Interpreting Dynamic Space-Time Panel Data Models

Nicolas Debarsy, Cem Ertur, James P. Lesage

Laboratoire d'Economie d'Orléans

Document de Recherche

n° 2010-22

« Interpreting Dynamic Space-Time Panel Data Models»

Cem ERTUR
Nicolas DEBARSY
James P. LeSAGE

Laboratoire d'Economie d'Orléans – UMR CNRS 6221 Faculté de Droit, d'Economie et de Gestion, Rue de Blois, B.P. 6739 – 45067 Orléans Cedex 2 - France
Tél : 33 (0)2 38 41 70 37 – 33 (0)2 38 49 48 19 – Fax : 33 (0)2 38 41 73 80
E-mail : leo@univ-orleans.fr - http://www.univ-orleans.fr/DEG/LEO

Interpreting dynamic space-time panel data models

Nicolas Debarsy[1,a], Cem Ertur[2], and James P. LeSage[3,*]

[1] *CERPE, University of Namur, Rempart de la vierge, 8, 5000 Namur, Belgium, ndebarsy@fundp.ac.be*

[2] *Université d'Orléans, Rue de Blois - BP 6739, 45067 Orléans Cedex 2, France, cem.ertur@univ-orleans.fr*

[3] *Texas State University-San Marcos, Department of Finance & Economics, 601 University Drive, San Marcos, TX 78666, USA, jlesage@spatial-econometrics.com*

Abstract

There is a great deal of literature regarding the asymptotic properties of various approaches to estimating simultaneous space-time panel models, but little attention has been paid to how the model estimates should be interpreted. The motivation for use of space-time panel models is that they can provide us with information not available from cross-sectional spatial regressions. [8] show that cross-sectional simultaneous spatial autoregressive models can be viewed as a limiting outcome of a dynamic space-time autoregressive process. A valuable aspect of dynamic space-time panel data models is that the own- and cross-partial derivatives that relate changes in the explanatory variables to those that arise in the dependent variable are explicit. This allows us to employ parameter estimates from these models to quantify dynamic responses over time and space as well as space-time diffusion impacts. We illustrate our approach using the demand for cigarettes over a 30 year period from 1963-1992, where the motivation for spatial dependence is a bootlegging effect where buyers of cigarettes near state borders purchase in neighboring states if there is a price advantage to doing so.

JEL Classification: C11, C23

Keywords: Dynamic space-time panel data model,MCMC estimation, dynamic responses over time and space.

---

a. Nicolas Debarsy is research fellow at the F.R.S - FNRS and gratefully acknowledges their financial support.

∗. Corresponding author.

Résumé

La littérature économétrique récente fait une place croissante à l'étude des propriétés asymptotiques des différentes méthodes d'estimation des modèles de données de panel spatio-temporels. Toutefois, force est de constater que peu d'attention est consacrée à l'interprétation économique de tels modèles malgré leur grand intérêt pour la modélisation des phénomènes économiques dans une dimension spatio-temporelle et le rôle qu'ils pourraient jouer dans l'évaluation des politiques économiques dans cette même dimension. Nous montrons dans ce papier que les coefficients estimés de ces modèles permettent d'expliciter non seulement la dynamique temporelle des impacts mais également leur dynamique spatiale et surtout de quantifier la diffusion spatio-temporelle de l'impact d'une variation d'une variable explicative. La méthode proposée est illustrée par une étude de la demande de cigarettes dans 46 Etats américains sur la période 1963-1992 en utilisant une base de données bien connue dans la littérature économétrique. La présence d'autocorrélation spatiale est ici motivée par un effet de " contrebande ". Les consommateurs proches des frontières d'un état achèteront en effet leurs cigarettes dans les états voisins si le prix des cigarettes y est inférieur à celui pratiqué dans leur propre Etat.

Classification JEL : C11, C23
Mots clés : Modèles de données de panel dynamiques spatio-temporels, Estimation par MCMC, Réponses dynamiques temporelles et spatiales.

1 Introduction

There are obvious linkages between cross-sectional and dynamic models. [8] begin with the relationship in (1), where the time lag of spatially weighted neighboring values $\mathbf{W}\mathbf{y}_{t-1}$ is introduced, in addition to a matrix of explanatory variables $\mathbf{X}_t$.

$$\mathbf{y}_t = \rho\mathbf{W}\mathbf{y}_{t-1} + \mathbf{X}_t\beta + \varepsilon_t \tag{1}$$

The $N \times N$ matrix $\mathbf{W}$ is a spatial weight matrix whose i, jth element takes some positive value if regions i and j are neighbors and zero otherwise. Main diagonal elements are set to zero and the matrix is normalized to have row sums of unity. This means that the vector $\mathbf{W}\mathbf{y}_{t-1}$ represents a linear combination of previous period values from neighboring regions. The $N \times K$ matrix $\mathbf{X}_t$ on the right-hand-side of (1) (which might include an intercept) is assumed to represent explanatory variables in the relationship that do not change over time (or more generally follow some deterministic time path). Recursive substitution of $\mathbf{y}_{t-1}$ in (1) over q periods leads to:

$$\begin{aligned} \mathbf{y}_t &= \left(\mathbf{I}_N + \rho\mathbf{W} + \rho^2\mathbf{W}^2 + \ldots + \rho^{q-1}\mathbf{W}^{q-1}\right)\mathbf{X}\beta \\ &+ \rho^q\mathbf{W}^q\mathbf{y}_{t-q} + \mathbf{u}_t \qquad (2) \\ \mathbf{u}_t &= \varepsilon_t + \rho\mathbf{W}\varepsilon_{t-1} + \rho^2\mathbf{W}^2\varepsilon_{t-2} + \ldots + \rho^{q-1}\mathbf{W}^{q-1}\varepsilon_{t-(q-1)} \qquad (3) \end{aligned}$$

[8] show that when q is large, the expected value of (2), shown in (5), corresponds to the mean of the cross-sectional simultaneous spatial lag model, expressed in (4), which can be viewed as the outcome of a long-run equilibrium or steady state.

$$\begin{aligned} \mathbf{y} &= \rho\mathbf{W}\mathbf{y} + \mathbf{X}\beta + \epsilon \qquad (4) \\ \lim_{q\to\infty} E(\mathbf{y}) &= (\mathbf{I}_N - \rho\mathbf{W})^{-1}\mathbf{X}\beta \qquad (5) \end{aligned}$$

The model in (4) has been labeled a SAR model in the spatial econometrics literature and it serves as the workhorse of cross-sectional spatial regression modeling. In our application, cigarette sales in state i depend on those of neighboring states because of the "bootlegging" phenomena, where buyers of cigarettes near state borders purchase in neighboring states if there is a price advantage to doing so. This provides a motivation for the spatial lag variable $\mathbf{Wy}$.

The own- and cross-partial derivatives: $\partial \mathbf{y}/\partial \mathbf{x}_r'$ for the SAR model take the form of an $N \times N$ matrix that can be expressed as:

$$\partial \mathbf{y}/\partial \mathbf{x}_r' \;=\; (\mathbf{I}_N - \rho \mathbf{W})^{-1} \mathbf{I}_N \beta_r \qquad (6)$$

These partial derivatives show how changes in say the price of cigarettes (x_{jr}) in state j influence cigarette sales in state i. [8] propose using the average of the main diagonal elements of this $N \times N$ matrix as a scalar summary measure of the own-partial derivatives that they label a direct (own-region) effect. [1] As in the case of regression coefficients, the direct effect averages over all observations/regions to produce a scalar summary measure of the own-partial derivative for all regions. They also propose an average of the (cumulative) off-diagonal elements over all rows (observations) to produce a scalar summary that corresponds to the cross-partial derivative or indirect (other region) effect associated with changes in the rth explanatory variable. In our application this would measure the average magnitude of the bootlegging effect.

In addition to proposing scalar summary measures for the effects, [8] provide a computationally efficient approach to determining measures of dispersion for these scalar summary effects estimates. These can be used to draw inferences regarding the statistical significance of the direct and indirect effects estimates for the explanatory variables in the model. Since the matrix inverse $(\mathbf{I}_N - \rho \mathbf{W})^{-1}$ from (6) can be expressed as: $(\mathbf{I}_N + \rho \mathbf{W} + \rho^2 \mathbf{W}^2 + \rho^3 \mathbf{W}^3 \ldots)$, the partial derivatives can also be used to investigate what [8] label *marginal effect/impacts* which show how spillovers decay with respect to order of the neighbors. In these models, $\mathbf{W}$ represents (first-order) neighboring observations while $\mathbf{W}^2$ reflects neighbors to these neighbors (second-order), and so on for higher powers $\mathbf{W}^p$.

Our focus is on extending this reasoning to the case of dynamic space-time panel data models. This specification allows us to compute own- and cross-partial derivatives that trace the effects (own-region and other-region) through time and space. Space-time dynamic models produce a situation where a change in the ith observation of the rth explanatory variable at

1. The direct effect for region i includes some feedback loop effects that arise as a result of impacts passing through neighboring regions j and back to region i.

time t will produce contemporaneous and future responses in all regions' dependent variables y_{it+T} as well as other region future responses y_{jt+T}. This is due to the presence of an individual time lag (capturing time dependence), a spatial lag (that accounts for spatial dependence) and a cross-product term reflecting the space-time diffusion.

To the best of our knowledge, [9] is the only study dealing with impacts coefficients for both space and time. They consider a time-space dynamic model that relates commuting times to highway expenditures. It seems clear that expenditures for an improvement in a single highway segment at time t (say segment i) will improve commuting times for those traveling on this highway segment (say y_{it}).[2] Improvements in the segment i will also produce future benefits of improved travel times to those using segment i $(y_{it+T}, T = 1, \ldots)$. Equally important is the fact that commuting times on neighboring roadways will also improve in current and future time periods, which we might denote as: y_{jt} and y_{jt+T} where $j \neq i$. This is because less congestion on one highway segment will improve traffic flow on neighboring segments. It might also be the case that commuters adjust commuting patterns over time to take advantage of the improvements made in highway segment i and their impact on lessening congestion of nearby arteries.

Dynamic space-time panel data models have the ability to quantify these changes which should prove extremely useful in numerous applied modeling situations. We show that the partial derivatives $\partial \mathbf{y}_t / \partial \mathbf{x}'_{rt}$ for these models take the form of an $N \times N$ matrix for time t and those for the cumulative effects of a change taking place in time t at future time horizon T take the form of a sum of T different $N \times N$ matrices. We derive explicit forms for these as a function of the dynamic space-time panel data model parameter estimates. This allows us to calculate the dynamic responses over time and space that arise from changes in the explanatory variables. In addition to setting forth expressions for the partial derivatives we also propose scalar summary measures for these and take up the issue of efficient calculation of measures of dispersion.

Section 2 of the paper describes the dynamic space-time panel data model along with Bayesian Markov Chain Monte Carlo (MCMC) estimation procedures. Section 3 sets forth analytical expressions for the partial derivatives

2. We abstract from the issue of time scale here and assume that measurements are taken over a sufficient period of time (say one year) to allow the improvements to be made in time t and for commuters to travel on the highway segment during some part of the year (time t).

along with our proposed scalar summary measures for the own- and cross-partials. We also discuss interpretative considerations related to whether one is interested in responses to one-period changes in the explanatory variables or permanent changes in the level of these variables. Section 4 illustrates our approach using a panel dataset from [4] that relates state-level cigarette sales to prices and income over time. A final section contains our conclusions.

2 The dynamic space-time panel data model

[1] and [14] consider a dynamic *spatial lag* panel model that allows for both time and spatial dependence as well as a cross-product term reflecting spatial dependence at a one-period time lag. We add spatially lagged exogenous variables to the set of covariates, leading to a dynamic *spatial Durbin model* shown in (7).

$$\begin{aligned} \mathbf{y}_t &= \phi \mathbf{y}_{t-1} + \rho \mathbf{W} \mathbf{y}_t + \theta \mathbf{W} \mathbf{y}_{t-1} + \iota_N \alpha + \mathbf{x}_t \beta + \mathbf{W} \mathbf{x}_t \gamma + \eta_t \\ \eta_t &= \mu + \varepsilon_t \quad t = 1, \ldots, T, \end{aligned} \tag{7}$$

Where $\mathbf{y}_t$ is the N-dimensional vector of the dependent variable, $\mathbf{x}_t$ the $N \times K$ matrix of explanatory variables, and β and γ, K-dimensional vectors of coefficients associated with the covariates and their spatial lag ($\mathbf{W}\mathbf{x}_t$). $\mathbf{W}$ is the $N \times N$ spatial weight matrix that identifies neighboring regions, ι_N is an $N \times 1$ column vectors of ones with α the associated intercept parameter, ρ the spatial dependence parameter, ϕ the autoregressive time dependence parameter, and θ the spatio-temporal diffusion parameter. We assume ε_t is *i.i.d.* across i and t with zero mean and variance $\sigma_\varepsilon^2 \mathbf{I}_N$. The $N \times 1$ column vector μ represents individual effects with $\mu_i \sim N(0, \sigma_\mu^2)$, and it is typically assumed that μ is uncorrelated with ε_t.

In this paper, we use a one way error component to model individual heterogeneity. However, our results would also apply to a (time-space dynamic panel) model with fixed effects such as that from [14]. [9] and [10] propose a general framework for specifying space-time dependence that involves applying space and time filter expressions to the dependent variable vector Y or the disturbances.

Let $\mathbf{Y}_a = (\mathbf{y}_0', \ldots, \mathbf{y}_T')'$, and $\mathbf{A}$ be the $T{+}1 \times T{+}1$ time filter matrix shown in (8), which includes the term ψ from the Prais-Winsten transformation for the initial period.

$$\mathbf{A} = \begin{pmatrix} \psi & 0 & \dots & 0 \\ -\phi & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & -\phi & 1 \end{pmatrix} \tag{8}$$

Specification of ψ, the (1,1) element in $\mathbf{A}$ depends on whether the first period is modeled or assumed to be known. We will not model this but rather condition on the initial period, since our focus is on interpretation not estimation of these models. [3] Assuming the process is stationary, ψ is given by:

$$\psi = \sqrt{1 - \phi^2}, \quad |\phi| < 1. \tag{9}$$

The filter for spatial dependence is defined as a nonsingular matrix $\mathbf{B} = (\mathbf{I}_N - \rho\mathbf{W})$. As already noted, $\mathbf{W}$ defines dependence between the cross-sectional (spatial) observations. We will also assume that $\mathbf{W}$ is row-normalized from a symmetric matrix, so that all eigenvalues are real and less than or equal to one.

The two filter expressions are combined using the Kronecker product of the matrices $\mathbf{A}$ and $\mathbf{B}$:

$$\mathbf{A} \otimes \mathbf{B} = \mathbf{I}_{N,T+1} - \rho\mathbf{I}_{T+1} \otimes \mathbf{W} - \phi\mathbf{L} \otimes \mathbf{I}_N + (\rho \times \phi)\mathbf{L} \otimes \mathbf{W} \tag{10}$$

where $\mathbf{L}$ is the $(T+1) \times (T+1)$ matrix time-lag operator. This filter implies a restriction that θ, the parameter associated with spatial effects from the previous period ($\mathbf{L}\otimes\mathbf{W}$) is equal to $-\rho\times\phi$. [10] show that applying this space-time filter to the error terms greatly simplifies estimation and [9] illustrate that interpretation of these models is also simplified by this restriction. The restriction produces a situation where space and time are separable, leading to simplifications in the space-time covariance structure as well as the own- and cross-partial derivatives used to interpret the model. We will have more to say about this later.

We consider the more general case shown in (11), where the simplifying restriction is not imposed, leading to three parameters ϕ, ρ, θ which will be estimated.

$$\mathbf{A} \otimes \mathbf{B} = \mathbf{I}_{N,T+1} - \rho\mathbf{I}_{T+1} \otimes \mathbf{W} - \phi\mathbf{L} \otimes \mathbf{I}_N - \theta\mathbf{L} \otimes \mathbf{W}, \tag{11}$$

3. See [11] for a discussion of issues pertaining to this.

Applying the filter to the dependent variable results in a model specification:

$$\begin{aligned}
(\mathbf{A}\otimes\mathbf{B})\mathbf{Y}_a &= \iota_{N,T+1}\alpha + \mathbf{Z}\beta + \mathbf{WZ}\gamma + \eta \qquad (12)\\
\eta &\sim N(\mathbf{0},\tilde{\mathbf{\Omega}})\\
\tilde{\mathbf{\Omega}} &= \sigma_\mu^2(\mathbf{J}_{T+1}\otimes\mathbf{I}_N) + \sigma_\varepsilon^2\mathbf{I}_{N,T+1}\\
\mathbf{J}_{T+1} &= \iota_{T+1}\iota'_{T+1}
\end{aligned}$$

where $\mathbf{Z} = (\mathbf{x}'_0,\ldots,\mathbf{x}'_T)'$, and we note that all model parameters are assumed to be constant across time and spatial units.

For the case we deal with here where we condition on initial period observations, we work with the new filter $\mathbf{P}$ shown in (13), which corresponds to the filter in (12) where explanatory variable observations for the first time period are deleted:

$$\mathbf{P}_{NT,N(T+1)} = \begin{pmatrix} -(\phi\mathbf{I}_N+\theta\mathbf{W}) & \mathbf{B} & & \mathbf{0} \\ & \ddots & \ddots & \\ \mathbf{0} & & -(\phi\mathbf{I}_N+\theta\mathbf{W}) & \mathbf{B}\end{pmatrix}, \qquad (13)$$

which allows us to rewrite the model in terms of: $\mathbf{e} = (\mathbf{PY_a} - \mathbf{X}\beta - \mathbf{WX}\gamma - \iota_{NT}\alpha)$, with $\mathbf{X} = (\mathbf{x}'_1,\ldots,\mathbf{x}'_T)'$ so the log-likelihood function of the complete sample size (NT) is given by: [4]

$$\begin{aligned}
\ln L_T(\upsilon) &= -\frac{NT}{2}\ln(2\pi) - \frac{1}{2}\ln|\mathbf{\Omega}| + T\sum_{i=1}^{N}\ln[(1-\rho\varpi_i)] - \frac{1}{2}\mathbf{e}'\mathbf{\Omega}^{-1}\mathbf{e}\\
\mathbf{\Omega} &= (T\sigma_\mu^2+\sigma_\varepsilon^2)(\bar{\mathbf{J}}_T\otimes\mathbf{I_N}) + \sigma_\varepsilon^{\mathbf{2}}\left[(\mathbf{I_T}-\bar{\mathbf{J}}_\mathbf{T})\otimes\mathbf{I_N}\right] \qquad (14)\\
\bar{\mathbf{J}}_T &= \mathbf{J}_T/T
\end{aligned}$$

where $\upsilon = (\beta',\gamma',\alpha,\sigma_\varepsilon^2,\sigma_\mu^2,\phi,\rho,\theta)$, and $\varpi_i, i=1,\ldots,N$ represent eigenvalues of the matrix $\mathbf{W}$ which are real and less than or equal to one given our assumptions regarding the row-stochastic matrix $\mathbf{W}$.

4. The random effects parameters have been integrated out and we use the decomposition proposed by [13] to replace $\mathbf{J}_T$ by its idempotent counterpart.

For this specification, stationary conditions are satisfied only if $|\mathbf{AB}^{-1}| < 1$, which requires [10]:

$$\begin{array}{lll} \phi + (\rho+\theta)\varpi_{max} < 1 & \text{if} & \rho+\theta \geq 0 \\ \phi + (\rho+\theta)\varpi_{min} < 1 & \text{if} & \rho+\theta < 0 \\ \phi - (\rho-\theta)\varpi_{max} > -1 & \text{if} & \rho-\theta \geq 0 \\ \phi - (\rho-\theta)\varpi_{min} > -1 & \text{if} & \rho-\theta < 0 \end{array} \quad (15)$$

where ϖ_{min} and ϖ_{max} are the minimum and maximum eigenvalues of W respectively.

This specification where the first period is not modeled allows us to use conventional matrix expressions and decompositions from the panel data literature that reduce the dimensionality of matrices requiring manipulation during estimation. As indicated, we will rely on a Bayesian Markov Chain Monte Carlo estimation scheme to produce estimates of the parameters in the model. Complete details can be found in [11], but we make note of one issue that arises here. The priors for the space-time parameters ϕ, ρ and θ should be defined over the stationary interval in (15). A uniform joint prior distribution over this interval does not produce vague marginal priors. [12] propose different approaches to define priors on a constrained parameter space. Since we are concerned with the parameter vector $(\rho,\ \phi,\ \theta)$, a prior can be constructed that takes the form $p(\rho,\ \phi,\ \theta) = p(\rho)\, p(\phi|\rho)\, p(\theta|\rho, \phi)$.

Assuming that the parameter space for ρ is a compact subset of (-1,1), we can define the following conditional prior $p(\phi|\rho,\theta) \sim U(-1+|\rho-\theta|, 1-|\rho+\theta|)$ based on the stationary interval defined in (15). [5] Then focusing only on the parameters θ and ρ it is easy to show that the conditional prior $p(\theta|\rho) \sim U(-1+|\rho|, 1-|\rho|)$. The last prior is therefore $p(\rho) \sim U(-1, 1)$. Note that the joint prior is a uniform distribution and equal to 1/2 over the parameter space define by stationary interval (15).

For estimation purposes, we assign a prior distribution $p(\alpha, \beta, \gamma, \sigma_u^2, \sigma_\varepsilon^2, \rho, \phi, \theta)$ such that all parameters are a priori independent. Concerning the parameters (α, β'), we estimate separately the intercept term α and the parameters β assuming a non-hierarchical prior of the independent Normal-Gamma variety. Thus,

5. This assumption is also used in [14], and [6] discuss transformations that can be applied to the matrix $\mathbf{W}$ to produce this restricted parameter space for ρ.

$$\begin{aligned} \alpha &\sim N(\alpha_0, M_\alpha^{-1}) & (16)\\ \beta &\sim N(\beta_0, \mathbf{M}_\beta^{-1}) \\ \sigma_\varepsilon^{-2} &\sim G(v_0/2, S_0/2) & (17)\\ \sigma_\mu^{-2} &\sim G(v_1/2, S_1/2) \end{aligned}$$

We use diffuse priors with prior means α_0 and β_0 set to zero, the precision parameter M_α^{-1} set to 10^{12} and $\mathbf{M}_\beta^{-1}$ to $10^{12} I_K$. Parameters for the Gamma priors are all set to 0.001. Having the posterior distribution of the explanatory variables β conditional on the random effects μ is not desirable because these two sets of parameters tend to be highly correlated which can create problems with mixing for the Markov Chain estimation procedure. We use the method proposed by [5] who suggest first sampling β marginalized over μ and then sampling μ conditioned on β. Posterior distributions are standard and can be found in [7].

3 Interpreting the model estimates

Our focus here is on the partial derivative effects associated with changing the explanatory variables in model (7). This model has own- and cross-partial derivatives that measure the impact on y_{it} that arises from changing the value of the rth explanatory variable at time t in region i. Specifically, $\partial y_{it}/\partial X_{it}^r$, represents the contemporaneous direct effect on region i's dependent variable arising from a change in the rth explanatory variable in region i. There is also a cross-partial derivative $\partial y_{jt}/\partial X_{it}^r$ that measures the contemporaneous spatial spillover effect on region j, $j \neq i$.

We are most interested in partial derivatives that measure how region i's dependent variable responds over time to changes in the initial period levels of the explanatory variables. The model allows us to calculate partial derivatives that can quantify the magnitude and timing of dependent variable responses in each region at various time horizons $t + T$ to changes in the explanatory variables at time t. Expressions for these are presented and discussed in the sequel. We simply note here that we are referring to $\partial y_{it+T}/\partial X_{it}^r$ which measures the $T-$horizon own-region i dependent variable response to changes in own-region explanatory variable r, and $\partial y_{jt+T}/\partial X_{it}^r$, that reflects spillovers/diffusion effects over time that impact the dependent variable in other regions when region i's initial period explanatory variables are changed. We distinguish between two different interpretative scenarios, one where the change in explanatory variables represents a permanent or sustained change in the level and the other where we have a transitory (or one-period) change.

We condition on the initial period observation and assume that this period is only subject to spatial dependence. This implies that the dependent variable for the whole sample is written as $\mathbf{Y} = (\mathbf{y}_1', \ldots, \mathbf{y}_T')'$. In this case, the data generating process (DGP) for our model can be expressed by replacing the $NT \times N(T+1)$ space-time filter $\mathbf{P}$ by the $NT \times NT$ matrix $\mathbf{Q}$ as in (18), with $\mathbf{H} = \iota_T \otimes \mathbf{I}_N$, a matrix that assigns the same N random effects to each region for all time periods.

$$\mathbf{Y} = \mathbf{Q}^{-1}[\iota_{NT}\alpha + \mathbf{X}\beta + \mathbf{W}\mathbf{X}\gamma + \mathbf{H}\mu + \varepsilon] \tag{18}$$

$$\mathbf{Y} = \sum_{r=1}^{K} \mathbf{Q}^{-1}(\mathbf{I}_{NT}\beta_r + \mathbf{W}\gamma_r)\mathbf{X}^{(r)} + \mathbf{Q}^{-1}[\iota_{NT}\alpha + \mathbf{H}\mu + \varepsilon] \tag{19}$$

$$\mathbf{Q} = \begin{pmatrix} \mathbf{B} & \mathbf{0} & \ldots & & \mathbf{0} \\ \mathbf{A} & \mathbf{B} & & & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \ddots & & \vdots \\ \vdots & & \ddots & & \\ \mathbf{0} & \ldots & & \mathbf{A} & \mathbf{B} \end{pmatrix} \tag{20}$$

$$\mathbf{A} = -(\phi\mathbf{I}_N + \theta\mathbf{W})$$

$$\mathbf{B} = (\mathbf{I}_N - \rho\mathbf{W})$$

In (19) we let $\mathbf{X}^{(r)}$ denote the rth column from the $NT \times K$ matrix $\mathbf{X}$, allowing us to express this DGP in a form suitable for considering the partial derivative impacts that arise from changes in the rth explanatory variable. For future reference we note that the matrix $\mathbf{Q}^{-1}$ takes the form of a lower-triangular block matrix, containing blocks with $N \times N$ matrices.

$$\mathbf{Q}^{-1} = \begin{pmatrix} \mathbf{B}^{-1} & \mathbf{0} & \ldots & & \mathbf{0} \\ \mathbf{D}_1 & & & & \vdots \\ \mathbf{D}_2 & \mathbf{D}_1 & \ddots & & \\ \vdots & & \ddots & & \mathbf{0} \\ \mathbf{D}_{T-1} & \mathbf{D}_{T-2} & \ldots & \mathbf{D}_1 & \mathbf{B}^{-1} \end{pmatrix} \tag{21}$$

$$\mathbf{D}_s = (-1)^s(\mathbf{B}^{-1}\mathbf{A})^s\mathbf{B}^{-1}, \quad s = 0, \ldots, T-1$$

One implication of this is that we need only calculate $\mathbf{A}$ and $\mathbf{B}^{-1}$ to analyze the partial derivative impacts for any time horizon T. This means we can use a panel involving say 10 years to analyze the cumulative impacts arising from a permanent (or transitory) change in explanatory variables at any time t extending to future horizons $t + T$.

The one-period-ahead impact of a (permanent) change in the rth variable at time t is:

$$\partial\mathbf{Y}_{t+1}/\partial\mathbf{X}_t^{r'} = \left(\mathbf{D}_1 + \mathbf{B}^{-1}\right)[\mathbf{I}_N\beta_r + \mathbf{W}\gamma_r] \tag{22}$$

and more generally, the T-period-ahead (cumulative) impact arising from a permanent change at time t in $\mathbf{X}_t^r$ takes the form in (23). Note that we are cumulating down the columns (or rows) of the matrix in (21). [6]

$$\begin{aligned} \partial\mathbf{Y}_{t+T}/\partial\mathbf{X}_t^{r'} &= \sum_{s=0}^{T}\mathbf{D}_s[\mathbf{I}_N\beta_r + \mathbf{W}\gamma_r] \\ \mathbf{D}_s &= (-1)^s(\mathbf{B}^{-1}\mathbf{A})^s\mathbf{B}^{-1} \end{aligned} \tag{23}$$

By analogy to [8], the main diagonal elements of the $N \times N$ matrix sums in (23) for time horizon T represent (cumulative) own-region impacts that arise from both time and spatial dependence. The sum of off-diagonal elements of this matrix reflect diffusion over space and time. We note that it is not possible to separate out the time from space and space-time diffusion effects in this model.

Of course, the $T-$horizon impulse response to a transitory change in the rth explanatory variable at time t would be given by the main- and off-diagonal elements of:

$$\begin{aligned} \partial\mathbf{Y}_{t+T}/\partial\mathbf{X}_t^{r'} &= \mathbf{D}_T[\mathbf{I}_N\beta_r + \mathbf{W}\gamma_r] \\ \mathbf{D}_T &= (-1)^T(\mathbf{B}^{-1}\mathbf{A})^T\mathbf{B}^{-1} \end{aligned} \tag{24}$$

We note that (24) also corresponds to the marginal effect in period $t + T$ of a permanent change in the rth explanatory variable in time t.

6. Since row-sums and columns-sums of our matrix are the same, we can do either. However, for interpretative purposes we follow [8] who note that the columns represent a partial derivative change arising from a change in a single region, whereas the rows reflect changes in all regions.

A special case of the model and associated effects estimates was considered by [9] where the restriction $-\phi\rho = \theta$ holds. This allows the matrix $\mathbf{Q}^{-1}$ to be expressed as:

$$\begin{aligned}
\mathbf{Q}^{-1} &= \begin{pmatrix} \mathbf{B}^{-1} & \mathbf{0} & \dots & & \mathbf{0} \\ \mathbf{D}_1 & \ddots & & & \vdots \\ \mathbf{D}_2 & \mathbf{D}_1 & \ddots & & \\ \vdots & & \ddots & \ddots & \mathbf{0} \\ \mathbf{D}_{T-1} & \mathbf{D}_{T-2} & \dots & \mathbf{D}_1 & \mathbf{B}^{-1} \end{pmatrix} \\
\mathbf{D}_s &= \phi^s \times \mathbf{B}^{-1}, s = 0, \dots, T-1
\end{aligned} \tag{25}$$

In this case, we have simple geometric decay over time periods of the spatial effects captured by the matrix $\mathbf{B}^{-1}$. The computationally efficient approach to calculating the effects for cross-sectional spatial regression models described in [8] can be used in conjunction with a scalar weighting term: $\phi^s, s = 0, \dots, T-1$.

In any application of the model it is possible to test if the restriction $-\phi\rho = \theta$ holds, which suggests that the sample data is consistent with a model based on space-time separability. We illustrate this in our application in the next section.

4 Application to state-level smoking behavior

We use a panel consisting of 45 (of the lower 48) states plus the District of Columbia covering 30 years from 1963-1992 taken from [4]. The model is a simple (logged) demand equation for (packs of) cigarettes as a function of the (logged) cigarette prices (per pack) and (logged) state-level income per capita. [7] We have observations for 30 years on (logged) real per capita sales of cigarettes measured in packs per person aged 14 years or older (the dependent variable). The two explanatory variables are the (logged) average retail price of a pack of cigarettes and (logged) real per capita disposable income in each state and time period.

Their motivation for spatial dependence (in their model disturbances) was a bootlegging effect where buyers of cigarettes near state borders purchase in

7. Colorado, North Carolina and Oregon are the three missing states.

Table 1: Dynamic space-time model parameter estimates

Parameters	lower 0.01	lower 0.05	mean	upper 0.95	upper 0.99
ϕ	0.8099	0.8125	0.8326	0.8554	0.8584
ρ	0.2825	0.2855	0.3040	0.3299	0.3333
θ	-0.2801	-0.2751	-0.2511	-0.2293	-0.2211
$-\phi\rho$	-0.2758	-0.2734	-0.2531	-0.2367	-0.2353
σ_μ^2	0.0006	0.0007	0.0011	0.0018	0.0021
σ_ε^2	0.0012	0.0012	0.0013	0.0014	0.0015
Variables	lower 0.01	lower 0.05	mean	upper 0.95	upper 0.99
price	-0.3540	-0.3406	-0.2982	-0.2555	-0.2432
income	0.0348	0.0500	0.0989	0.1479	0.1645
$W\times$ price	0.1216	0.1376	0.1862	0.2323	0.2464
$W\times$ income	-0.0883	-0.0717	-0.0206	0.0324	0.0492

neighboring states if there is a price advantage to doing so. They did not allow for time dependence in the model disturbances. [3] use a panel covering the period from 1963 to 1980 to estimate a non-spatial dynamic demand equation for cigarettes and find a significant negative price elasticity of -0.2 but no significant income elasticity. This model accounted for the bootlegging effect by incorporating the lowest price for cigarettes from neighboring states as an explanatory variable. Bootlegging was found to be statistically significant.

We report estimates for the model parameters in Table 1 based on 200,000 MCMC draws with the first 100,000 discarded to account for burn-in of the sampler. The table reports the posterior mean as well as lower 0.01 and 0.05 and upper 0.95 and 0.99 percentiles constructed using the retained draws. [8] Large variances were assigned to the prior distributions so these estimates should reflect mostly sample data information and be roughly equivalent to those from maximum likelihood estimation. [9]

The estimates for the parameters of the space-time filter indicate strong time dependence using the 0.01 and 0.99 intervals and weaker spatial dependence whose 0.01 and 0.99 intervals point to positive dependence. The cross

8. Every tenth draw from the 100,000 retained draws was used to construct the posterior estimates reported in the tables to reduce serial dependence in the sampled values.

9. This was checked and found to be the case.

product term θ is negative and the 0.01 and 0.99 intervals point to a difference from zero. We report the posterior distribution for the product $-\phi\rho$ constructed using the draws from the MCMC sampler. This distribution appears consistent with the restriction that can be used to simplify the model along with the effects estimates. It appears the sample data and model are consistent with space-time separability.

The coefficients associated with price, income and their spatial lags cannot be directly interpreted as if they were partial derivatives that measure the response of the dependent variable to changes in the regressors. As already shown, the partial derivatives take the form of $N \times N$ matrices for each time horizon and are non-linear functions of these coefficient estimates and the space-time filter parameters.

Table 2 shows the *direct effect* estimates for the contemporaneous time period out to a time horizon T of 29 years. [10]

Except for the first row of both panels that show pure feedbacks effects, these effects should capture mostly impacts arising from time dependence of region i on changes in its own explanatory variables plus some of the feedback loop (spatial) effects, which will be fed forward in time. Since all variables in the model have been log-transformed, we can interpret our direct, indirect and total effects estimates in elasticity terms. The table reports the posterior mean of the period-by-period effects along with credible intervals for these constructed from the MCMC draws. The column labeled 'Cumulative' shows the cumulation of these period-by-period effects that would reflect the time horizon $t + T$ response to a permanent change in the explanatory variables at time t. Since our estimates for the the space-time filter parameters are consistent with model stability (the sum of the spatial filter parameters being less than one), we will see the (period-by-period) direct effects die down to zero over time.

Consistent with microeconomic theory we see a greater long-run elasticity response of cigarettes sales to both price and income. [11] The direct effect period 0 price elasticity estimate of -0.29 is consistent with the estimate of -0.2 from [3]. The high level of time dependence in the estimate for ϕ leads

10. The general expressions in (22) were used to produce these effects estimates despite the fact that the space-time separability restriction appears consistent with the model and data. These expressions collapse (approximately) to the simpler expressions in (25) in this case.

11. Since it takes time for people to adjust behavior in response to price and income changes, the long-run elasticity is larger than short-run.

Table 2: Space-time direct effect estimates

	Price (elasticity)					
Horizon T	Cumulative	lower 0.01	lower 0.05	mean	upper 0.95	upper 0.99
0	-0.2898	-0.3432	-0.3302	-0.2898	-0.2494	-0.2378
1	-0.5311	-0.2821	-0.2725	-0.2412	-0.2096	-0.2001
2	-0.7318	-0.2341	-0.2259	-0.2007	-0.1755	-0.1675
3	-0.8990	-0.1959	-0.1888	-0.1671	-0.1461	-0.1394
4	-1.0381	-0.1651	-0.1586	-0.1391	-0.1209	-0.1152
5	-1.1541	-0.1400	-0.1340	-0.1159	-0.0996	-0.0948
6	-1.2507	-0.1192	-0.1135	-0.0966	-0.0817	-0.0777
7	-1.3312	-0.1017	-0.0964	-0.0805	-0.0669	-0.0635
8	-1.3983	-0.0870	-0.0821	-0.0671	-0.0547	-0.0518
9	-1.4543	-0.0745	-0.0699	-0.0559	-0.0446	-0.0422
10	-1.5010	-0.0638	-0.0597	-0.0466	-0.0364	-0.0344
15	-1.6412	-0.0298	-0.0272	-0.0189	-0.0131	-0.0122
20	-1.6982	-0.0141	-0.0125	-0.0077	-0.0047	-0.0043
25	-1.7216	-0.0067	-0.0058	-0.0031	-0.0016	-0.0015
29	-1.7299	-0.0037	-0.0031	-0.0015	-0.0007	-0.0006
	Income (elasticity)					
Horizon T	Cumulative	lower 0.01	lower 0.05	mean	upper 0.95	upper 0.99
0	0.0996	0.0391	0.0536	0.0996	0.1460	0.1618
1	0.1825	0.0327	0.0449	0.0829	0.1207	0.1334
2	0.2515	0.0275	0.0376	0.0689	0.1000	0.1104
3	0.3089	0.0230	0.0314	0.0574	0.0831	0.0916
4	0.3567	0.0192	0.0262	0.0477	0.0692	0.0763
5	0.3965	0.0161	0.0218	0.0397	0.0578	0.0638
6	0.4296	0.0134	0.0182	0.0331	0.0483	0.0533
7	0.4573	0.0112	0.0151	0.0276	0.0404	0.0447
8	0.4803	0.0093	0.0126	0.0230	0.0339	0.0376
9	0.4995	0.0078	0.0105	0.0191	0.0285	0.0317
10	0.5155	0.0065	0.0087	0.0160	0.0240	0.0268
15	0.5635	0.0025	0.0033	0.0064	0.0103	0.0118
20	0.5830	0.0009	0.0012	0.0026	0.0045	0.0053
25	0.5910	0.0003	0.0004	0.0010	0.0020	0.0024
29	0.5939	0.0001	0.0002	0.0005	0.0010	0.0013

to a much more responsive long-run price elasticity of -1.73. The would be close to the long-run value, since at a time horizon of 29 years, the period-by-period effects appear to have nearly died down to zero (the upper 0.99

interval value is -0.0006). Similarly for the income elasticity we see a period zero value of 0.10 and a thirty-year horizon value of 0.59, where again this is close to the long-run elasticity (since the lower 0.01 interval value is 0.0001 at the $T = 29$ horizon).

These results suggest that a 10 percent increase in (per pack) cigarette prices would lead to a short-run decrease in sales (of packs per capita) by 3 percent, but a long-run decrease in sales of 17.3 percent. Since the income elasticity is positive, increases in state-level per capita income leads to increased sales of cigarettes. [12] In the short-run a 10 percent increase in income leads to a 1 percent increase in cigarette sales, whereas in the long-run sales are more responsive showing a 6 percent increase.

Table 3 shows the *indirect effects* in a format identical to that of Table 2. These effects represent spatial spillovers plus contagion or diffusion that takes place over time. The magnitude of these effects is likely to be small since the estimate for the spatial dependence parameter ρ was small. Following [3], one motivation for the presence of spatial spillover and contagion effects is the bootlegging phenomena where buyers of cigarettes living near state borders purchase these at lower prices when possible.

The indirect effects for price are positive and different from zero up to a time horizon of 14 years using the 0.01 and 0.99 intervals. The positive sign is consistent with bootlegging since the (scalar summary) indirect effects estimates tell us that a positive change in own-state prices will lead to increased cigarette sales in neighboring states. Recall, this is cumulated over all first- second- and higher-order contiguous neighbors to each state and averaged to produce the scalar summary effects reported in the table. Since the marginal or period-by-period positive spillover effects die down to zero by year $T = 15$, where the cumulative effects take a value of 0.71, we can conclude that bootlegging serves to offset a substantial portion of the cumulative negative own-price elasticity effect of -1.64 that we see for year 15. A 10 percent increase in own-state cigarette prices would lead to a 7.1 percent increase in bootleg sales from neighboring states. The cumulative spillover/bootlegging impact is around 0.75 which in conjunction with the negative cumulative direct price impact of -1.72 suggests a long-run total impact from price changes that would be close to unit-elastic. This means a 10 percent increase in price would lead to a 10 percent decrease in cigarette sales. Ignoring the spatial spillover/bootlegging impact would lead to an over-

12. Economists label commodities having positive income elasticities *normal goods*.

Table 3: Space-time indirect effect estimates

	Price spillover (elasticity)					
Horizon T	Cumulative	lower 0.01	lower 0.05	mean	upper 0.95	upper 0.99
0	0.1290	0.0630	0.0786	0.1290	0.1759	0.1897
1	0.2361	0.0577	0.0694	0.1071	0.1437	0.1547
2	0.3250	0.0502	0.0595	0.0889	0.1181	0.1272
3	0.3988	0.0420	0.0495	0.0737	0.0977	0.1055
4	0.4600	0.0339	0.0405	0.0611	0.0818	0.0886
5	0.5107	0.0267	0.0325	0.0507	0.0694	0.0756
6	0.5528	0.0208	0.0257	0.0420	0.0596	0.0657
7	0.5877	0.0159	0.0201	0.0348	0.0516	0.0579
8	0.6166	0.0119	0.0155	0.0288	0.0450	0.0513
9	0.6405	0.0087	0.0118	0.0239	0.0394	0.0455
10	0.6603	0.0062	0.0089	0.0198	0.0346	0.0403
15	0.7186	-0.0000	0.0013	0.0077	0.0182	0.0214
20	0.7412	-0.0015	-0.0006	0.0029	0.0093	0.0111
25	0.7499	-0.0015	-0.0008	0.0011	0.0046	0.0056
29	0.7527	-0.0012	-0.0007	0.0005	0.0026	0.0032
	Income spillover (elasticity)					
Horizon T	Cumulative	lower 0.01	lower 0.05	mean	upper 0.95	upper 0.99
0	0.0127	-0.0540	-0.0379	0.0127	0.0660	0.0825
1	0.0235	-0.0435	-0.0302	0.0108	0.0533	0.0667
2	0.0327	-0.0351	-0.0243	0.0091	0.0434	0.0543
3	0.0405	-0.0285	-0.0197	0.0077	0.0358	0.0449
4	0.0471	-0.0235	-0.0161	0.0066	0.0299	0.0373
5	0.0528	-0.0195	-0.0134	0.0056	0.0251	0.0314
6	0.0577	-0.0164	-0.0112	0.0048	0.0213	0.0267
7	0.0619	-0.0141	-0.0095	0.0042	0.0181	0.0228
8	0.0656	-0.0121	-0.0082	0.0036	0.0155	0.0195
9	0.0687	-0.0105	-0.0071	0.0031	0.0134	0.0168
10	0.0714	-0.0092	-0.0062	0.0027	0.0116	0.0145
15	0.0804	-0.0049	-0.0033	0.0013	0.0059	0.0073
20	0.0848	-0.0027	-0.0017	0.0006	0.0031	0.0040
25	0.0871	-0.0014	-0.0009	0.0003	0.0017	0.0023
29	0.0880	-0.0008	-0.0005	0.0001	0.0010	0.0014

estimate of the sensitivity of sales to price changes.

Turning to the indirect effects for the income variable, these are small and not different from zero based on the 0.01 and 0.99 credible intervals. This suggests that increases in state-level income do not exert an influence on bootlegging behavior.

One point to note regarding our dynamic space-time model compared to models that deal with space and time dependence in the disturbances is that we have an explicit measure of spatiotemporal spillovers. The scalar summary effects estimates we propose here can be used to produce a quantitative assessment of the magnitude, timing and statistical significance of these spillovers.

A second point is that in the general space-time dynamic model considered here, the restriction $-\phi\rho = \theta$ is not imposed. This implies that except from the contemporaneous effects that represent pure spatial effects, future time horizons contain both time and space diffusion effects, which cannot be distinguished from each other. As noted by [9] when this restriction is consistent with the model and sample data, it is possible to separate out spatial, temporal and spatiotemporal impact magnitudes.

Table 4 reports the *total effects/impacts* estimates in a format identical to that used for Tables 2 and 3. These effects are the sum of the direct and indirect effects, so they reflect the long-run elasticity associated with the price and income variables from the broader perspective of society at large. Individual state leaders or policy makers would be interested in the direct effects on cigarette sales from changes in own-state prices and incomes. The bootlegging spillovers impacting individual states are likely to be small and of little consequence. However, from the broader perspective of national policy makers the (cumulative) total effects estimates would be the relevant estimates for national policy purposes.

The total effects for both price and income are different from zero at all 29 time horizons reported in the table. However, the marginal effects die down to nearly zero based on an examination of the 0.01 and 0.99 interval magnitudes for the horizon $T = 29$.

The negative direct effect (elasticity) of -1.72 from changes in price are offset somewhat by the positive effect of 0.75 on cigarette sales from bootlegging, leading to a total effect long-run elasticity of -0.977. Of course, this represents a much more elastic long-run relationship relative to the short-run elasticity of -0.16. A similar result occurs for the income elasticity where we see the short-run elasticity of 0.11 increased to 0.68 over time.

Table 4: Space-time total effect estimates

	Price total (elasticity)					
Horizon T	Cumulative	lower 0.01	lower 0.05	mean	upper 0.95	upper 0.99
0	-0.1608	-0.2181	-0.2048	-0.1608	-0.1286	-0.1196
1	-0.2949	-0.1691	-0.1614	-0.1340	-0.1103	-0.1040
2	-0.4067	-0.1343	-0.1301	-0.1118	-0.0943	-0.0895
3	-0.5001	-0.1129	-0.1076	-0.0933	-0.0801	-0.0764
4	-0.5781	-0.0962	-0.0905	-0.0779	-0.0673	-0.0644
5	-0.6433	-0.0820	-0.0767	-0.0651	-0.0554	-0.0529
6	-0.6978	-0.0700	-0.0656	-0.0545	-0.0444	-0.0424
7	-0.7435	-0.0601	-0.0563	-0.0456	-0.0350	-0.0333
8	-0.7817	-0.0517	-0.0485	-0.0382	-0.0274	-0.0261
9	-0.8137	-0.0446	-0.0418	-0.0320	-0.0214	-0.0203
10	-0.8406	-0.0387	-0.0361	-0.0268	-0.0168	-0.0158
15	-0.9225	-0.0198	-0.0175	-0.0112	-0.0049	-0.0043
20	-0.9570	-0.0103	-0.0086	-0.0047	-0.0014	-0.0012
25	-0.9717	-0.0054	-0.0043	-0.0020	-0.0004	-0.0003
29	-0.9771	-0.0032	-0.0024	-0.0010	-0.0001	-0.0001
	Income total (elasticity)					
Horizon T	Cumulative	lower 0.01	lower 0.05	mean	upper 0.95	upper 0.99
0	0.1124	0.0785	0.0860	0.1124	0.1490	0.1595
1	0.2061	0.0683	0.0737	0.0937	0.1172	0.1235
2	0.2843	0.0590	0.0630	0.0781	0.0944	0.0976
3	0.3495	0.0505	0.0535	0.0652	0.0778	0.0822
4	0.4039	0.0427	0.0453	0.0544	0.0652	0.0702
5	0.4494	0.0358	0.0378	0.0454	0.0550	0.0599
6	0.4874	0.0293	0.0310	0.0380	0.0466	0.0512
7	0.5193	0.0235	0.0248	0.0318	0.0398	0.0438
8	0.5459	0.0186	0.0195	0.0266	0.0342	0.0375
9	0.5682	0.0146	0.0153	0.0223	0.0294	0.0321
10	0.5869	0.0114	0.0120	0.0187	0.0253	0.0276
15	0.6440	0.0031	0.0035	0.0078	0.0122	0.0136
20	0.6679	0.0008	0.0010	0.0033	0.0059	0.0070
25	0.6781	0.0002	0.0003	0.0014	0.0029	0.0037
29	0.6819	0.0000	0.0001	0.0007	0.0016	0.0022

5 Conclusion

We have extended the approach taken by [8] for measuring own- and cross-partial derivative impacts that arise in (cross-sectional) spatial regression models to the case of dynamic space-time panel data models. They propose scalar summary measures along with measures of dispersion for these that allow the $N \times N$ matrices of impacts for each explanatory variable in the model to be summarized. Their approach is consistent with treatment of regression coefficient estimates where we view these as reflecting how changes in the explanatory variables impact the dependent variable on average over the sample. The extension results in a series of $N \times N$ matrix products for future horizons that can be cumulated to measure the dependent variable response over any time horizon. We follow [8] and produce scalar summary measures using averages of the main diagonal elements of the sequence of $N \times N$ matrices for direct or own-partial derivatives and averages of the cumulated off-diagonal elements for the cross-partials.

A re-examination of the 30 year space-time panel data set on state-level cigarette sales, prices and income from [3] demonstrated the usefulness of our dynamic space-time elasticities/responses. In particular, we are able to capture spillovers attributed to bootlegging as part of the model. We found that over the period 1963 to 1992 positive spatial spillovers attributed to bootlegging reduced the short-run price elasticity of sales response from -0.29 to -0.16, and the long-run price elasticity of sales response from -1.73 to around -0.98. Spatial spillovers played no significant role in affecting the income elasticity, which exhibited a short-run elasticity of 0.10 and a long-run elasticity of 0.60.

References

[1] Anselin, L. Spatial Econometrics. In: B.H. Baltagi (Ed.), A Companion to Theoretical Econometrics, Blackwell Publishers Ltd, Massashusetts, 2001, pp. 310-330.

[2] Baltagi, B.H. *Econometric Analysis of Panel Data*, 4th edition. John Wiley & Sons, Chichester, 2008.

[3] Baltagi, B.H. and D. Levin. Estimating dynamic demand for cigarettes using panel data: The effects of bootlegging, taxation and advertising reconsidered. The Review of Economics and Statistics 68 (1986) 148-155.

[4] Baltagi, B.H. and D. Li. Prediction in the panel data model with spatial correlation. In: L. Anselin, R.J.G.M. Florax, S.J. Rey (Eds.), Advances in Spatial Econometrics: Methodology, Tools and Applications, Berlin, 2004, pp. 283-295.

[5] Chib, S. and B.P. Carlin. On MCMC sampling in hierarchical longitudinal models, *Statistics and Computing* 9 (1999) 17-26.

[6] Kelejian, H.H and I.R. Prucha. Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. Journal of Econometrics 157 (2010) 53-67.

[7] Koop, G. *Bayesian Econometrics*, John Wiley & Sons, West Sussex, England, 2003.

[8] LeSage, J.P. and R. K. Pace. *Introduction to Spatial Econometrics*, Taylor & Francis/CRC Press, New York, 2009.

[9] Parent, O. and J. P. LeSage. A spatial dynamic panel model with random effects applied to commuting times. Transportation Research Part B: Methodological. 44B (2010a) 633-645.

[10] Parent, O. and J. P. LeSage. A space-time filter for panel data models containing random effects. Computational Statistics & Data Analysis. (2010b) forthcoming.

[11] Parent, O. and J. P. LeSage. Spatial Dynamic Panel Data Models With Random Effects. Working paper. (2010c) Available at SSRN: http://ssrn.com/abstract=1609748

[12] Sun, D. and J.O. Berger. Reference priors with partial information. Biometrika. 85 (1998) 55Ű71.

[13] Wansbeek, T.J. and A. Kapteyn. A simple way to obtain the spectral decomposition of variance components models for balanced data. Communications in Statistics Part A - Theory Methods. 11 (1982) 2105-2112.

[14] Yu, J., R. de Jong and L.F. Lee. Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and T are large. Journal of Econometrics. 146 (2008) 118-134.

[15] Zellner, A. *An introduction to Bayesian inference in econometrics*, John Wiley & Sons, New York, 1971.