda)=(Z,h(\mathbf{X},\lambda))$ is a $TN\times k$ matrix whose rows contain the values of the covariates (transformed or untransformed), $\boldsymbol{1}_T$ is a T-vector of ones and $\otimes$ denotes the Kronecker product. The covariance matrix of $\boldsymbol{u}$ is $\mathbf{E}(\boldsymbol{u}\boldsymbol{u}')=\sigma_v^2\boldsymbol{\Omega}$ with

$$\Omega=\phi(J_T\otimes\mathbf{I}_N)+\mathbf{I}_T\otimes(B'B)^{-1}$$

where $\phi=\sigma_\mu^2/\sigma_v^2$, and $J_T=\boldsymbol{1}_T\boldsymbol{1}_T'$. Putting $\boldsymbol{\theta}=(\phi,\ \delta,\ \lambda,\ \rho)'$ and assuming the initial observations $\boldsymbol{Y}_{0i}$, $i=1,\ldots,\ N$, contained in the lagged vector $\boldsymbol{Y}_{-1}$ are constants,[2] the log likelihood function (after dropping the constant term) has the form

$$\ell\left(\beta,\sigma_v^2,\boldsymbol{\theta}\right)=-\frac{TN}{2}\log\left(\sigma_v^2\right)-\frac{1}{2}\log|\Omega|-\frac{1}{2\sigma_v^2}\boldsymbol{u}'\left(\beta,\sigma_v^2,\boldsymbol{\theta}\right)\Omega^{-1}\boldsymbol{u}\left(\beta,\sigma_v^2,\boldsymbol{\theta}\right)+J(\lambda), \tag{5}$$

where $\boldsymbol{u}\left(\beta,\sigma_v^2,\boldsymbol{\theta}\right)=\boldsymbol{Y}(\lambda,\rho)-\mathbf{X}(\lambda)\beta$, and $J(\lambda)=\sum_{t=1}^{T}\sum_{i=1}^{N}\log h_{\boldsymbol{Y}}(\boldsymbol{Y}_{ti},\lambda)$.

Maximizing the log likelihood (Eq. (5)) gives the MLEs of the model parameters. To make this optimization problem feasible, the following procedures are developed. Firstly, the dimension of maximization can be reduced by concentrating out the parameters β and σ_v^2 from $\ell(\beta,\sigma_v^2,\boldsymbol{\theta})$. It is easy to show that given $\boldsymbol{\theta}$, ℓ is maximized at

$$\hat\beta(\boldsymbol{\theta})=\left[\mathbf{X}'(\lambda)\Omega^{-1}\mathbf{X}(\lambda)\right]^{-1}\mathbf{X}'(\lambda)\Omega^{-1}\boldsymbol{Y}(\lambda,\rho),$$
$$\hat\sigma_v^2(\boldsymbol{\theta})=\tfrac{1}{NT}\tilde{\boldsymbol{u}}'(\boldsymbol{\theta})\Omega^{-1}\tilde{\boldsymbol{u}}(\boldsymbol{\theta}),$$

where $\tilde{\boldsymbol{u}}(\boldsymbol{\theta})=\boldsymbol{Y}(\lambda,\rho)-\mathbf{X}(\lambda)\hat\beta(\boldsymbol{\theta})$. Substituting $\hat\beta(\boldsymbol{\theta})$ and $\hat\sigma_v^2(\boldsymbol{\theta})$ back into the log likelihood function (Eq. (5)) for β and σ_v^2, respectively, gives the following concentrated log likelihood (after dropping the constant)

$$\ell_c(\theta)=-\frac{TN}{2}\log\left[\tilde{\boldsymbol{u}}'(\boldsymbol{\theta})\Omega^{-1}\tilde{\boldsymbol{u}}(\boldsymbol{\theta})\right]-\frac{1}{2}\log|\Omega|+J(\lambda). \tag{6}$$

Maximizing $\ell_c(\boldsymbol{\theta})$, subject to $|\delta|<1$ and $|\rho|<1$, gives the MLE $\hat{\boldsymbol{\theta}}$, which upon substitution gives the unconstrained MLEs $\hat\beta=\hat\beta(\hat{\boldsymbol{\theta}})$ and $\hat\sigma_v^2=\hat\sigma_v^2(\hat{\boldsymbol{\theta}})$ for β and σ_v^2, respectively. Further, the unconstrained MLE of σ_μ^2 is given by $\hat\sigma_\mu^2=\hat\phi\hat\sigma_v^2$.

Secondly, maximization of $\ell_c(\boldsymbol{\theta})$ involves repeated evaluations of $\boldsymbol{\Omega}^{-1}$ and $|\boldsymbol{\Omega}|$ for the $NT\times NT$ matrix $\boldsymbol{\Omega}$, which can be a great burden when N or T or both are large. The following identities, given in Magnus (1982),

$$|\Omega|=|(B'B)^{-1}+\phi T\mathbf{I}_N|\cdot|B|^{2(T-1)} \tag{7}$$

$$\Omega^{-1}=(1/T)J_T\otimes\left[(B'B)^{-1}+\phi T\mathbf{I}_N\right]^{-1}+[\mathbf{I}_T-(1/T)J_T]\otimes(B'B) \tag{8}$$

[2] See Hsiao (2003, Ch. 4) for detailed discussion on the effect of the assumptions concerning the initial observations on the likelihood inferences. Our estimation procedure can be adapted to account for different assumptions for the initial observations.

Appeared in: Economics Letters, 2006, 91, 138-145.

reduce the calculations of the inverse and determinant of an $NT \times NT$ matrix to the calculations of the inverse and determinants of several $N \times N$ matrices. Following Griffith (1988), calculation of the determinants can be further simplified by using

$$|B| = \prod_{i=1}^{N} (1 - \delta w_i), \text{ and } |(B'B)^{-1} + \phi T \mathbf{I}_N| = \prod_{i=1}^{N} \left[(1 - \delta w_i)^{-2} + T\phi \right]. \tag{9}$$

where w_i are the eigenvalues of $\mathbf{W}$.

Thirdly, maximization of $\ell_c(\boldsymbol{\theta})$ can be facilitated using the analytical gradients or the concentrated score vector $\boldsymbol{S}_c(\boldsymbol{\theta})$, which can be obtained by either substituting $\beta(\boldsymbol{\theta})$ and $\sigma_v^2(\boldsymbol{\theta})$ into the last four elements of the score function $\mathbf{S}(\beta, \sigma_v^2, \boldsymbol{\theta})$ given in the Appendix, or directly taking partial derivatives of $\ell_c(\boldsymbol{\theta})$:

$$\frac{\partial \ell_c(\boldsymbol{\theta})}{\partial \phi} = \frac{1}{2} \left(\frac{TN \tilde{\boldsymbol{u}}'(\boldsymbol{\theta}) \Omega^{-1} (J_T \otimes \mathbf{I}_N) \Omega^{-1} \tilde{\boldsymbol{u}}(\boldsymbol{\theta})}{\tilde{\boldsymbol{u}}'(\boldsymbol{\theta}) \Omega^{-1} \tilde{\boldsymbol{u}}(\boldsymbol{\theta})} - tr\left[\Omega^{-1} (J_T \otimes \mathbf{I}_N) \right] \right) \tag{10}$$

$$\frac{\partial \ell_c(\boldsymbol{\theta})}{\partial \delta} = \frac{1}{2} \left(\frac{TN \tilde{\boldsymbol{u}}'(\boldsymbol{\theta}) \Omega^{-1} (\mathbf{I}_T \otimes A) \Omega^{-1} \tilde{\boldsymbol{u}}(\boldsymbol{\theta})}{\tilde{\boldsymbol{u}}'(\boldsymbol{\theta}) \Omega^{-1} \tilde{\boldsymbol{u}}(\boldsymbol{\theta})} - tr\left[\Omega^{-1} (\mathbf{I}_T \otimes A) \right] \right) \tag{11}$$

$$\frac{\partial \ell_c(\boldsymbol{\theta})}{\partial \lambda} = J_\lambda(\lambda) - \frac{TN \tilde{\boldsymbol{u}}_\lambda{}'(\boldsymbol{\theta}) \Omega^{-1} \tilde{\boldsymbol{u}}(\boldsymbol{\theta})}{\tilde{\boldsymbol{u}}'(\boldsymbol{\theta}) \Omega^{-1} \tilde{\boldsymbol{u}}(\boldsymbol{\theta})} \tag{12}$$

$$\frac{\partial \ell_c(\boldsymbol{\theta})}{\partial \rho} = \frac{TN h(\boldsymbol{Y}_{-1}, \lambda) \Omega^{-1} \tilde{\boldsymbol{u}}(\boldsymbol{\theta})}{\tilde{\boldsymbol{u}}'(\boldsymbol{\theta}) \Omega^{-1} \tilde{\boldsymbol{u}}(\boldsymbol{\theta})} \tag{13}$$

where $A = (\partial / \partial \delta)(B'B)^{-1} = (B'B)^{-1}(\mathbf{W}'B + B'\mathbf{W})(B'B)^{-1}$ and $\tilde{\boldsymbol{u}}_\lambda(\boldsymbol{\theta}) = (\partial / \partial \lambda)\tilde{\boldsymbol{u}}(\boldsymbol{\theta})$.

The above MLE procedure has been implemented using GAUSS 6.0 with its CO (constrained optimization) and CML (constrained maximum likelihood) procedures. It turns out that the programs work very well for moderate sized data.[3] From the expressions (7)–(9), we see that the main computational task in the iterative maximization process is the repeated evaluation of the matrix $[(B'B)^{-1} + \phi T \mathbf{I}_N]^{-1}$ and the computation of the eigenvalues w_i. As pointed out by Anselin (2001, p. 325), computation of eigenvalues becomes unstable when the $\mathbf{W}$ matrix becomes larger than 1000×1000, and much remains to be done to develop efficient algorithms and data structure to allow for the analysis of very large spatial data sets.

Covariance matrix of $(\hat{\beta}', \hat{\sigma}_v^2, \hat{\boldsymbol{\theta}}')'$ can be estimated by $-\mathbf{H}^{-1}(\hat{\beta}, \hat{\sigma}_v^2, \hat{\boldsymbol{\theta}})$ where $\mathbf{H}(\beta, \sigma_v^2, \boldsymbol{\theta})$ is the Hessian matrix given in the Appendix. With this, inferences on model parameters can be conveniently carried out following the Wald procedure. Of particular interest are the model specification tests related to the parameter vector $\boldsymbol{\theta}$. Likelihood ratio and Lagrange multipliers procedures can also be used for inferences.

[3] The programs are available from the first author upon request.

Appeared in: Economics Letters, 2006, 91, 138-145.

3. An empirical illustration

In this section, we use a well known data set: the cigarettes demand for the United States, to illustrate the applications of our model and inference method. The data set contains a panel of 46 states over 30 time periods (1963–1992) and is listed as CIGAR.TXT on the Wiley web site associated with Baltagi (2001). The response variable Y=Cigarette sales in packs per capita. The covariates are X_1=Price per pack of cigarettes; X_2=Population (Pop); X_3=Population above the age of 16; X_4=Consumer price index with (1983=100); X_5=Per capita disposable income; and X_6=Minimum price in adjoining states per pack of cigarettes.

Earlier studies regarding demand for cigarettes include Hamilton (1972), McGuiness and Cowling (1975), Baltagi and Levin (1986, 1992), and Baltagi et al. (2000), where a habit-persistence type of dynamic demand model is developed and followed. However, no formal consideration is given to explicitly model the spatial dependence. Also, the functional form used in these studies is the fixed log–log form. Baltagi and Li (2004) argued how the spatial autocorrelation may arise in the demand for cigarettes and considered prediction problem using a random effect model with spatial error. But dynamic effect and functional form transformation were not considered in the model. It is thus important to formally assess the existence/nonexistence of the dynamic effect, spatial dependence as well as the functional form transformation in the context of cigarettes demand.

We consider fitting of three models: (I) both response and covariates are log transformed; (III) response is Box–Cox transformed (Box and Cox, 1964), and covariates are log transformed; and (III) both response and covariates are Box–Cox transformed. For the spatial weighting matrix **W**, we follow the first-order rook's contiguity relations. See Kelejian and Robinson (1995) for a good discussion on the spatial weighing matrix. The results are summarized in Table 1. The maximum of the concentrated log likelihood without the constant term is listed in the last row labeled as loglik.

From the results in Table 1 we see that there exists strong evidence for the existence of dynamic effect, spatial dependence, as well as Box–Cox functional form, in particular the dynamic effect. It is interesting to note that three models give quite consistent estimates of spatial error correlation and

Table 1
Estimation results for the cigarette demand data

	Model I		Model II		Model III	
	Par. est.	*t*-stat	Par. est.	*t*-stat	Par. est.	*t*-stat
β_0	0.6600	7.1847	0.4669	6.6540	−0.1594	−0.8946
β_1	−0.2143	−10.6937	−0.2233	−3.5331	−0.2072	−12.3892
β_2	−0.2457	−3.9874	−0.2627	−2.8806	−0.1685	−1.6363
β_3	0.2406	3.9130	0.2575	2.8581	0.1463	1.5328
β_4	−0.0240	−1.1467	−0.0266	−1.1780	−0.0200	−1.1120
β_5	0.0793		0.0865	2.7881	0.3172	4.6179
β_6	0.0951	5.5560	0.0954	3.0094	0.0628	4.3113
σ	0.0365	691.6175	0.0161	114.8971	0.0117	221.3993
ϕ	0.2922	2.8772	0.3307	2.9748	0.5932	3.3161
δ	0.2338	9.1928	0.2397	9.3655	0.2515	9.7891
λ	–	–	−0.1695	−3.0222	−0.2329	−5.8438
ρ	0.8644	62.4733	0.8577	61.0391	0.8256	54.7972
loglik	−3723.30		−3718.75		−3709.07	

Appeared in: Economics Letters, 2006, 91, 138-145.

dynamic effect coefficient. Model I is embedded in both Model II and Model III with λ specified as zero. Wald and likelihood ratio tests of Model I against Model II, or against Model III strongly reject the Model I specification. Thus, the conventional Cobb–Douglas functional form specification for the cigarettes demand is not supported by the data. Furthermore, all three models show the significance of random individual effects. According to the loglik values, Model III fits the data the best.

4. Conclusions

A generalized error components model is proposed that explicitly takes into account three major issues in the modelling of economic panel data: dynamics, spatial dependence and functional form, where $h(\cdot,\lambda)$ can be any smooth monotonic function. This model and its MLE procedures developed in this paper allow easy testing of these three effects as well as other standard effects incorporated in the model. We emphasize that when a priori information or theoretical foundation are lacking, it is important to start the empirical investigations with a more general model and then test for the existence/nonexistence of certain effects.

Consideration of a flexible functional form in spatial panel framework is fairly new, although there have been some considerations in a cross sectional set up. See, for example, van Gastel and Paelinck (1995), Griffith et al. (1998), Baltagi and Li (2001), and Pace et al. (2004). Much remains to be done to develop other methods of estimation for the current model, and to include flexible functional form in other spatial panel models such as spatial lag models.

Acknowledgements

Z. Yang acknowledges research support (C208/MSS3E110) from the Wharton-SMU Research Center, Singapore Management University.

Appendix A. The score and Hessian functions

Some of the matrix differential formulas that are useful in our derivation can be found in Magnus (1982) or Magnus and Neudecker (1999). Denote $\boldsymbol{u}=\boldsymbol{u}(\beta,\lambda,\rho)$. Let $\boldsymbol{u}_\lambda=\partial\boldsymbol{u}/\partial\lambda$ and $\boldsymbol{u}_{\lambda\lambda}=\partial^2\boldsymbol{u}/\partial\lambda^2$. We obtain the score function $\mathbf{S}(\beta,\sigma_v^2,\boldsymbol{\theta})=\partial\ell(\beta,\sigma_v^2,\boldsymbol{\theta})/\partial(\beta',\sigma_v^2,\boldsymbol{\theta}')'$ as

$$\mathbf{S}(\beta,\sigma_v^2,\boldsymbol{\theta})=\begin{cases}\frac{1}{\sigma_v^2}\mathbf{X}'(\lambda)\Omega^{-1}\boldsymbol{u}\\ \frac{1}{2\sigma_v^4}\boldsymbol{u}'\Omega^{-1}\boldsymbol{u}-\frac{NT}{2\sigma_v^2}\\ \frac{1}{2\sigma_v^2}\boldsymbol{u}'\Omega^{-1}(J_T\otimes\mathbf{I}_N)\Omega^{-1}\boldsymbol{u}-\frac{1}{2}tr\left[\Omega^{-1}(J_T\otimes\mathbf{I}_N)\right]\\ \frac{1}{2\sigma_v^2}\boldsymbol{u}'\Omega^{-1}(\mathbf{I}_T\otimes A)\Omega^{-1}\boldsymbol{u}-\frac{1}{2}tr\left[\Omega^{-1}(\mathbf{I}_T\otimes A)\right]\\ J_\lambda(\lambda)-\frac{1}{\sigma_v^2}\boldsymbol{u}_\lambda{}'\Omega^{-1}\boldsymbol{u}\\ \frac{1}{\sigma_v^2}h(\boldsymbol{Y}_{-1},\lambda)\Omega^{-1}\boldsymbol{u}\end{cases}$$

and the Hessian matrix $\mathbf{H}(\beta,\sigma_v^2,\boldsymbol{\theta})=\partial\mathbf{S}(\beta,\sigma_v^2,\boldsymbol{\theta})/\partial(\beta',\sigma_v^2,\boldsymbol{\theta}')$ with elements

Appeared in: Economics Letters, 2006, 91, 138-145.

$$\mathbf{H}_{\beta\beta} = -\frac{1}{\sigma_v^2}\mathbf{X}'(\lambda)\Omega^{-1}\mathbf{X}(\lambda)$$

$$\mathbf{H}_{\beta\sigma_v^2} = -\frac{1}{\sigma_v^4}\mathbf{X}'(\lambda)\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\beta\phi} = -\frac{1}{\sigma_v^2}\mathbf{X}'(\lambda)\Omega^{-1}(J_T\otimes\mathbf{I}_N)\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\beta\delta} = -\frac{1}{\sigma_v^2}\mathbf{X}'(\lambda)\Omega^{-1}(\mathbf{I}_T\otimes A)\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\beta\lambda} = \frac{1}{\sigma_v^2}\left[\mathbf{X}_\lambda{}'(\lambda)\Omega^{-1}\boldsymbol{u}+\mathbf{X}'(\lambda)\Omega^{-1}\boldsymbol{u}_\lambda\right]$$

$$\mathbf{H}_{\beta\rho} = -\frac{1}{\sigma_v^2}\mathbf{X}'(\lambda)\Omega^{-1}h(\boldsymbol{Y}_{-1},\lambda)$$

$$\mathbf{H}_{\sigma_v^2\sigma_v^2} = \frac{NT}{2\sigma_v^4}-\frac{1}{\sigma_v^6}\boldsymbol{u}'\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\sigma_v^2\phi} = -\frac{1}{2\sigma_v^4}\boldsymbol{u}'\Omega^{-1}(J_T\otimes\mathbf{I}_N)\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\sigma_v^2\delta} = -\frac{1}{2\sigma_v^4}\boldsymbol{u}'\Omega^{-1}(\mathbf{I}_T\otimes A)\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\sigma_v^2\lambda} = \frac{1}{\sigma_v^4}\boldsymbol{u}_\lambda'\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\sigma_v^2\rho} = -\frac{1}{\sigma_v^4}\boldsymbol{u}'\Omega^{-1}h(\boldsymbol{Y}_{-1},\lambda)$$

$$\mathbf{H}_{\phi\phi} = \frac{1}{2}tr\left[\Omega^{-1}(J_T\otimes\mathbf{I}_N)\Omega^{-1}(J_T\otimes\mathbf{I}_N)\right]-\frac{1}{\sigma_v^2}\boldsymbol{u}'\Omega^{-1}(J_T\otimes\mathbf{I}_N)\Omega^{-1}(J_T\otimes\mathbf{I}_N)\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\phi\delta} = \frac{1}{2}tr\left[\Omega^{-1}(\mathbf{I}_T\otimes A)\Omega^{-1}(J_T\otimes\mathbf{I}_N)\right]-\frac{1}{\sigma_v^2}\boldsymbol{u}'\Omega^{-1}(\mathbf{I}_T\otimes A)\Omega^{-1}(J_T\otimes\mathbf{I}_N)\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\phi\lambda} = \frac{1}{\sigma_v^2}\boldsymbol{u}_\lambda'\Omega^{-1}(J_T\otimes\mathbf{I}_N)\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\phi\rho} = -\frac{1}{\sigma_v^2}\Omega^{-1}(J_T\otimes\mathbf{I}_N)\Omega^{-1}h(\boldsymbol{Y}_{-1},\lambda)$$

$$\mathbf{H}_{\delta\delta} = \frac{1}{2}tr\left[\Omega^{-1}(\mathbf{I}_T\otimes A)\Omega^{-1}(\mathbf{I}_T\otimes A)-\Omega^{-1}\left(\mathbf{I}_T\otimes\frac{\partial A}{\partial\delta}\right)\right]-\frac{1}{\sigma_v^2}\boldsymbol{u}'\Omega^{-1}(\mathbf{I}_T\otimes A)\Omega^{-1}(\mathbf{I}_T\otimes A)\Omega^{-1}\boldsymbol{u}$$
$$+\frac{1}{2\sigma_v^2}\boldsymbol{u}'\Omega^{-1}\left(\mathbf{I}_T\otimes\frac{\partial A}{\partial\delta}\right)\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\delta\lambda} = \frac{1}{\sigma_v^2}\boldsymbol{u}_\lambda'\Omega^{-1}(\mathbf{I}_T\otimes A)\Omega^{-1}\boldsymbol{u}$$

$$\mathbf{H}_{\delta\rho} = -\frac{1}{\sigma_v^2}\Omega^{-1}(\mathbf{I}_T\otimes A)\Omega^{-1}h(\boldsymbol{Y}_{-1},\lambda)$$

$$\mathbf{H}_{\lambda\lambda} = -\frac{1}{\sigma_v^2}\left(\boldsymbol{u}_{\lambda\lambda}'\Omega^{-1}\boldsymbol{u}+\boldsymbol{u}_\lambda'\Omega^{-1}\boldsymbol{u}_\lambda\right)+J_{\lambda\lambda}(\lambda)$$

$$\mathbf{H}_{\lambda\rho} = \frac{1}{\sigma_v^2}\left[h_\lambda'(\boldsymbol{Y}_{-1},\lambda)\Omega^{-1}\boldsymbol{u}+\boldsymbol{u}_\lambda'\Omega^{-1}h(\boldsymbol{Y}_{-1},\lambda)\right]$$

$$\mathbf{H}_{\rho\rho} = -\frac{1}{\sigma_v^2}h'(\boldsymbol{Y}_{-1},\lambda)\Omega^{-1}h(\boldsymbol{Y}_{-1},\lambda)$$

where A is given in Eq. (12) and $\partial A/\partial\delta = 2(B'B)^{-1}[\mathbf{W}'B+B'\mathbf{W})A-\mathbf{W}'\mathbf{W}]$.

References

Anselin, L., 2001. Spatial econometrics. In: Baltagi, B.H. (Ed.), A Companion to Theoretical Econometrics. Blackwell.

Baltagi, B.H., 2001. Econometric Analysis of Panel Data. John Wiley and Sons, New York.

Baltagi, B.H., Levin, D., 1986. Estimating dynamic demand for cigarettes using panel data: the effects of bootlegging, taxation, and advertising reconsidered. Review of Economics and Statistics 68, 148–155.

Appeared in: Economics Letters, 2006, 91, 138-145.

Baltagi, B.H., Levin, D., 1992. Cigarette taxation: raising revenue and reducing consumption. Structure Change and Economic Dynamics 3, 321–335.

Baltagi, B.H., Li, D., 2001. LM tests for functional form and spatial correlation. International Regional Science Review 24, 194–225.

Baltagi, B.H., Li, D., 2004. Prediction in the panel data model with spatial correlation. In: Anselin, L., Florax, R., Rey, S.J. (Eds.), Advances in Spatial Econometrics. Springer-Verlag, Berlin.

Baltagi, B.H., Griffin, J.M., Xiong, W., 2000. To pool or not to pool: homogeneous versus heterogeneous estimators applied to cigarette demand. Review of Economics and Statistics 82, 117–126.

Box, G.E.P., Cox, D.R., 1964. An analysis of transformations (with discussion). Journal of the Royal Statistical Society B. 26, 211–246.

Davidson, R., MacKinnon, J.G., 1993. Estimation and Inference in Econometrics. Oxford University Press, Oxford.

Elhorst, J.P., 2003. Specification and estimation of spatial panel data models. International Regional Science Review 26, 244–268.

Griffith, D.A., 1988. Advanced Spatial Statistics. Kluwer, Dordrecht, the Netherlands.

Griffith, D.A., Paelinck, J.H.P., van Gastel, R.A., 1998. The Box–Cox transformation: computational and interpretation features of the parameters. In: Griffith, D.A., Amrhein, C. (Eds.), Econometric Advances in Spatial Modelling and Methodology. Kluwer, Amsterdam.

Hamilton, J.L., 1972. The demand for cigarettes: advertising, the health care, and the cigarette advertising ban. Review of Economics and Statistics 54, 401–411.

Hsiao, C., 2003. Analysis of Panel Data, 2nd ed. Cambridge University Press, Cambridge.

Kelejian, H., Robinson, D.P., 1995. Spatial correlations: a suggested alternative to the autoregressive model. In: Anselin, L., Florax, R. (Eds.), New Directions in Spatial Econometrics. Springer, Berlin.

Magnus, J.R., 1982. Multivariate error components analysis of linear and nonlinear regression models by maximum likelihood. Journal of Econometrics 19, 239–285.

Magnus, J.R., Neudecker, H., 1999. Matrix Differential Calculus with Applications in Statistics and Econometrics. Revised edition. John Wiley and Sons.

McGuiness, T., Cowling, K., 1975. Advertising and the aggregate demand for cigarettes. European Economic Review 6, 311–328.

Pace, R.K., Barry, R., Slawson Jr., V.C., Sirmans, C.F., 2004. Simultaneous spatial and functional form transformation. In: Anselin, L., Florax, R., Rey, S.J. (Eds.), Advances in Spatial Econometrics. Springer-Verlag, Berlin.

van Gastel, R.A., Paelinck, J.H.P., 1995. Computation of Box–Cox transform parameter: a new method and its applications to spatial econometrics. In: Anselin, L., Florax, R. (Eds.), New Directions in Spatial Econometrics. Springer-Verlag.