RUHR
ECONOMIC PAPERS

James P. LeSage
Yao-Yu Chih
Colin Vance

Markov Chain Monte Carlo Estimation of Spatial Dynamic Panel Models for Large Samples

rwi

#769

Imprint

Ruhr Economic Papers

Published by

RWI – Leibniz-Institut für Wirtschaftsforschung
Hohenzollernstr. 1-3, 45128 Essen, Germany

Ruhr-Universität Bochum (RUB), Department of Economics
Universitätsstr. 150, 44801 Bochum, Germany

Technische Universität Dortmund, Department of Economic and Social Sciences
Vogelpothsweg 87, 44227 Dortmund, Germany

Universität Duisburg-Essen, Department of Economics
Universitätsstr. 12, 45117 Essen, Germany

Editors

Editorial Office

Ruhr Economic Papers #769

Responsible Editor: Manuel Frondel

ISSN 1864-4872 (online) – ISBN 978-3-86788-897-4

Ruhr Economic Papers #769

James P. LeSage, Yao-Yu Chih, and Colin Vance

Markov Chain Monte Carlo Estimation of Spatial Dynamic Panel Models for Large Samples

Bibliografische Informationen der Deutschen Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available on the Internet at http://dnb.dnb.de

RWI is funded by the Federal Government and the federal state of North Rhine-Westphalia.

http://dx.doi.org/10.4419/86788897
ISSN 1864-4872 (online)
ISBN 978-3-86788-897-4

James P. LeSage, Yao-Yu Chih, and Colin Vance[1]

Markov Chain Monte Carlo Estimation of Spatial Dynamic Panel Models for Large Samples

Abstract

Focus is on efficient estimation of a dynamic space-time panel data model that incorporates spatial dependence, temporal dependence, as well as space-time covariance and can be implemented in large N and T situations, where N is the number of spatial units and T the number of time periods. Quasi-maximum likelihood (QML) estimation in cases involving large N and T poses computational challenges because optimizing the (log) likelihood requires: 1) evaluating the log-determinant of an NT x NT matrix that appears in the likelihood, 2) imposing stability restrictions on parameters reflecting space-time dynamics, as well as 3) simulations to produce an empirical distribution of the partial derivatives used to interpret model estimates that require numerous inversions of large matrices. We set forth a Markov Chain Monte Carlo (MCMC) estimation procedure capable of handling large problems, which we illustrate using a sample of T = 487 daily fuel prices for N = 12, 435 German gas stations, resulting in N x T over 6 million. The procedure produces estimates equivalent to those from QML and has the additional advantage of producing a Monte Carlo integrated estimate of the log-marginal likelihood, useful for purposes of model comparison. Our MCMC estimation procedure uses: 1) a Taylor series approximation to the logdeterminant based on traces of matrix products calculated prior to MCMC sampling, 2) block sampling of the spatiotemporal parameters, which allows imposition of the stability restrictions, and 3) a Metropolis-Hastings guided Monte Carlo integration of the log-marginal likelihood. We also provide an efficient approach to simulations needed to produce the empirical distribution of the partial derivatives for model interpretation.

JEL Classification: C23, D40

Keywords: Dynamic panel models; spatial dependence; Markov Chain Monte Carlo estimation

September 2018

---

1 James P. LeSage, Texas State University, USA; Yao-Yu Chih, Texas State University, USA; Colin Vance, RWI and Jacobs University, Bremen. *– The authors would like to thank Olivier Parent for helpful comments on an earlier version of this paper. The authors also express their gratitude to Dr. Alexander Kihm, whose contributions made the assembly of the data possible. This research was supported by a grant from the Fritz Thyssen Foundation, grant number 10.17.2.026WW.* *– All correspondence to: Colin Vance, RWI, Hohenzollernstr. 1-3, 45128 Essen, Germany, e-mail: colin.vance@rwi-essen.de*

1 Introduction

Dynamic panel data models that accommodate spatial dependence have attracted attention in the spatial econometrics panel data literature. A variety of models that control for various types of correlation across locations have been explored. Su and Yang (2015) consider a dynamic model that models spatial dependence in the error terms, whereas Yu et al. (2008) introduce a dynamic panel specification that treats spatial dependence in the dependent variable vector. Yu et al. (2012) implement a specification that allows for spatial and time dependence as well as component mixing of space and time dependence that we label space-time covariance. This aspect of the model allows spatial diffusion to take place over time.

We focus on fast, efficient Markov Chain Monte Carlo (MCMC) estimation of these models in cases where the sample size is large in both the cross-sectional units N, and time periods T. The approach taken draws on previous work by Parent and LeSage (2010, 2011, 2012) who introduce a space-time filter view of these models. Since we have a large number of cross section (N) and time series (T), we follow Yu et al. (2008) and treat the data generating process as conditional on the initial period cross-sectional observations. For models that treat spatial dependence in the disturbances rather than the dependent variable, treatment of the initial period observations in small T situations has been found to be important (Su and Yang, 2015). Parent and LeSage (2010, 2011, 2012) provide Monte Carlo results showing that when the initial cross-sectional observations are endogenous, but incorrectly treated as exogenous, estimates and inferences can be biased, a finding similar to that of Su and Yang (2015) for the case where dependence is modeled in the disturbances. The Monte Carlo study from Parent and LeSage (2012) also show that when the number of time periods becomes large, treatment of the initial period observations becomes irrelevant, which is the case we address here.

A matrix version of the model specification on which we focus is in equation (1), with a more complete motivation and development of this specification taken up in section 2. The matrix W is an $N \times N$ spatial connectivity matrix with zeros on the main diagonal and non-zero elements in the i, jth position when observations i and j exhibit dependence. The $T \times T$ matrix L is the time lag operator such that $(L \otimes I_N)y = y_{t-1}$, and $(L \otimes W)y = Wy_{t-1}$. The disturbance vector ε is assumed to be normally distributed with constant scalar variance (σ^2) over time and space. Exogenous explanatory variables are represented by the matrix X with associated parameter vector β.

$$y = \rho(I_T \otimes W)y + \phi(L \otimes I_N)y + \theta(L \otimes W)y + X\beta + \varepsilon \tag{1}$$

Quasi-maximum likelihood (QML) estimation involves determining parameters $\rho, \phi, \theta, \beta, \sigma^2$ that minimize the negative of the log-likelihood function, subject to stability restrictions on the spatiotemporal parameters ρ, ϕ, θ.[1] In cases involving large N and T computational challenges arise because optimizing the (log) likelihood requires: 1) evaluating the log-determinant of an $NT \times NT$ matrix that appears in the likelihood, 2) imposing stability restrictions on parameters reflecting space-time dynamics which requires use of a constrained optimization algorithm, as well as 3) simulations to produce an empirical distribution of the partial derivatives used to interpret model estimates that require numerous inversions of large matrices.

We set forth a Markov Chain Monte Carlo (MCMC) estimation procedure capable of handling large problems, which uses: 1) a Taylor series approximation to the log-determinant based on traces of matrix products calculated prior to MCMC sampling that allows rapid evaluation of the log-determinant, 2) block sampling of the spatiotemporal parameters that allows imposition of the stability restrictions via rejection sampling, and 3) a Metropolis-Hastings guided Monte Carlo integration of the joint posterior distribution resulting in an estimate of the log-marginal likelihood. We also provide an efficient approach to simulations needed to produce the empirical distribution of the partial derivatives for model interpretation for a special case where the spatiotemporal parameters obey a restriction that has been labeled space-time separability in the literature (Parent and LeSage, 2011, 2012).

An illustration of the method uses a sample of $T = 487$ daily gas station prices for more than $N = 12,000$ stations in Germany (described in LeSage et al. (2017)), resulting

[1] We are assuming a fixed effects transformation has been applied to the model relationship in (1) to eliminate space and time fixed effects parameters, e.g., Lee and Yu (2010).

in $N \times T$ greater than 6 million. Fixed effects in spatial panel data models have been extensively analyzed for the case of large T by Yu et al. (2008), and for small T in Su and Yang (2015). In our application, we introduce fixed effects for the brand configuration of each station and its nearest neighboring station. Using price markups (over cost) as the dependent variable, this allows us to explore brand competition/cooperation between six different brands of stations.

Section 2 sets forth space-time filter expressions for the model and discusses computational issues that arise, along with our proposed solutions. Section 3 provides results from a Monte Carlo study that examines the accuracy and speed of our approach. Section 4 applies the estimation method to a price markup model of German gas stations involving the analysis of large sample.

2 A dynamic space-time panel data model

The model we explore is a space-time dynamic extension of the panel data model in (2), where α_i denote observation-specific and λ_t time-period specific fixed effects.

$$\begin{aligned} y_{it} &= x_{it}\beta + \alpha_i + \lambda_t + \varepsilon_{it} \\ & \quad i = 1, \dots, N,\ t = 1, \dots, T \end{aligned} \tag{2}$$

We consider the space-time filter of Parent and LeSage (2010, 2011, 2012) which can be applied to the dependent variable vector in the basic panel data model for $t = 1, \dots, T$, shown in (3).[2]

$$y_t = x_t\beta + \varepsilon_t \tag{3}$$

Where $y_t = (y_{1t}, \dots, y_{Nt})'$ is the $N \times 1$ vector of observations for the tth time period, x_t denotes the $N \times k$ matrix of non-stochastic regressors and the random terms ε_t are assumed

[2] For simplicity, we assume that a fixed effects transformation was used to eliminate time and spatial unit-specific fixed effects from the sample data y_t, x_t, e.g., Lee and Yu (2010).

to be independent and identically distributed with zero mean and a variance $\sigma_\varepsilon^2 I_N$.

Parent and LeSage (2010, 2011, 2012) suggest use of $C = (I_{T+1} - \phi L)$ as a time filter where ϕ is the time dependence coefficient and L is the $T \times T$ time-lag operator (matrix) such that

$$L = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & \ddots & & 0 \\ 0 & \dots & 1 & 0 \end{pmatrix}. \tag{4}$$

This time-dependence filter C is defined as:

$$C = \begin{pmatrix} \psi & 0 & \dots & 0 \\ -\phi & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & -\phi & 1 \end{pmatrix}. \tag{5}$$

Specification of ψ, the (1,1) element in C depends on whether the first period is modeled or assumed to be known. Since our focus is on large T samples, we do not model the first period.

The spatial dependence filter is defined as a nonsingular matrix $B = (I_N - \rho W)$, where ρ is a scalar spatial dependence parameter and W is a known, row-normalized $N \times N$ spatial weight matrix whose diagonal elements are zero. This matrix defines the dependence between cross sectional (spatial) observational units.

The space-time filter that Parent and LeSage (2010, 2011, 2012) proposed corresponds to the Kronecker product of the matrices C and B:

$$C \otimes B = I_{N,T} - \rho I_T \otimes W - \phi L \otimes I_N + (\rho \times \phi) L \otimes W, \tag{6}$$

where L is the $T \times T$ time-lag matrix. Parent and LeSage (2010) note that the filter implies a restriction on the parameter associated with spatial effects from the previous period $(L \otimes W)$, which should equal to $-\rho \times \phi$. However, they argue for relaxing this restriction and introducing an unrestricted parameter θ for the space-time covariance term in the model. Their argument is that imposing the restriction $\theta = -\rho \times \phi$ produces simplistic space-time impacts on the dependent variable y_t arising from changes in the explanatory variables x_t. Specifically, Debarsy et al. (2012) note that the restriction implies simplistic spatial spillovers decay over time according to $\phi^s B^{-1}$. Following LeSage and Pace (2009), the mean of diagonal elements of $(I_N - \rho W)^{-1} I_N \beta$, represent scalar summary measures of direct (own-partial derivative) effects, and the average of cumulative row-sums off-diagonal elements, spillover (cross-partial derivatives). Of course, having estimated the unrestricted specification involving the parameter θ, one can always test for the possible restriction that $\theta = -\rho\phi$ using posterior estimates, about which we will have more to say later. We can use the filter matrices to write our model in the matrix form shown in (1).

2.1 Computational challenges for estimation

The specification in (1) contains multiple dependence parameters to be estimated (ρ, ϕ, θ), which means that QML estimation requires a multivariate optimization routine to maximize the likelihood. Another aspect of the model is that the dependence parameters ρ, ϕ, θ associated with stable processes require $\rho + \phi + \theta < 1$, and for cases where $\rho - \theta > 0$, stability requires that $\phi - \rho + \theta > -1$ (Parent and LeSage, 2011). This means that constrained optimization must be used.

Another challenge to maximum likelihood estimation of this model is the log-determinant term that arises in the (log) likelihood function, specifically (log): $|I_{NT} - \rho(I_T \otimes W) - \phi(L \otimes I_N) - \theta(L \otimes W)|$. In the case of conventional spatial regression models involving a single weight matrix, there is a great deal of literature on approaches to efficiently calculating or approximating the log-determinant term that appears in the (log) likelihood ($|I_N - \rho W|$), (see LeSage and Pace (2009), Chapter 4). These approaches are not directly applicable to the model considered here, since the log-determinant expression needs to be evaluated for multiple dependence parameter values during each iteration of the optimization algorithm. In the case of Markov Chain Monte Carlo estimation, the log-determinant term appears in the conditional distribution for the dependence parameters requiring multiple evaluations, one for each trip through the MCMC sampler.

Another aspect of the model relates to proper interpretation of the partial derivative impacts on the dependent variable vector arising from changes in the explanatory variables, e.g., $\partial E(y)/\partial X^r$ for the rth explanatory variable. For the model in (1) we have: $\partial E(y)/\partial X^r = [I_{NT} - \hat{\rho}(I_T \otimes W) - \hat{\phi}(L \otimes I_N) - \hat{\theta}(L \otimes W)]^{-1} I_{NT}\hat{\beta}^r$, where $\hat{\beta}^r$ is the coefficient on the rth explanatory variable, (see Debarsy, Ertur and LeSage, 2011). Calculating the matrix inverse only once to determine point estimates of the partial derivative effects is not particularly challenging as the matrices W, L are sparse. However, conventional practice in maximum likelihood or MCMC estimation produces empirical estimates for measures of dispersion associated with the partial derivative effects estimates based on simulating the dependence parameters ρ, ϕ, θ as well as $\hat{\beta}$ from the (estimated) variance-covariance matrix involving either the analytical or numerical hessian, or using MCMC draws for these parameters (LeSage and Pace, 2009). A number of such simulated or MCMC draws of these parameters (say 1,000) are then evaluated in the partial derivative expressions to produce an empirical measure of dispersion, which requires numerous matrix inversions (say 1,000). This of course makes calculation of the empirical measures of dispersion for the partial derivative effects estimates on which inference is based a computationally challenging problem.

2.2 A computationally efficient representation of the model

The model in (1) can be written more compactly as in (7):

$$
\begin{aligned}
\tilde{y}\omega &= X\beta + \varepsilon \qquad (7)\\
\tilde{y} &= \begin{pmatrix} y & W_1 y & W_2 y & W_3 y \end{pmatrix}\\
W_1 &= I_T \otimes W,\ W_2 = L \otimes I_N,\ W_3 = L \otimes W\\
\omega &= \begin{pmatrix} 1 \\ -\rho \\ -\phi \\ -\theta \end{pmatrix} = \begin{pmatrix} 1 \\ -\Gamma \end{pmatrix},\ \Gamma = \begin{pmatrix} \rho \\ \phi \\ \theta \end{pmatrix}
\end{aligned}
$$

A key feature of $\tilde{y}$ is that this expression separates dependence parameters to be estimated from sample data describing the spatiotemporal dependence process, with the scalar spatiotemporal dependence parameters in the vector ω. MCMC estimation proceeds by sampling sequentially from the conditional distributions of each parameter (or set of parameters). The conditional distributions for the parameters $\beta, \sigma^2 I_N, \omega$ needed to implement MCMC estimation are set forth in the next section. We will see that a virtue of expressing the model as in (7) is that numerous quantities arise that involve sample data $\tilde{y}$, which can be calculated once prior to MCMC sampling.

2.3 Conditional distributions for parameters of the model

MCMC estimation involves sampling from the complete sequence of conditional distributions for the model parameters, which are set forth here. Since our focus is on large samples N, T, we can rely on uninformative priors for the parameters β, as these would not likely impact posterior estimates. For the same reason, we rely on an uninformative inverse Gamma$(\bar{a}, \bar{b})$, where we let $\bar{a}, \bar{b} \rightarrow 0$ for σ^2. Since the dependence parameters in ω are a focus of inference, we employ uniform priors for these dependence parameters, which must be constrained to lie in the stable region.

Given the limited prior information, the conditional distributions for the parameters β and σ^2 take the forms described in the following. The conditional distribution for the parameters β is multivariate normal with mean and variance-covariance shown in (8).

$$
\begin{aligned}
p(\beta|\sigma^2, \omega) &= \aleph(\tilde{\beta}, \tilde{\Sigma}_\beta) \\
\tilde{\beta} &= (X'X)^{-1}(X'\tilde{y}\omega) \\
\tilde{\Sigma}_\beta &= \sigma^2 (X'X)^{-1}
\end{aligned} \tag{8}
$$

We note that $(X'X)^{-1}X'\tilde{y}$ consists of only sample data information, so this expression can be calculated once prior to MCMC sampling, and this is true of $(X'X)^{-1}$ as well. This means that sampling new values of the parameters β (given values for the parameters σ^2, ω) can take place in a rapid, computationally efficient way.

The conditional posterior for σ^2 (given β, ω) takes the form in (9), when we set the prior parameters $\bar{a} = \bar{b} = 0$.

$$
\begin{aligned}
p(\sigma^2|\beta,\omega) &\propto (\sigma^2)^{-(\frac{NT}{2})}\exp\left(-\frac{1}{2\sigma^2}(e'e)\right) \quad (9)\\
e &= (\tilde{y}\omega - X\beta)\\
&\sim IG(\tilde{a},\tilde{b})\\
\tilde{a} &= NT/2\\
\tilde{b} &= (e'e)/2
\end{aligned}
$$

We sample the spatiotemporal parameters (ρ,ϕ,θ) as a block using Metropolis-Hastings. Block sampling means that a vector of the spatiotemporal dependence parameters in ω are proposed and compared to the current vector of spatiotemporal dependence parameters. The proposed vector is either accepted or rejected. This allows proposals of dependence parameters that obey the stability restriction, so any vectors that are accepted by the Metropolis-Hastings (M-H) procedure will always obey the needed restrictions. This requires evaluating the joint conditional distribution for the block of the spatiotemporal parameters (ρ,ϕ,θ). The stability constraint on the spatiotemporal parameters ρ,ϕ,θ requires that the sum of these parameters must be less than one, and in cases where where $\rho-\theta > 0$, stability requires that $\phi-\rho+\theta > -1$, Parent and LeSage (2011).

We rely on a reversible jump procedure to produce proposal values for the vector of parameters ρ,ϕ,θ. For each scalar parameter we rely on a three-headed coin flip. By this we mean a uniform random number on the open interval *coin flip* $= U(0,1)$, with head #1 equal to a value $\leq 1/3$, head #2 a value $> 1/3 \leq 2/3$ and head #3 a value $> 2/3 < 1$. Given a head #1 result, we set a proposal ρ^p using a uniform random draw on the open interval $(-1 < \rho^p < \rho^c)$, where ρ^c is the current value. A head #2 results in setting the proposal value equal to the current value $(\rho^p = \rho^c)$, while a head #3 selects a proposal value based on a uniform random draw on the open interval $(\rho^c < \rho^p < 1)$. Of course, a similar approach is used to produce proposals for the parameters ϕ,θ. Proposed vectors of these parameters inconsistent with the stability restrictions are eliminated via rejection sampling.

The reversal jump approach to proposing the block of spatiotemporal parameters has the virtue that accepted vectors will obey the stability restriction and will also reduce autocorrelation in the MCMC draws for these parameters. However, proposals from the reversible jump procedure based on the large intervals between $(-1 < \rho^c)$ and $(\rho^c < 1)$ will not produce candidates likely to be accepted when these parameters are estimated with a great deal of precision, as would be the case for problems involving large N, T. This can result in a failure to move the chain adequately over the parameter space. To address this issue, standard deviations, $\sigma_\rho, \sigma_\phi, \sigma_\theta$ for each parameter are calculated based on the first 1,000 draws (and thereafter using an interval of $m = 1,000$ draws). These are used in a tuned random-walk procedure to produce candidate/proposed values. Specifically, we use a tuning vector cc for each parameter that is adjusted based on acceptance rates for each parameter. This is used in conjunction with the standard deviations to produce proposals: $\rho^p = \rho^c + ccN(0,1)\sigma_\rho$, with the same approach used for ϕ_p, θ_p.

The block of proposed dependence parameters is then accepted or rejected using a Metropolis-Hastings step. We report results from a series of Monte Carlo experiments in Section 3 that allow us to assess the efficacy of this approach to sampling the parameters ρ, ϕ, θ.

2.4 The joint conditional for the spatiotemporal parameters of the model

The joint conditional distribution for the spatiotemporal dependence parameters in ω can be obtained by analytically integrating out β, σ^2 leading to a (log kernel) expression for the joint posterior of the dependence parameters in ω. For notational purposes, we express the sample data part of $\tilde{y}$ as: $\begin{bmatrix} y & W_1 y & W_2 y & W_3 y \end{bmatrix}$, where $W_1 = (I_T \otimes W), W_2 = (L \otimes I_N), W_3 = (L \otimes W)$.

$$
\begin{aligned}
log\ p(\omega|y, X, W_1, W_2, W_3) &\propto log[D(\omega)] - (NT/2)log(\omega' F \omega) \qquad (10)\\
F &= (\tilde{y} - X\beta_d)'(\tilde{y} - X\beta_d)\\
\beta_d &= (X'X)^{-1}X'\tilde{y}
\end{aligned}
$$

$log[D(\omega)]$ is a Taylor series approximation to the log-determinant term that arises in the model, described in detail in the next section. For now we note that this log-determinant term depends on the parameters in the vector ω, indicated by $D(\omega)$. We note that F consists of only sample data, so this expression can be calculated prior to MCMC sampling, leading to a computationally efficient expression reflecting a quadratic form: $log(\omega' F \omega)$, that can be easily evaluated for any vector of spatiotemporal dependence parameters ω.

2.5 A Monte Carlo estimate of the log-marginal likelihood

An advantage of analytically integrating out the parameters β, σ^2, is that further integration of the joint conditional posterior over the spatiotemporal dependence parameters in ω, yields the log-marginal likelihood for this model. We can use Monte Carlo integration to accomplish this task. Monte Carlo integration evaluates the expression to be integrated using random draws of the parameter values. A drawback to this approach is inefficiency because many of the random draws for the parameters are not in areas of high density of the function being integrated. In our case, the Metropolis-Hastings sampling procedure used to produce draws of the dependence parameters steers these parameter values to areas of high density of the joint posterior. This allows us to produce an efficient Monte Carlo integration of the log-marginal likelihood.

Given an estimate of the log-marginal likelihood for a model M_i ($LogM_i$), we can calculate: $prob(M_i) = LogM_i / \sum_{i=1}^{Q} LogM_i$ (in the case of Q different models). Of course, there is a great deal of interest in comparing alternative models, for example, models based on different spatial weight matrices. Given uncertainty regarding the appropriate weight matrix to use, a Bayesian solution to this problem is to produce estimates that average over models based on different W specifications, each weighted by their posterior model probability. This extends the model estimates to incorporate *uncertainty* regarding the correct weight matrix to use. Another use for model comparison regards the appropriate model specification. LeSage (2014) shows how to use $LogM$ in the case of single $W-$matrix models (cross-section and panel models) to compare models based on SAR, SDM, SDEM, SLX, SEM specifications, obtaining a straightforward answer to the question of which specification is most consistent with the sample data. The answer is unconditional on the parameter values because these have been integrated out to produce the log-marginal likelihood and associated model probabilities.

2.6 A Taylor's series approximation to the log-determinant

LeSage and Pace (2009) discuss Taylor's series as well as Chebyshev polynomial approximations for the log-determinant term that arises in spatial regression models. In our case, this takes the form shown in (11).

$$\begin{aligned} \ln|I_{N\times T} - \rho W_1 - \phi W_2 - \theta W_3| &= -\sum_{j=1}^{\infty} \frac{\Gamma^j \,\mathrm{tr}\, \tilde{W}^j}{j} \\ &\simeq -\sum_{j=1}^{q} \frac{\Gamma^j \mathrm{tr}(\tilde{W}^j)}{j} \\ \tilde{W} &= \begin{pmatrix} W_1 & W_2 & W_3 \end{pmatrix} \end{aligned} \tag{11}$$

From the definitions of W_1, W_2, W_3 we see that the first-order trace of $(\tilde{W}^1)$ for the matrix $W_1 = I_T \otimes W$ is zero because diagonal elements of the matrix W are zero. The same is true for the first-order trace of $W_2 = L \otimes I_N$ and $W_3 = L \otimes W$. Higher-order traces are shown in (12), where LeSage and Pace (2009) discuss computationally efficient ways to calculate these.

$$\begin{aligned} \mathrm{tr}(\tilde{W}^2) &= \sum_{i=1}^{3}\sum_{j=1}^{3} \mathrm{tr}(W_i W_j) \\ \mathrm{tr}(\tilde{W}^3) &= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3} \mathrm{tr}(W_i W_j W_k) \\ \mathrm{tr}(\tilde{W}^4) &= \sum_{i=1}^{3}\sum_{j=1}^{3}\sum_{k=1}^{3}\sum_{l=1}^{3} \mathrm{tr}(W_i W_j W_k W_l) \end{aligned} \tag{12}$$

In conclusion, for a set of pre-calculated traces, during MCMC sampling from the conditional (or joint posterior) distributions, we need only calculate powers of the parameters Γ^j and multiply these times the (already calculated) traces. A fourth-order Taylor's series approximation seems sufficient to produce accurate estimates. Specifically, for the case of our three matrices the second-order traces interacted with the parameter vector Γ takes the form in (13). Of course, the 3rd order terms increase to 27 and 4th order to 81 terms.

$$
\begin{aligned}
\Gamma^2 \operatorname{tr}(\tilde{W}^2) &= \Gamma' Q^2 \Gamma & (13)\\
\Gamma' &= \begin{pmatrix} \rho & \phi & \theta \end{pmatrix} \\
Q^2 &= \begin{pmatrix} \operatorname{tr}(W_1^2) & \operatorname{tr}(W_1 W_2) & \operatorname{tr}(W_1 W_3) \\ \operatorname{tr}(W_2 W_1) & \operatorname{tr}(W_2^2) & \operatorname{tr}(W_2 W_3) \\ \operatorname{tr}(W_3 W_1) & \operatorname{tr}(W_3 W_2) & \operatorname{tr}(W_3^2) \end{pmatrix} & (14)
\end{aligned}
$$

LeSage and Pace (2009) point out that accelerated computation of traces can be accomplished using sums of matrix Haddamard products. This involves some storage of previously calculated matrix products that avoids calculating these more than once. For the case of asymmetric matrices we use matrix products $\sum_i^3 \sum_j^3 W_i \odot W_j'$, and we note that row-normalized weight matrices typically used in spatial econometric models would be an example of asymmetric matrices.

2.7 Calculating effects estimates for model interpretation

LeSage and Pace (2009) point out that for the case of a cross-sectional SAR model involving N observations, partial derivatives take the form in (15) for the r^{th} explanatory variable. They propose scalar summary measures of the own- and cross-partial derivatives that they label *direct* and *indirect* effects, shown in (16) and (18), where ι_N is an $N \times 1$ vector of ones.

A scalar summary measure of the direct effect is constructed using the average of the main diagonal of the matrix $S_r(W)$, and a scalar summary indirect effect from the cumulative sum of off-diagonal elements in each row, averaged across rows.

$$\begin{aligned} \partial E(y)/\partial x^r &= (I_N - \rho W)^{-1}\beta^r = S_r(W) & (15) \\ &= I_N\beta_r + \rho W\beta_r + \rho^2 W^2\beta_r + \ldots \\ \bar{M}(r)_{direct} &= N^{-1}\,\mathrm{tr}(S_r(W)) & (16) \\ \bar{M}(r)_{total} &= N^{-1}\iota_N' S_r(W)\iota_N & (17) \\ \bar{M}(r)_{indirect} &= \bar{M}(r)_{total} - \bar{M}(r)_{direct} & (18) \end{aligned}$$

Expressions in (16), (17), (18) produce point estimates for the scalar summary measures of effects (own- and cross-partial derivatives) on dependent variable outcomes, but we also require a measure of dispersion for the purpose of statistical tests regarding the significance of these effects. Use of an empirical distribution constructed by simulating the non-linear expressions in (15) using (say 1,000) draws from the posterior distribution of the underlying parameters ρ, β_r is suggested by LeSage and Pace (2009), Note that a naive approach to such a simulation-based empirical distribution would require calculation of the $n \times n$ matrix inverse $(I_n - \rho W)^{-1}$ a large number of times, for varying values of the parameters ρ, which would be computationally intensive.

The required quantity for constructing the empirical distribution of the effects is $\mathrm{tr}(S_r(W))$, which can be estimated without a great deal of computational effort. For the purpose of calculating the effects estimates LeSage and Pace (2009) set forth a procedure that relies on a $(1 \times (q+1))$ vector R in (19) containing average diagonal elements from powers of W, and the $(1 \times (q+1))$ vector g in (20) and corresponding diagonal matrix G shown in (21), and $(q+1) \times 1$ vector of ones ι_{q+1}.[3]

[3] Computational implementations of this approach in the *Spatial Econometrics Toolbox* and the $R-$ language *Spdep* package set $q = 100$.

$$
\begin{aligned}
R &= \begin{pmatrix} 1 & 0 & \text{tr}(W^2)/N & \text{tr}(W^3)/N & \dots & \text{tr}(W^q)/N \end{pmatrix} && (19)\\
g &= \begin{pmatrix} 1 & \rho & \rho^2 & \dots & \rho^q \end{pmatrix} && (20)\\
G &= \begin{pmatrix} 1 & 0 & \dots & & 0 \\ 0 & \rho & 0 & \dots & 0 \\ 0 & 0 & \rho^2 & \dots & 0 \\ \vdots & & & \ddots & \\ 0 & & & \dots & \rho^q \end{pmatrix} && (21)\\
\bar{M}(r)_{direct} &= \beta^r RG\iota_{q+1} && (22)\\
\bar{M}(r)_{total} &= \beta^r g\iota_{q+1} && (23)\\
\bar{M}(r)_{indirect} &= \bar{M}(r)_{total} - \bar{M}(r)_{direct} && (24)
\end{aligned}
$$

Given the (pre-calculated) traces, empirical measures of dispersion for the effects can be constructed by using MCMC draws for the parameter ρ in g, G and β^r in expressions (22), (23), where we note that the total effects are the sum of the direct plus indirect effects.

In practice, computational implementations do not actually calculate traces but rather rely on estimates of these which do not require much computational effort. Specifically, starting with an $N \times m$ matrix v of *iid* normal deviates, where m is the number of vectors used in the estimation procedure (Barry and Pace, 1999),[4] set initial values $v_{(0)} = v$, and produce Monte Carlo estimates of the diagonals using:

$$
\begin{aligned}
v_{(j)} &= W v_{(j-1)} \\
\text{tr}(W^j) &\simeq (v \odot v_{(j)})\frac{\iota_m}{m} && (25)
\end{aligned}
$$

We can utilize aspects of this approach with some modifications to calculate effects estimates for our space-time dynamic panel data model. First note that the data generating process for our model can be expressed for each time period as shown in (26), where y_t

[4] We use $m = 150$, and $j = 100$.

represents an $N \times 1$ vector of observations at time period t, x_t an $N \times k$ matrix of explanatory variables with associated $k \times 1$ parameter vector β, and ε_t an $N \times 1$ vector of time t disturbances.

$$
\begin{aligned}
y_t &= Ay_{t-1} + (I_N - \rho W)^{-1}(x_t\beta + \varepsilon_t) \qquad (26) \\
A &= (I_N - \rho W)^{-1}(\phi I_N + \theta W)
\end{aligned}
$$

This formulation suggests that the current (period t) partial derivative impact of changes in the rth variable at time t, x_t^r on y_t take the form in (27). Since this is a dynamic model that includes time dependence as well as space-time covariance, changes in the explanatory variable at time t will have future impacts as well. The expression in (28) shows the one-period ahead impact, which depends on the time dependence parameter ϕ as well as the space-time covariance parameter θ, given the definition of the $N \times N$ matrix A.

$$
\begin{aligned}
\partial E(y_t)/\partial x_t^r &= (I_N - \rho W)^{-1}\beta^r \qquad (27) \\
\partial E(y_{t+1})/\partial x_t^r &= A[\partial E(y_t)/\partial x_t^r] \qquad (28) \\
\partial E(y_{t+T})/\partial x_t^r &= A[\partial E(y_t)/\partial x_{t+T-1}^r] = A^T[\partial E(y_t)/\partial x_t^r] \qquad (29)
\end{aligned}
$$

The current period effect in (27) can be calculated using the computationally efficient approach from LeSage and Pace (2009), which leads to direct (own-partial) impacts on own-region, and indirect (cross-partial, spatial spillover) impacts on neighboring regions. We can also produce empirical measures of dispersion using MCMC draws of the parameters β^r, ρ as described in the discussion surrounding equations (22), (23), (24).

One could use expressions like (29) to calculate $t + T$ period dynamic responses to changes in the rth explanatory variable at time t, using posterior means of the parameters ρ, ϕ, θ in the matrix A and $(I_N - \rho W)^{-1}$, as well as posterior means of β^r. These could take the form of impulse responses to a one-period change in the value of x_t^r, or a sustained change in the magnitude of x_t^r, $x_{t+1}^r, \ldots, x_{t+T}^r$. Impacts cumulated over time as well as marginal impacts (changes in each time period) could be considered. Calculating mean $t + T$ period dynamic responses would not be challenging, as this would involve inverting the $N \times N$ matrix A only once. Producing measures of dispersion for these responses would involve calculating the inverse of the $N \times N$ matrix A thousands of times for each set of MCMC draws for the parameters ρ, ϕ, θ.

In situations where posterior estimates of the parameters produce an indication that the restriction: $-\rho\phi = \theta$ is consistent with the sample data, dynamic period $t + T$ responses along with measures of dispersion can be calculated easily. In this restricted version of the model specification, the period $t + T$ response takes the form: $\phi^T(I_N - \rho W)^{-1}\beta^r$, and the long-run response is: $[1/(1 - \phi)](I_N - \rho W)^{-1}\beta^r$. This means that we can avoid the need to invert the $N \times N$ matrix A, and simply evaluate the long-run responses using draws for the parameters ϕ, ρ, β^r, where the computationally efficient approach to constructing $(I_N - \rho W)^{-1}\beta^r$ is used. Surprisingly, this restriction appears to hold for a great many space-time data samples. As Parent and LeSage (2010) point out, the restriction implies a type of space-time separability, where spatial effects decay over future time periods T according to the simple rate ϕ^T.

The matrix A needed for future period impacts might be of separate interest since it describes the one-period ahead impact arising from previous period changes in the dependent variable vector, $\partial E(y_{t+1})/\partial y_t = (I_N - \rho W)^{-1}(\phi I_N + \theta W)$. A change in period t values of y_t has both a direct and indirect impact on the period $t + 1$ outcomes, embodied in this $N \times N$ matrix. Given the structure of the model, changes in period t values of y_t also involve changes in neighboring region values, Wy_t, which gives rise to the term θW in this partial derivative. The main diagonal elements of the matrix A reflect own-region impacts arising from changes in previous period values of y_t, whereas the off-diagonal elements reflect other-region space-time diffusion impacts.

We can calculate scalar summary measures of these two types of impact, by extending the approach from LeSage and Pace (2009). This involves calculation of trace estimates shown in (30), in conjunction with the following expressions.

$$
\begin{aligned}
S &= \begin{pmatrix} 1 & 0 & \text{tr}(W^2)/N & \text{tr}(W^3)/N & \ldots & \text{tr}(W^q)/N \\ 0 & \text{tr}(W^2)/n & \text{tr}(W^3)/N & \text{tr}(W^4)/N & \ldots & \text{tr}(W^{q+1})/N \end{pmatrix} && (30) \\
g &= \begin{pmatrix} 1 & \rho & \rho^2 & \ldots & \rho^q \end{pmatrix} && (31) \\
G &= \begin{pmatrix} 1 & 0 & \ldots & & 0 \\ 0 & \rho & 0 & \ldots & 0 \\ 0 & 0 & \rho^2 & \ldots & 0 \\ \vdots & & & \ddots & \\ 0 & & & \ldots & \rho^q \end{pmatrix} && (32) \\
\bar{M}(A)_{direct} &= \begin{pmatrix} \phi & \theta \end{pmatrix} SG\iota_{q+1} && (33) \\
\bar{M}(A)_{total} &= (\phi + \theta) g\iota_{q+1} && (34) \\
\bar{M}(A)_{indirect} &= \bar{M}(A)_{total} - \bar{M}(A)_{direct} && (35)
\end{aligned}
$$

Here again, we would use estimated traces produced using the procedure from (25). Given pre-calculated trace estimates, empirical measures of dispersion for the (scalar summary) direct and indirect impacts based on MCMC draws of the parameters ρ, ϕ, θ in g and G, as well as ϕ, θ in expressions (33), (34) and (35) could be constructed.

3 Experiments to assess model performance

This section presents results regarding the time required to produce estimates for varying sample data sizes as well as results from a Monte Carlo experiment to assess the accuracy of our MCMC approach to estimation.

3.1 Timing results

Prior to reporting results from a Monte Carlo experiment to evaluate the performance of our sampling method, we report some timing tests for varying sample sizes and numbers of draws in Table 1. The number of explanatory variables used was fixed at four. The table reports total times required to produce estimates, including the MH-MC integration estimate for the log-marginal. These are for a Dell XPS 15 Laptop, with an Intel i7-7700HQ CPU 2.80GHz, 32 GB RAM and a 64-bit operating system. Times reported include that for calculating the traces and calculating effects estimates. The largest sample of $N = 5,000, T = 100$ involves working with a vector y equal to 500,000 and a matrix of size $500,000 \times 4$.

Table 2 provides a set of relative times, where the times reported in Table 1 were normalized by the $N = 1,000, T = 10$, number of draws = 10,000 times to facilitate analysis of how changes in N, T and the number draws impacts calculation times.

From the columns of the Table 2, we see that the time required is slightly less than linear in the number of draws. For example, tripling the number of draws from 10,000 to 30,000 increases the time in seconds required to produce estimates by 2.71 in the case of $N = 1,000$ and by $66.2266/23.8677 = 2.86$ in the worst case of $N = 5,000$.

Looking across the rows of Table 2, we see that a ten-fold increase in the number of time periods T from 10 to 100 results in around a 5.4 increase in time required for $N = 1,000$, but performance declines for $N = 2,000$ to around an 18 times increase and for $N = 5,000$ to around a 25-fold increase in time required.

Finally, comparing sample sizes of $N = 1,000$ to $N = 2,000$ and $N = 5,000$, (holding T and the # of draws constant), we see that a five-fold increase in the sample size along this dimension leads to a smaller than three-fold increase in time required to produce estimates.

Table 1: Times (in seconds) required to produce estimates

	$N = 1,000, T = 10$	$N = 1,000, T = 20$	$N = 1,000, T = 50$	$N = 1,000, T = 100$
10,000 draws	4.7040	6.1720	11.8610	25.7680
20,000 draws	8.5950	12.3480	22.0030	48.2400
30,000 draws	12.7830	17.4790	32.2610	70.8830
	$N = 2,000, T = 10$	$N = 2,000, T = 20$	$N = 2,000, T = 50$	$N = 2,000, T = 100$
10,000 draws	6.4560	9.9230	26.4850	110.5290
20,000 draws	11.8430	18.1870	48.7020	213.6800
30,000 draws	17.4240	26.1980	72.2170	318.0660
	$N = 5,000, T = 10$	$N = 5,000, T = 20$	$N = 5,000, T = 50$	$N = 5,000, T = 100$
10,000 draws	12.4230	26.2530	139.7670	296.5090
20,000 draws	22.6120	49.2050	271.7590	555.8490
30,000 draws	32.0970	72.5950	395.6430	822.7330

Table 2: Relative times for estimates

	$N=1,000, T=10$	$N=1,000, T=20$	$N=1,000, T=50$	$N=1,000, T=100$
10,000 draws	1.0000	1.3121	2.5215	5.4779
20,000 draws	1.8272	2.6250	4.6775	10.2551
30,000 draws	2.7175	3.7158	6.8582	15.0687
	$N=2,000, T=10$	$N=2,000, T=20$	$N=2,000, T=50$	$N=2,000, T=100$
10,000 draws	1.0000	1.5370	4.1024	17.1204
20,000 draws	1.8344	2.8171	7.5437	33.0979
30,000 draws	2.6989	4.0579	11.1860	49.2667
	$N=5,000, T=10$	$N=5,000, T=20$	$N=5,000, T=50$	$N=5,000, T=100$
10,000 draws	1.0000	2.1133	11.2507	23.8677
20,000 draws	1.8202	3.9608	21.8755	44.7435
30,000 draws	2.5837	5.8436	31.8476	66.2266

3.2 Monte Carlo experiments

For our Monte Carlo experiments we generated the dependent variable using: $y = (I_{NT} - \rho W_1 - \phi W_2 - \theta W_3)^{-1}(X\beta + \varepsilon)$, using a range of ρ values 0.2, 0.4, 0.6, ϕ values of 0.5, 0.7 and θ values of -0.3, -0.5. The spatial weight matrix used was a six nearest neighbor matrix, produced using random normal latitude-longitude coordinates. The $N \times T \times K$ matrix X consisted of standard normal deviates, with the values of the parameters $\beta' = \begin{pmatrix} 1 & -1 & 1 & -1 \end{pmatrix}$, which produced a stationary dependent variable vector centered on zero. Noise variances of $\sigma^2 = 1, 10$ were used, and sample sizes of $N = 1,000, 2,000$ with $T = 10, 50$ were used. We analyzed results separately for the 24 cases where $N = 1,000, T = 10$ and the 24 cases of $N = 2,000$ with $T = 50$. A set of 6,000 MCMC draws were used with the first 1,000 discarded for burn-in of the sampler.

Of the 24 cases two cases were included that reflect parameter values of $\rho + \phi + \theta = 1$, on the boundary of the stability restriction that: $\rho + \phi + \theta < 1$, specifically, $\rho = 0.6 + \phi = 0.7 + \theta = -0.3 = 1$, with $\sigma^2 = 1, 10$. This was done to explore performance of the MCMC sampler when we encounter situations where rejections of the proposed draws occurs because they violate the stability restriction. This would result in the posterior distribution being truncated at the boundary of the viable parameter space.

Table 3 shows results based on 1,000 Monte Carlo trials, for the various combinations of space-time dependence parameters and the two noise variance parameters used. The table shows actual and estimated values for the sum of the space-time parameters $((\rho + \phi + \theta))$ to conserve on space.[5] The absolute bias, mean-squared error (MSE) and 95% coverage results reported represent averages over the three dependence parameters. The 95% coverage results reflect the proportion of times (out of 1,000 trials) that the posterior 2.5% and 97.5% intervals constructed from the 5,000 retained MCMC draws, covered the true parameter value. The last column of the table shows an R^2 estimate (averaged over the 1,000 trials) based on $\hat{y} = (I_n - \hat{\rho}W_1 - \hat{\phi}W_2 - \hat{\theta}W_3)^{-1}X\hat{\beta}$, where the estimates used were posterior means calculated from the retained 5,000 MCMC draws.[6] This provides an indication of the relative signal-to-noise for the generated data used in the Monte Carlo experiments.

Table 3: Monte Carlo results ($n = 1,000, T = 10$), Space-time parameters

Space-time parameters ($\rho + \phi + \theta$)	True	Estimated	$\sum_{i=1}^{3}$ abs(Bias)	$\sum_{i=1}^{3}$ MSE	Mean (95% coverage)	R^2
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =1	0.4	0.4000	0.000994	0.000171	0.942333	0.847
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =1	0.2	0.1999	0.000343	0.000154	0.960667	0.852
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =1	0.6	0.5995	0.000643	0.000173	0.938000	0.888
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =1	0.4	0.4006	0.000625	0.000146	0.950667	0.895
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =1	0.6	0.6008	0.002300	0.000133	0.945667	0.858
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =1	0.4	0.4012	0.001501	0.000131	0.946667	0.857
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =1	0.8	0.8004	0.003539	0.000129	0.942333	0.898
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =1	0.6	0.6007	0.001830	0.000123	0.939667	0.893
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =1	0.8	0.8026	0.008604	0.000135	0.895333	0.892
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =1	0.6	0.6044	0.006196	0.000116	0.908667	0.881
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =1	0.8	0.8024	0.009286	0.000129	0.893333	0.955
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =10	0.4	0.3994	0.001213	0.000529	0.948333	0.912
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =10	0.2	0.2004	0.000492	0.000564	0.941667	0.455
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =10	0.6	0.6001	0.000127	0.000503	0.949000	0.472
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =10	0.4	0.3999	0.000597	0.000507	0.943000	0.602
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =10	0.6	0.6018	0.006450	0.000462	0.938667	0.625
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =10	0.4	0.4027	0.005652	0.000430	0.942000	0.493
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =10	0.8	0.8005	0.006917	0.000401	0.942333	0.491
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =10	0.6	0.6015	0.006743	0.000406	0.935667	0.637
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =10	0.8	0.8080	0.025760	0.000677	0.800333	0.618
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =10	0.6	0.6140	0.021136	0.000589	0.810333	0.613
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =10	0.8	0.8076	0.027739	0.000676	0.783333	0.577
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =1 †	1.0	0.9971	0.244113	0.028345	0.795000	0.687
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =10 †	1.0	0.9951	0.184305	0.020593	0.913667	0.687

† Cases on the boundary of $\rho + \phi + \theta < 1$

From the table we see good results in terms of bias and MSE, for cases not involving the two sets of parameters on the boundary of the parameter space (indicated in the table

[5]Appendix A shows results for all parameters for the larger sample of $N = 2,000, T = 40$.

[6]Specifically, $R^2 = \hat{y}'\hat{y}/\tilde{y}'\tilde{y}$, where $\tilde{y} = y - \bar{y}$, which is the usual R^2 formula from ordinary regression adjusted to accommodate the predicted value from the space-time dynamic model.

with the † symbol. For these two cases, we see a degradation in the 95% coverage as well as a large increase in MSE. We see a few other cases (the 3rd, 4th and 5th to the last row in the table), where absolute bias increases and coverage declines. These were associated with situations where the signal-to-noise was relatively low and the sum of the space-time parameters equalled 0.6 or 0.8. We might expect to see this type of result arise due to errors made by the 4th-order Taylor series approximation to the log-determinant term. This will become more inaccurate for values of the space-time parameters at the boundaries of the stationary region. Nonetheless, coverage intervals around 80% for the space-time parameters are relatively decent, since these are not the main focus of inference in these models.

Table 4 shows bias, MSE and coverage results for the current period direct effects estimates, which are a focus of inference in these models. Recall, these estimates are based on the average of the diagonal from the matrix: $(I_N - \rho W)^{-1} I_N \beta_i, i = 1, \ldots, 4$. For each of the 1,000 trials a set of 5,000 estimates of these four sets of effects were produced based on the retained draws. For each trial, posterior means and credible intervals were constructed for the scalar summary measures using the set of 5,000 estimates. Since the direct effects estimates are non-linear functions of the underlying model parameters ρ, β_i, we should not produce effects estimates based on posterior mean estimates of ρ, β_i, a point made by LeSage and Pace (2018). This approach to calculating bias, MSE and coverage for the direct effects estimates should accurately reflect results experienced by practitioners.

Given the four different direct effects estimates that sum to zero, we report the sum of absolute bias and MSE for these four different effects estimates along with coverage results averaged over the four sets of estimates. From the table, we see good estimates based on all three measures, and the 95% coverage results do not appear to be severly impacted for the two cases at the boundary of the space-time parameter space.

Table 5 shows bias, MSE and coverage results for the current period indirect effects estimates, which are also focus of inference in these models. These estimates are based on the average of cumulated row sums of off-diagonal elements from the matrix: $(I_N - \rho W)^{-1} I_N \beta_i, i = 1, \ldots, 4$. As in the case of the direct effects, scalar summary estimates of the indirect effects were calculated for each of the 1,000 trials based on the set of 5,000

Table 4: Monte Carlo results ($n = 1,000, T = 10$), direct effects parameters

Current period $(I_N - \rho W)^{-1}\beta_i$ Direct effects parameters	$\sum_{i=1}^4$ abs(Bias)	$\sum_{i=1}^4$ MSE	Mean (95% coverage)
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =1	0.001071	0.000407	0.9425
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =1	0.001382	0.000405	0.9475
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =1	0.000769	0.000406	0.9507
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =1	0.001016	0.000405	0.9480
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =1	0.001550	0.000426	0.9465
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =1	0.001177	0.000442	0.9430
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =1	0.001277	0.000431	0.9490
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =1	0.002172	0.000421	0.9532
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =1	0.007497	0.000504	0.9475
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =1	0.006519	0.000472	0.9545
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =1	0.008307	0.000499	0.9515
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =10	0.004002	0.004035	0.9487
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =10	0.001742	0.004091	0.9517
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =10	0.004505	0.004019	0.9490
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =10	0.003658	0.003971	0.9532
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =10	0.004872	0.004251	0.9490
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =10	0.003439	0.004321	0.9490
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =10	0.004758	0.004329	0.9482
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =10	0.002233	0.004208	0.9532
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =10	0.021542	0.004968	0.9485
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =10	0.024125	0.004953	0.9452
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =10	0.022947	0.004839	0.9482
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =1 †	0.041975	0.001050	0.9267
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =10 †	0.022515	0.004885	0.9607

† Cases on the boundary of $\rho + \phi + \theta < 1$

estimates of these four sets of effects.

Given the four different indirect effects estimates, we report the sum of absolute bias and MSE for these four different effects estimates along with coverage results averaged over the four sets of estimates. From the table, we see that the two cases on the boundary of the parameter space where the space-time parameters suffered from reduced 95% coverage accuracy had an impact on coverage of the current period indirect effects. This is also true for the few other cases (the 3rd, 4th and 5th to the last row in the table) noted in our discussion of Table 3, where bias increased and coverage declined for the space-time parameter estimates.

Table 6 shows bias, MSE and coverage results for the direct effects estimates associated with the long-run (LR) multiplier matrix $A = (I_N - \rho W)^{-1}(\phi I_N + \theta W)$, which might also be a focus of inference in these models. These estimates are based on the average of the

Table 5: Monte Carlo results ($n = 1,000, T = 10$), indirect effects parameters

Current period $(I_N - \rho W)^{-1}\beta_i$ Indirect effects parameters	$\sum_{i=1}^{4}$ abs(Bias)	$\sum_{i=1}^{4}$ MSE	Mean (95% coverage)
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =1	0.003086	0.000649	0.935500
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =1	0.000369	0.000598	0.948000
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =1	0.001309	0.000669	0.931000
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =1	0.001642	0.000578	0.944250
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =1	0.015486	0.001398	0.948250
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =1	0.015241	0.001450	0.943750
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =1	0.018897	0.001490	0.945250
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =1	0.013503	0.001445	0.945250
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =1	0.130151	0.008779	0.845750
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =1	0.119633	0.007982	0.862250
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =1	0.136403	0.008989	0.837500
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =10	0.003102	0.001824	0.942000
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =10	0.002526	0.001857	0.943750
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =10	0.003609	0.001878	0.945000
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =10	0.002108	0.001820	0.948250
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =10	0.045130	0.005709	0.940750
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =10	0.042904	0.005476	0.940250
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =10	0.042217	0.005436	0.946000
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =10	0.046498	0.005852	0.930500
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =10	0.401235	0.063220	0.705250
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =10	0.395035	0.062077	0.712000
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =10	0.405367	0.063024	0.708000
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =1 †	0.900585	0.280634	0.820750
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =10 †	0.528925	0.155829	0.946000

† Cases on the boundary of $\rho + \phi + \theta < 1$

diagonal elements from the matrix A. As in the case of the current period direct effects, scalar summary estimates of these LR direct effects were calculated for each of the 1,000 trials based on the set of 5,000 estimates of the single set of effects.

We report the bias and MSE for these LR direct effects estimates along with coverage results. From the table, we see that the two cases on the boundary of the parameter space where the space-time parameters suffered from reduced 95% coverage accuracy had an impact on coverage of the LR direct effects. These results are consistent with those for the current period direct effects estimates in Table 4.

Table 7 shows bias, MSE and coverage results for the indirect effects estimates associated with the long-run (LR) multiplier matrix $A = (I_N - \rho W)^{-1}(\phi I_N + \theta W)$, which might also be a focus of inference in these models. Recall, these estimates are based on the average of cumulated row sums of off-diagonal elements from the matrix A. As in the case of the

Table 6: Monte Carlo results ($n = 1,000, T = 10$), A matrix direct effects parameter

$A = (I_N - \rho W)^{-1}(\phi I_N + \theta W)$	True	Estimate	Bias	MSE	(95% coverage)
Direct effects parameters					
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =1	0.493024	0.493187	0.000163	0.000000	0.941000
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =1	0.486048	0.486083	0.000035	0.000000	0.961000
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =1	0.694419	0.694280	-0.000139	0.000000	0.946000
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =1	0.687443	0.687666	0.000223	0.000000	0.944000
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =1	0.491747	0.491864	0.000117	0.000000	0.943000
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =1	0.475241	0.475041	-0.000200	0.000000	0.964000
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =1	0.698349	0.698567	0.000217	0.000000	0.945000
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =1	0.681844	0.681919	0.000076	0.000000	0.941000
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =1	0.500000	0.499953	-0.000047	0.000000	0.949000
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =1	0.468146	0.468014	-0.000132	0.000000	0.950000
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =1	0.687258	0.687158	-0.000100	0.000000	0.960000
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =10	0.493024	0.493149	0.000125	0.000000	0.949000
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =10	0.486048	0.486190	0.000142	0.000000	0.951000
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =10	0.694419	0.694421	0.000002	0.000000	0.960000
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =10	0.687443	0.687236	-0.000207	0.000000	0.942000
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =10	0.491747	0.491805	0.000058	0.000000	0.947000
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =10	0.475241	0.475412	0.000170	0.000000	0.955000
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =10	0.698349	0.698420	0.000071	0.000000	0.946000
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =10	0.681844	0.681389	-0.000455	0.000000	0.950000
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =10	0.500000	0.500016	0.000016	0.000000	0.958000
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =10	0.468146	0.468487	0.000341	0.000000	0.952000
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =10	0.687258	0.687631	0.000373	0.000000	0.954000
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =1 †	0.719112	0.654505	-0.064608	0.004734	0.762000
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =10 †	0.719112	0.663227	-0.055885	0.004073	0.849000

† Cases on the boundary of $\rho + \phi + \theta < 1$

current period indirect effects, scalar summary estimates of these LR indirect effects were calculated for each of the 1,000 trials based on the set of 5,000 estimates of the single set of effects.

It is not surprising that we see increased bias and MSE for the two cases on the boundary. More surprising is that coverage of the estimates are very good for the non-boundary cases, and not bad for the two boundary cases. An explanation for this type of result is that the matrix A includes all of the space time parameters, making it possible for poor results (bias) in any single parameter to be offset or compensated by bias in the opposite direct for the other space-time dependence parameters.

Table 8 shows results for the noise variance parameter σ^2, in the same format as the previous tables. We see results consistent with those from Table 3 where the two cases on the boundary of the space-time parameter space exhibit increased bias and MSE, and

Table 7: Monte Carlo results ($n = 1,000, T = 10$), A matrix indirect effects parameter

$A = (I_N - \rho W)^{-1}(\phi I_N + \theta W)$	True	Estimate	Bias	MSE	(95% coverage)
Indirect effects parameters					
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =1	-0.243024	-0.243438	-0.000414	0.000000	0.958000
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =1	-0.486048	-0.485939	0.000109	0.000000	0.964000
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =1	-0.194419	-0.194573	-0.000154	0.000000	0.951000
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =1	-0.437443	-0.437071	0.000372	0.000000	0.960000
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =1	-0.158414	-0.158757	-0.000343	0.000000	0.951000
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =1	-0.475241	-0.475295	-0.000054	0.000000	0.946000
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =1	-0.031683	-0.032164	-0.000481	0.000000	0.960000
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =1	-0.348510	-0.348694	-0.000184	0.000000	0.951000
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =1	0.000000	-0.000444	-0.000444	0.000000	0.953000
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =1	-0.468146	-0.470023	-0.001877	0.000004	0.961000
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =1	-0.187258	-0.188292	-0.001034	0.000001	0.955000
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =10	-0.243024	-0.244030	-0.001006	0.000003	0.952000
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =10	-0.486048	-0.485748	0.000300	0.000002	0.925000
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =10	-0.194419	-0.194355	0.000064	0.000001	0.944000
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =10	-0.437443	-0.437271	0.000172	0.000001	0.940000
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =10	-0.158414	-0.159836	-0.001423	0.000004	0.938000
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =10	-0.475241	-0.477168	-0.001927	0.000006	0.945000
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =10	-0.031683	-0.032879	-0.001196	0.000002	0.944000
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =10	-0.348510	-0.350186	-0.001676	0.000004	0.948000
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =10	0.000000	-0.000853	-0.000853	0.000002	0.964000
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =10	-0.468146	-0.474764	-0.006618	0.000046	0.964000
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =10	-0.187258	-0.189893	-0.002635	0.000008	0.956000
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =1 †	0.280888	0.339081	0.058193	0.003921	0.877000
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =10 †	0.280888	0.325009	0.044121	0.002824	0.947000

† Cases on the boundary of $\rho + \phi + \theta < 1$

decreased coverage. There are lesser problems for cases shown in the (the 3rd, 4th and 5th to the last row in the table), where we see an increase in bias and MSE, but only slightly degraded coverage.

We also carried out Monte Carlo experiments for $N = 2,000, T = 50$, which produce lower bias, MSE and better coverage consistent with our knowledge about asymptotic performance of these estimates (see Yu et al. (2008)). Tables of results from those experiments are presented in the Appendix, where all parameters are shown instead of the summaries presented here.

Table 8: Monte Carlo results ($n = 1,000, T = 10$), σ^2 parameter

$A = (I_N - \rho W)^{-1}(\phi I_N + \theta W)$	True	Estimate	Bias	MSE	(95% coverage)
σ^2 parameters					
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =1	1	1.000489	0.000489	0.000203	0.957000
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =1	1	1.000238	0.000238	0.000204	0.956000
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =1	1	0.999983	-0.000017	0.000197	0.946000
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =1	1	1.000666	0.000666	0.000203	0.940000
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =1	1	0.998873	-0.001127	0.000208	0.943000
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =1	1	1.000390	0.000390	0.000200	0.956000
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =1	1	1.000131	0.000131	0.000202	0.945000
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =1	1	0.999255	-0.000745	0.000201	0.956000
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =1	1	0.998563	-0.001437	0.000204	0.952000
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =1	1	0.997882	-0.002118	0.000196	0.953000
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =1	1	0.999161	-0.000839	0.000206	0.943000
ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =10	10	10.005732	0.005732	0.020375	0.947000
ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =10	10	10.011446	0.011446	0.018679	0.959000
ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =10	10	9.998280	-0.001720	0.018571	0.957000
ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =10	10	10.006694	0.006694	0.020538	0.950000
ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =10	10	9.995011	-0.004989	0.019573	0.948000
ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =10	10	9.997400	-0.002600	0.019889	0.961000
ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =10	10	10.002006	0.002006	0.021833	0.940000
ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =10	10	9.988898	-0.011102	0.019691	0.958000
ρ =0.6, ϕ =0.5, θ =-0.3, σ^2 =10	10	9.962699	-0.037301	0.022270	0.946000
ρ =0.6, ϕ =0.5, θ =-0.5, σ^2 =10	10	9.952259	-0.047741	0.022429	0.931000
ρ =0.6, ϕ =0.7, θ =-0.5, σ^2 =10	10	9.951466	-0.048534	0.022895	0.933000
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =1 †	1	1.168741	0.168741	0.036968	0.877000
ρ =0.6, ϕ =0.7, θ =-0.3, σ^2 =10 †	10	10.483834	0.483834	0.384246	0.938000

† Cases on the boundary of $\rho + \phi + \theta < 1$

4 Space-time analysis of German gas station pricing

A 2012 report by the International Energy Agency[7] describes Germany as having a largely deregulated and competitive oil market, with a large number of independents in the refining and retail sectors. The German government does not have an ownership stake in any of the companies operating in the oil sector. The report also indicates that in 2010 the largest refining operator in Germany was Shell Deutschland Oil (Shell brand name) with a 25.6% share of overall German refining capacity. Next largest was BP Europa (Aral brand name) (14.5%), ConocoPhillips Germany (Jet brand name) (13.9%) and Total Deutschland (Total brand name) (11.8%). Next in terms of refining capacity share at 9.8% was (Esso brand name) owned by Rosneft via the joint venture Ruhr Oel (with BP). This leaves a remaining share of 35% for other brands, which we label Noname.

[7] Oil & Gas Security Emergency Response of IEA Countries, which we reference as IEA (2012).

In terms of the retail fuel sector, the IEA (2012) report, (page 9) states that there are more than 14,300 roadside filling stations in Germany, and another 350 filling stations on the autobahns. Aral and Shell have the highest market shares (22.5% and 21% of fuel sales respectively), followed by Jet with 10.5% and Total and Esso with 7.5% each. This leaves a share of 41% for numerous other refinery companies, and independent and medium-sized oil companies are active in the retail fuel market, including Avia, Westfalen and Freie Tankstellen (bft), all of which we simply label Noname.

The IEA (2012, page 9) report also states that: "Company market shares in the German retail fuel sector have remained relatively steady over the past few years. However, the number of filling stations in Germany is declining, with 475 fewer filling stations than at beginning of 2006 and approximately 1600 fewer than at beginning of 2001." and that: "Demand for diesel increased by around 16% between 2001 and 2011 while demand for gasoline dropped by nearly 30% during the same period." (IEA (2012), page 10).

To investigate competition in Germany's retail gasoline market, Kihm et al. (2016) use a panel of daily price data from 2012 to 2013, and find that the effects of price varies by brands. They show that Total and Shell have the highest deviations from cost-based pricing, while the non-majors have the lowest. In addition, they identify some factors, such as the absence of nearby competitors and regional market concentration, that play a significant role in mediating the influence of the oil price.

In this vein, LeSage et al. (2017) describe a balanced panel of daily station-level prices created for over 14,000 filling stations in Germany, operating over the period from June 1, 2014 to September 30, 2015, or $T = 487$ days. Price for both diesel and e5 fuels are in nominal terms and include excise and value-added taxes. They also provide a cost variable, constructed from daily refined diesel and gas prices reported in Rotterdam, where one of the major pipelines into Germany originates. By applying a static heterogeneous coefficient spatial autoregressive panel model to produce station-level estimates of gas station rivalry, they show that non-competitive pricing prevails in Germany's retail gasoline market, which is in contrast to the IEA (2012) report but is consistent with a report issued by Germany's Federal Cartel Office finding evidence for oligopolistic pricing among Germany's retail gas stations.

While both Kihm et al. (2016) and LeSage et al. (2017) consider the degree of competition in the vicinity of the station, they do not further examine the competition/non-competition behavior across different neighboring brands. Other things equal, proximity to stations with dominating brand name should lead to brand competition on price, while proximity to the non-major brand stations might be expected to reduce the tension of brand competition. For a better understanding of the role played by brand competition in Germany's retail gasoline market, we analyze interaction between different brands based on spatial proximity of each station to those of own- and other-brands. In the case of our six brands we require a large sample of stations to have an adequate sample needed to produce a six by six matrix of interaction between stations of each brand with the nearest neighboring own- or other-brand station. We use a large sample in conjunction with the dynamic space time model to explore the role played by brand configuration of stations in close spatial proximity on price markup behavior in the German retail gasoline market.

4.1 German gas station markup pricing model

We model the dependent variable as the daily before-tax markup in price over refining cost for the sample of 12,050 stations selling $e5$ (unleaded) fuel and 12,435 stations selling diesel fuel, that have a nearest neighboring station within 3 miles (radial distance). With 487 days, this produces a vector y of length $NT = 5,868,350$ for $e5$ and $NT = 6,055,845$ in the case of diesel fuel. As explanatory variables we include a set of 36 indicator variables for the brand configuration. This involves an indicator for brand i of the dependent variable station and brand j of the nearest neighboring station.[8] These can be interpreted as fixed effects that capture the impact on price markup associated with the various brand configurations. In addition, we include a variable vector for the distance to the nearest neighboring station. The dependent variable is transformed to deviations from time period means to allow use of all 36 brand indicators, which of course sum to one.

We can view the brand configuration matrix that we explore as in Table 9, where the number of stations of type brand i versus brand j are presented. The main diagonal elements

[8] Relying only on the nearest neighboring station allows us to isolate the brand configuration impact, whereas use of more than a single neighboring station would be problematical if the brands of the two nearest stations were different.

of the matrix reflect cases where the nearest neighboring station is of the same brand type. Off diagonal elements i, j show cases where the price markup dependent variable represents that of a station of brand type i and the nearest neighboring station is of type brand j.

Table 9: Brand configuration of $e5$ and diesel fuel stations

Number of $e5$ fuel stations						
Dependent (market share)/independent	Aral	Shell	Jet	Total	Esso	Noname
Aral (22.5%)	216	199	139	325	149	1014
Shell (21%)	201	43	57	135	61	439
Jet (10.5%)	118	48	15	92	28	272
Total (7.5%)	343	141	113	148	97	799
Esso (7.5%)	162	57	33	102	27	311
Noname (41.0%)	1078	487	314	784	295	3208
Number of diesel fuel stations						
Dependent (market share)/independent	Aral	Shell	Jet	Total	Esso	Noname
Aral (22.5%)	212	200	136	322	148	1043
Shell (21%)	201	41	57	137	61	447
Jet (10.5%)	117	49	15	89	28	278
Total (7.5%)	338	145	111	150	96	817
Esso (7.5%)	159	57	33	98	27	320
Noname (41.0%)	1107	502	320	811	299	3464

Table 9 also shows a degree of asymmetry in the brand configuration relationships between stations. For instance, in the e5 fuels portion of the table, we see that 199 Aral brand stations have a Shell brand station as the nearest neighboring station. However, there are 201 Shell brand stations with an Aral station as the nearest neighbor. Given that our spatial weight matrix W is based on the nearest neighboring station, a *symmetric* spatial configuration (spatial weight matrix) arising from situations where stations i and j and stations j and i were nearest neighbors for all stations would have implications for the nature of spatial spillovers. A strictly symmetric matrix W implies a block diagonal weight matrix, and as noted, $\partial E(y_t)/\partial x_t^r = (I_N - \rho W)^{-1}\beta^r$. This expression would also be block diagonal, implying that indirect effects for all stations would extend only to the single nearest neighboring stations and have no spillover impact on other stations. From the counts of cases in Table 9, we see that strict symmetry is not the case. That is, there are *asymmetric* situations where station i is the nearest neighbor to station j, while station j is not the nearest neighbor to station i. Asymmetry in the spatial proximity of stations indicated by Table 9 means that the pricing decision of a station can affect not only the markups of its nearest neighboring station, but also that of higher order neighboring stations. To see this, note that $(I_N - \rho W)^{-1} = I_N + \rho W + \rho^2 W^2 + \rho^3 W^3 \ldots$, where W^2 contains non-zero elements that identify neighbors to neighbors, and W^3 has non-zero elements reflecting neighbors to the neighbors of the neighbors, and so on. In the case of an *asymmetric* matrix W, we will see diffusion to higher-order neighbors reflected in non-zero off-diagonal elements of the matrix inverse, with a decline in magnitude of these impacts because $\rho < 1$, so $\rho^3 < \rho^2 < \rho$.

The model estimated is of the form shown in (36), where the only explanatory variables are the set of $N \times 36$ (fixed effects) brand indicators (B) and the $N \times 1$ vector D of distances to the nearest neighboring station.

$$y = \rho(I_T \otimes W)y + \phi(L \otimes I_N)y + \theta(L \otimes W)y + \iota_T \otimes B\beta + \iota_T \otimes D\psi + \varepsilon \qquad (36)$$

We would expect that increasing distance to the nearest neighboring station has a positive impact on the fuel price markup, as this reflects a spatial monopoly situation. In terms of the brand configuration impacts, we would expect the own-brand nearest neighbors would have a positive impact on price markup in cases where prices are coordinated by a company main office or formal pricing policy. In the case of Noname brand stations, we might expect to see competition between these independent stations, leading to a negative impact on price markup. For cases involving brand i versus brand j, we might expect positive impacts on price markup for the large market share brands if there is a price leadership scheme in play. Negative impacts for brand i versus brand j of course point to competition that leads to a decrease in the price markup.

It turns out that for both diesel and $e5$ fuels, the restriction that $-\rho\phi = \theta$ is consistent with the sample data. This allows us to easily calculate not only the current period (daily) impact of the brand configuration on price markups, but also the long-run implications of the various brand configurations on the price markup. As noted, given an asymmetric weight matrix, our dynamic space-time panel model implies that current period impacts on price markups will spillover to neighboring stations in the current period, and will diffuse over time to have impacts on stations that are neighbors to the neighboring stations, neighbors to the neighbors, and so on. The indirect effects estimates *cumulate* spatial spillover impacts over not simply the nearest neighbors, but other higher-order neighbors as well.

The period $t + T$ impacts arising from a brand i versus brand j station configuration in this restricted version of the model takes the form: $\phi^T \times (I_N - \rho W)I_N\beta^{(i,j)}$, where $\beta^{(i,j)}$ is the coefficient associated with a station of brand i having a nearest neighboring station of brand j. Of course, as in all regression models, the coefficients reflect an average over all sample observations, in our case NT. The long-run impacts can also be calculated as: $[1/(1-\phi)] \times (I_N - \rho W)I_N\beta^{(i,j)}$. Empirical measures of dispersion for these long-run impacts can easily be calculated using the MCMC draws for ϕ, ρ and $\beta^{(i,j)}$.

Table 10: Space-time and distance parameter estimates for fuel stations

			*e*5 fuel stations		
Parameter	Posterior mode	mean	MC error	Geweke	
distance (ψ)	0.7596	0.7588	0.00074328	0.995550	
ρ	0.3891	0.3891	0.00008079	0.996896	
ϕ	0.8701	0.8701	0.00002530	0.999658	
θ	-0.3309	-0.3308	0.00010400	0.995410	
	lower 0.01	lower 0.05	median	upper 0.95	upper 0.99
distance (ψ)	0.5899	0.6281	0.7592	0.8891	0.9257
ρ	0.3865	0.3875	0.3891	0.3896	0.3897
ϕ	0.8692	0.8697	0.8701	0.8705	0.8706
θ	-0.3317	-0.3315	-0.3309	-0.3297	-0.3276
$-\rho\phi$			-0.3385		
			diesel fuel stations		
Parameter	Posterior mode	mean	MC error	Geweke	
distance (ψ)	0.9763	0.9751	0.00094484	0.993080	
ρ	0.3340	0.3338	0.00020675	0.991698	
ϕ	0.8853	0.8853	0.00004910	0.999358	
θ	-0.2945	-0.2943	0.00025001	0.988810	
	lower 0.01	lower 0.05	median	upper 0.95	upper 0.99
distance (ψ)	0.7950	0.8431	0.9752	1.1092	1.1439
ρ	0.3230	0.3328	0.3340	0.3345	0.3347
ϕ	0.8820	0.8847	0.8853	0.8857	0.8858
θ	-0.2953	-0.2952	-0.2946	-0.2934	-0.2792
$-\rho\phi$			-0.2956		

Table 10 shows posterior mean, median estimates based on MCMC draws for the space-time and distance parameters, along with an estimate of the mode taken from the joint posterior. Monte Carlo error estimates and Geweke's diagnostic for convergence are also shown.[9] As we would expect given the large sample size, these parameters are estimated to

[9]This statistic compares draws from the first 10 percent of the MCMC sampling (after burn-in) and the

have very little dispersion, based on our empirical percentiles constructed from the MCMC draws.

From the table, we can see that the value of $-\rho\phi$ lies within the lower 0.05 and upper 0.95 percentiles of the empirical distribution measuring dispersion of the unconstrained parameter estimate for θ in the case of the $e5$ fuel estimates. The same is not quite true for the diesel fuel estimates, where we see that $-\rho\phi$ lies just outside the lower 0.01 percentile of the empirical distribution for θ. Nonetheless, we will use the restricted model approach to calculate long-run effects on price markup arising from the various brand configurations.[10]

Table 10 also shows that distance parameter (ψ) is positive and statistically significant. The finding is consistent with Barron et al. (2004); Clemenz and Gugler (2006); Jaureguiberry (2010); Kihm et al. (2016) such that the presence of spatial isolation reduces price competition among brands and increases the ability of markups. The analysis suggests that, after controlling for brand characteristics, consumers will pay an additional 0.76 cents for an additional 1 mile between two retail gasoline stations. Indeed, when consumers are less aware of alternative choices because of higher search cost, the station's local market power increases.

Table 11 shows posterior means for the brand configuration parameter β estimates, all of which are significantly different from zero. In the context of our (non-linear) space-time dynamic model, these are not interpretable as partial derivatives showing the impact of changes in the brand configuration on the price markup (as in standard linear regression). As already noted, the partial derivatives showing the impact on price markups from changes in the brand configuration for our model take the form of own-partial and cross-partial derivatives that capture own-station impacts as well as feedback and spillover impacts on other stations.

The long-run estimates of direct impacts arising from the various brand configurations are shown in Table 12. These include feedback effects that arise in our model. Feedback effects are from neighboring stations reactions as well as reactions of neighbors to the neighboring stations, etc. Note that the matrices W^2, W^3 that arise in $(I_N - \rho W)^{-1}$ contain

last 50 percent of the draws. The test is whether the batched means are equal, which indicates convergence.

[10] Given MC error for the empirical distribution of the parameter θ, and the non-linear nature of the restriction $-\rho\phi$, we cannot conclude that we should reject the restriction.

Table 11: Brand configuration β parameter estimates (posterior means)

β estimates for $e5$ fuel stations						
Dependent/independent	Aral	Shell	Jet	Total	Esso	Noname
Aral	0.1436	0.2909	0.2059	0.2152	0.2116	0.3447
Shell	-0.1075	-0.0080	0.0759	-0.0197	-0.0292	0.0701
Jet	-0.2691	-0.2068	-0.1366	-0.2857	-0.2624	-0.1303
Total	0.1033	0.1761	0.3109	0.1139	0.1980	0.3227
Esso	0.0181	0.0686	0.1907	0.0558	0.1382	0.2177
Noname	-0.2692	-0.1792	-0.1116	-0.2627	-0.2221	-0.0935
β estimates for diesel fuel stations						
Dependent/independent	Aral	Shell	Jet	Total	Esso	Noname
Aral	0.1891	0.3022	0.1781	0.2497	0.2290	0.3293
Shell	-0.0373	0.0252	0.0884	0.0343	0.0070	0.0807
Jet	-0.2154	-0.1645	-0.1260	-0.2275	-0.2183	-0.1121
Total	0.1748	0.2252	0.3289	0.1669	0.2458	0.3358
Esso	0.0690	0.0842	0.1809	0.0981	0.1851	0.2080
Noname	-0.1984	-0.1415	-0.0954	-0.1984	-0.1729	-0.0788

non-zero diagonal elements, reflecting the fact that station i is a neighbor to its neighboring station j, which accounts for feedback. Impacts from brand i station on neighboring brand j station will feedback to influence the price markup of the brand i station, and this will in turn have additional spillover impacts on neighboring stations, neighbors to the neighboring stations, and so on.

The estimates in each column of Table 12 can be interpreted as indicating how the price markup of station i (in each row) is impacted by having the column branded station as its nearest neighbor. The diagonal elements show the impact of having a same-branded station as the nearest neighbor. For example the 1,1 element for the case of $e5$ fuel in the table shows that an Aral brand station with an Aral brand nearest neighboring station will enjoy around 1.24 (Euro) cents higher price markup. To put this in context, the average price markup for $e5$ was 78 cents and 60 cents for diesel over all stations and time periods. The first row of the $e5$ portion of the table shows that Aral stations (the market share leader) enjoy a positive markup irrespective of what other brand station is the nearest neighbor. Also, the first row shows Aral station markups are highest (2.97 cents) when the independent/Noname brand station is the nearest neighbor, (e.g., relative to other elements in row #1). This observation is consistent across all brands. Specifically, the markups are the highest for all brands when next to the independent/Nonme brand, with one exception, that Shell station markups are the highest when next to Jet (0.65 cents), followed by Noname brand (0.60 cents). In sum, stations next to Noname and Jet stations, the latter of which is marketed as a discount brand in Germany, exhibit the largest deviations from cost-based pricing.

In addition, in comparison with other brands, the first row of the $e5$ portion of the table shows that Aral stations have the highest price markups irrespective of which other brand stations are the nearest neighbor, while the third and last rows show that Jet and Noname stations have lowest price markups. The results from Table 12 imply that the brand effect plays a significant role on not only own-station but also other-station (neighboring) markups. Since Noname stations are not affiliated with a recognizable brand name, they appear to be less desirable in the eye of consumers, accounting for the relatively lower markups.[11] It also means that stations in close proximity to Noname stations can exercise more local market power and price discrimination, since consumers are less willing to switch to the Noname stations. In contrast, proximity of brand name stations to other brand name stations (notably Aral and Total), appears to decrease own-station market power, perhaps because customers view these as close substitutes, leading to a more competitive oil market as stated in IEA (2012).

Turning to the diesel fuel markup results, we see similar brand configuration results regarding diesel pricing strategies. Variation in markups are explained by both own- and nearest neighboring station brand names, with stations located next to Noname able to maintain higher markups. Notice that market demand for diesel fuel has been expanding, but the average markup is lower than for $e5$ at 60 cents. Furthermore, a great deal of the similarity in long-run effects is explained by the similarity of the space-time parameter estimates. We should also note that most of the stations in our sample of 12,050 stations selling $e5$ (unleaded) fuel and 12,435 stations selling diesel fuel are the *same* stations that sell both types of fuel. We might expect similar pricing policies/strategies for both types of fuel sold.

As motivated in the face of asymmetric proximity of brand configurations for neighboring stations, we can have indirect (spillover) effects on pricing decisions of higher-order neighboring stations. Indirect effects are shown in Table 13, where we see generally smaller

[11] Jaureguiberry (2010) argues that the variations of markups depend upon many factors than just gasoline. Gasoline sold at different station/brandname is heterogeneous product in the eye of consumers.

Table 12: Brand configuration long-run direct effect estimates (posterior means)

Long-run direct effects for $e5$ fuel stations						
Dependent/independent	Aral	Shell	Jet	Total	Esso	Noname
Aral	1.2382	2.5044	1.7743	1.8535	1.8233	2.9699
Shell	-0.9260	-0.0695	0.6538	-0.1705	-0.2544	0.6045
Jet	-2.3172	-1.7813	-1.1779	-2.4621	-2.2579	-1.1235
Total	0.8896	1.5177	2.6760	0.9807	1.7060	2.7795
Esso	0.1564	0.5902	1.6427	0.4812	1.1928	1.8753
Noname	-2.3193	-1.5434	-0.9616	-2.2631	-1.9141	-0.8057
Long-run direct effects for diesel fuel stations						
Dependent/independent	Aral	Shell	Jet	Total	Esso	Noname
Aral	1.7812	2.8466	1.6769	2.3509	2.1560	3.1010
Shell	-0.3512	0.2373	0.8328	0.3224	0.0651	0.7607
Jet	-2.0280	-1.5484	-1.1862	-2.1422	-2.0546	-1.0559
Total	1.6461	2.1207	3.0975	1.5712	2.3152	3.1613
Esso	0.6493	0.7933	1.7048	0.9249	1.7429	1.9594
Noname	-1.8687	-1.3334	-0.8989	-1.8680	-1.6289	-0.7424

indirect effects estimates than direct effects estimates for both $e5$ and diesel fuel. Recall that these represent cumulative spillovers over not just immediately neighboring stations but higher-order neighbors as well. The smaller magnitudes of (cumulative) impact suggests that the spatial scope of retail fuel pricing competition is limited to lower-order neighbors, which seems intuitively plausible. Smaller spillovers might also reflect the fact that a great deal of symmetry exists with regard to stations, where stations i and j are nearest neighbors and stations j and i are neighbors. The counts of brands in Table 9 imply a great deal of symmetry in this regard.

The long-run total effects are reported in Table 14, which provide a nice summary of the overall impacts associated with the various brand configurations. The range of $e5$ price markup impacts from different brand configurations is from -3.62 to 4.03 cents, or more than 7 Euro cents. For diesel fuel the range is -2.97 to 4.38 cents, both of these reflecting about 10 percent of the average markup.

The impact on $e5$ price markups of Aral (the market share leader) range from 1.82 to 4.37 cents, with the lowest markup for cases where another Aral station is the nearest neighbor and the highest being in cases where a Noname/independent station is the nearest neighbor. One of the lowest markups arises in the case of a Noname station with Aral as its nearest neighbor (-3.41 cents), and another in cases where a Jet station has an Aral station

Table 13: Brand configuration long-run indirect (spillover) effect estimates (posterior means)

Long-run indirect effects for $e5$ fuel stations						
Dependent/independent	Aral	Shell	Jet	Total	Esso	Noname
Aral	0.5839	1.1817	0.8368	0.8745	0.8599	1.4015
Shell	-0.4367	-0.0327	0.3081	-0.0802	-0.1197	0.2851
Jet	-1.0931	-0.8400	-0.5548	-1.1612	-1.0647	-0.5298
Total	0.4195	0.7158	1.2626	0.4624	0.8047	1.3115
Esso	0.0737	0.2778	0.7745	0.2264	0.5619	0.8849
Noname	-1.0943	-0.7283	-0.4534	-1.0679	-0.9030	-0.3801
Long-run indirect effects for diesel fuel stations						
Dependent/independent	Aral	Shell	Jet	Total	Esso	Noname
Aral	0.6901	1.1030	0.6500	0.9112	0.8353	1.2020
Shell	-0.1361	0.0920	0.3230	0.1250	0.0252	0.2948
Jet	-0.7859	-0.5999	-0.4596	-0.8304	-0.7961	-0.4091
Total	0.6381	0.8219	1.2003	0.6088	0.8971	1.2253
Esso	0.2518	0.3075	0.6607	0.3585	0.6755	0.7596
Noname	-0.7244	-0.5167	-0.3483	-0.7241	-0.6313	-0.2877

as the nearest neighbor, (also, -3.41 cents). Interestingly, a Jet station with a Total brand station as it neighbor also produces the lowest markup (-3.62 cents), and these two brands have lower market shares (10.5 and 7.5 percent respectively).

Table 14: Brand configuration long-run total effect estimates (posterior means)

Long-run total effects for $e5$ fuel stations						
Dependent/independent	Aral	Shell	Jet	Total	Esso	Noname
Aral	1.8216	3.6859	2.6102	2.7277	2.6826	4.3715
Shell	-1.3628	-0.1024	0.9619	-0.2507	-0.3742	0.8894
Jet	-3.4100	-2.6214	-1.7324	-3.6233	-3.3227	-1.6529
Total	1.3091	2.2332	3.9382	1.4431	2.5099	4.0910
Esso	0.2301	0.8676	2.4172	0.7075	1.7548	2.7600
Noname	-3.4135	-2.2717	-1.4149	-3.3309	-2.8169	-1.1856
Long-run total effects for diesel fuel stations						
Dependent/independent	Aral	Shell	Jet	Total	Esso	Noname
Aral	2.4713	3.9495	2.3269	3.2621	2.9913	4.3031
Shell	-0.4873	0.3292	1.1558	0.4474	0.0903	1.0555
Jet	-2.8139	-2.1483	-1.6458	-2.9725	-2.8507	-1.4650
Total	2.2842	2.9425	4.2978	2.1800	3.2123	4.3866
Esso	0.9010	1.1008	2.3655	1.2835	2.4184	2.7190
Noname	-2.5931	-1.8501	-1.2472	-2.5921	-2.2602	-1.0302

5 Conclusion

A computationally efficient approach to producing MCMC estimates for space-time dynamic panel models involving large N and T was set forth. In addition to providing the ability to handle large problems, the method also produces Metropolis-Hastings tuned Monte Carlo estimates of the log-marginal likelihood, which allow formal Bayesian model comparison of alternative specifications based on differing weight matrices. Monte Carlo experiments show the MCMC estimation method produces estimates with good coverage, low bias and mean-squared error.

An illustration applied the method to 487 daily fuel (both unleaded and diesel) prices for over 12,000 German gas stations, where $N \times T$ is over 6 million. The focus of the application was on pricing competition/cooperation between six different branded stations. The results reveal some evidence for market power resulting from the spatial configuration of branded and unbranded stations.

References

Barron, J. M., B. A. Taylor, and J. R. Umbeck (2004). Number of sellers, average prices, and price dispersion. *International Journal of Industrial Organization 22*(8-9), 1041–1066.

Barry, R. P. and R. K. Pace (1999). Monte carlo estimates of the log determinant of large sparse matrices. *Linear Algebra and its applications 289*(1-3), 41–54.

Clemenz, G. and K. Gugler (2006). Locational choice and price competition: some empirical results for the austrian retail gasoline market. *Empirical Economics 31*(2), 291–312.

Debarsy, N., C. Ertur, and J. P. LeSage (2012). Interpreting dynamic space–time panel data models. *Statistical Methodology 9*(1-2), 158–171.

Jaureguiberry, F. (2010). *An analysis of strategic price setting in retail gasoline markets*. The Pardee RAND Graduate School.

Kihm, A., N. Ritter, and C. Vance (2016). Is the german retail gasoline market competitive? a spatial-temporal analysis using quantile regression. *Land Economics 92*(4), 718–736.

Lee, L.-f. and J. Yu (2010). A spatial dynamic panel data model with both time and individual fixed effects. *Econometric Theory 26*(2), 564–597.

LeSage, J. and R. K. Pace (2009). *Introduction to spatial econometrics*. Chapman and Hall/CRC.

LeSage, J. P. and R. K. Pace (2018). Spatial econometric monte carlo studies: raising the bar. *Empirical Economics 55*(1), 17–24.

LeSage, J. P., C. Vance, and Y.-Y. Chih (2017). A bayesian heterogeneous coefficients spatial autoregressive panel data model of retail fuel duopoly pricing. *Regional Science and Urban Economics 62*, 46–55.

Parent, O. and J. P. LeSage (2010). A spatial dynamic panel model with random effects applied to commuting times. *Transportation Research Part B: Methodological 44*(5), 633–645.

Parent, O. and J. P. LeSage (2011). A space–time filter for panel data models containing random effects. *Computational Statistics & Data Analysis 55*(1), 475–490.

Parent, O. and J. P. LeSage (2012). Spatial dynamic panel data models with random effects. *Regional Science and Urban Economics 42*(4), 727–738.

Su, L. and Z. Yang (2015). Qml estimation of dynamic panel data models with spatial errors. *Journal of Econometrics 185*(1), 230–258.

Yu, J., R. De Jong, and L.-f. Lee (2008). Quasi-maximum likelihood estimators for spatial dynamic panel data with fixed effects when both n and t are large. *Journal of Econometrics 146*(1), 118–134.

Yu, J., R. de Jong, and L.-f. Lee (2012). Estimation for spatial dynamic panel data with fixed effects: The case of spatial cointegration. *Journal of Econometrics 167*(1), 16–37.

Appendix A

We present Monte Carlo results showing bias, mean-squared error, and 95% coverage for all parameters along with the true and estimated values. The results presented are for cases not on the boundary of the stationary parameter space. The results and discussion of boundary cases presented in the text hold true for the larger sample Monte Carlo experiments. That is, we see relatively poor performance for models where the true parameters lie on the boundary of the stationary parameter space, suggesting practitioners should pay attention to this issue in applied practice.

Comparing the results presented here for the larger sample to those from the text, we see the expected results, smaller bias and mean-squared errors.

Table 15: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	0.999888	-0.000112	0.000010	0.950000
β_2	-1.000000	-0.999815	0.000185	0.000009	0.954000
β_3	1.000000	1.000039	0.000039	0.000010	0.954000
β_4	-1.000000	-1.000048	-0.000048	0.000010	0.948000
ρ	0.200000	0.200185	0.000185	0.000007	0.952000
ϕ	0.500000	0.499933	-0.000067	0.000002	0.938000
θ	-0.300000	-0.300053	-0.000053	0.000008	0.964000
Current period effects					
direct 1	1.007136	1.007042	-0.000094	0.000010	0.946000
direct 2	-1.007136	-1.006968	0.000168	0.000010	0.958000
direct 3	1.007136	1.007194	0.000058	0.000010	0.954000
direct 4	-1.007136	-1.007203	-0.000066	0.000010	0.952000
indirect 1	0.242864	0.243134	0.000271	0.000016	0.954000
indirect 2	-0.242864	-0.243116	-0.000253	0.000016	0.946000
indirect 3	0.242864	0.243171	0.000307	0.000016	0.958000
indirect 4	-0.242864	-0.243173	-0.000309	0.000016	0.948000
total 1	1.250000	1.250177	0.000177	0.000032	0.946000
total 2	-1.250000	-1.250084	-0.000084	0.000031	0.958000
total 3	1.250000	1.250365	0.000365	0.000032	0.952000
total 4	-1.250000	-1.250375	-0.000375	0.000032	0.948000
$A-$ matrix effects					
A direct	0.492864	0.492790	-0.000073	0.000000	0.938000
A indirect	-0.242864	-0.242883	-0.000019	0.000000	0.960000
A total	0.250000	0.249907	-0.000093	0.000000	0.946000
σ^2	1.000000	1.000065	0.000065	0.000021	0.952000

Table 16: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	1.000207	0.000207	0.000011	0.956000
β_2	-1.000000	-1.000016	-0.000016	0.000010	0.958000
β_3	1.000000	1.000197	0.000197	0.000011	0.936000
β_4	-1.000000	-0.999964	0.000036	0.000010	0.960000
ρ	0.200000	0.199981	-0.000019	0.000007	0.950000
ϕ	0.500000	0.500066	0.000066	0.000001	0.956000
θ	-0.500000	-0.499985	0.000015	0.000007	0.950000
Current period effects					
direct 1	1.007136	1.007350	0.000213	0.000011	0.956000
direct 2	-1.007136	-1.007157	-0.000021	0.000010	0.956000
direct 3	1.007136	1.007340	0.000203	0.000011	0.934000
direct 4	-1.007136	-1.007104	0.000032	0.000010	0.966000
indirect 1	0.242864	0.242905	0.000041	0.000015	0.958000
indirect 2	-0.242864	-0.242859	0.000005	0.000015	0.946000
indirect 3	0.242864	0.242904	0.000040	0.000016	0.936000
indirect 4	-0.242864	-0.242846	0.000018	0.000015	0.950000
total 1	1.250000	1.250255	0.000255	0.000032	0.946000
total 2	-1.250000	-1.250016	-0.000016	0.000031	0.936000
total 3	1.250000	1.250244	0.000244	0.000035	0.930000
total 4	-1.250000	-1.249950	0.000050	0.000031	0.956000
$A-$ matrix effects					
A direct	0.485727	0.485795	0.000068	0.000000	0.954000
A indirect	-0.485727	-0.485697	0.000030	0.000000	0.952000
A total	0.000000	0.000098	0.000098	0.000000	0.950000
σ^2	1.000000	0.999979	-0.000021	0.000023	0.932000

Table 17: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	1.000423	0.000423	0.000010	0.956000
β_2	-1.000000	-1.000068	-0.000068	0.000009	0.966000
β_3	1.000000	0.999895	-0.000105	0.000010	0.936000
β_4	-1.000000	-1.000027	-0.000027	0.000010	0.952000
ρ	0.200000	0.200082	0.000082	0.000007	0.946000
ϕ	0.700000	0.700006	0.000006	0.000001	0.958000
θ	-0.300000	-0.300184	-0.000184	0.000008	0.940000
Current period effects					
direct 1	1.007136	1.007577	0.000441	0.000010	0.958000
direct 2	-1.007136	-1.007220	-0.000083	0.000010	0.968000
direct 3	1.007136	1.007046	-0.000090	0.000011	0.934000
direct 4	-1.007136	-1.007178	-0.000042	0.000010	0.952000
indirect 1	0.242864	0.243106	0.000243	0.000016	0.952000
indirect 2	-0.242864	-0.243020	-0.000156	0.000016	0.944000
indirect 3	0.242864	0.242978	0.000115	0.000016	0.946000
indirect 4	-0.242864	-0.243010	-0.000146	0.000016	0.946000
total 1	1.250000	1.250684	0.000684	0.000032	0.952000
total 2	-1.250000	-1.250239	-0.000239	0.000030	0.952000
total 3	1.250000	1.250024	0.000024	0.000033	0.936000
total 4	-1.250000	-1.250188	-0.000188	0.000032	0.942000
$A-$ matrix effects					
A direct	0.694291	0.694289	-0.000002	0.000000	0.958000
A indirect	-0.194291	-0.194461	-0.000170	0.000000	0.950000
A total	0.500000	0.499828	-0.000172	0.000000	0.952000
σ^2	1.000000	0.999912	-0.000088	0.000020	0.944000

Table 18: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	0.999870	-0.000130	0.000010	0.954000
β_2	-1.000000	-0.999976	0.000024	0.000010	0.952000
β_3	1.000000	0.999821	-0.000179	0.000010	0.950000
β_4	-1.000000	-0.999934	0.000066	0.000010	0.950000
ρ	0.200000	0.200125	0.000125	0.000007	0.940000
ϕ	0.700000	0.700035	0.000035	0.000001	0.946000
θ	-0.500000	-0.500052	-0.000052	0.000007	0.944000
Current period effects					
direct 1	1.007136	1.007018	-0.000119	0.000010	0.950000
direct 2	-1.007136	-1.007124	0.000012	0.000010	0.948000
direct 3	1.007136	1.006968	-0.000168	0.000010	0.950000
direct 4	-1.007136	-1.007082	0.000055	0.000010	0.950000
indirect 1	0.242864	0.243042	0.000178	0.000016	0.942000
indirect 2	-0.242864	-0.243067	-0.000203	0.000016	0.938000
indirect 3	0.242864	0.243029	0.000165	0.000016	0.952000
indirect 4	-0.242864	-0.243056	-0.000193	0.000016	0.942000
total 1	1.250000	1.250059	0.000059	0.000033	0.940000
total 2	-1.250000	-1.250191	-0.000191	0.000032	0.948000
total 3	1.250000	1.249997	-0.000003	0.000032	0.958000
total 4	-1.250000	-1.250138	-0.000138	0.000031	0.944000
$A-$ matrix effects					
A direct	0.687155	0.687182	0.000028	0.000000	0.946000
A indirect	-0.437155	-0.437168	-0.000014	0.000000	0.954000
A total	0.250000	0.250014	0.000014	0.000000	0.958000
σ^2	1.000000	0.999892	-0.000108	0.000020	0.956000

Table 19: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	0.999784	-0.000216	0.000011	0.940000
β_2	-1.000000	-0.999996	0.000004	0.000010	0.954000
β_3	1.000000	0.999866	-0.000134	0.000011	0.936000
β_4	-1.000000	-1.000030	-0.000030	0.000009	0.958000
ρ	0.400000	0.401408	0.001408	0.000007	0.904000
ϕ	0.500000	0.500008	0.000008	0.000001	0.952000
θ	-0.300000	-0.300449	-0.000449	0.000006	0.944000
Current period effects					
direct 1	1.033726	1.033799	0.000073	0.000011	0.940000
direct 2	-1.033726	-1.034019	-0.000293	0.000011	0.952000
direct 3	1.033726	1.033884	0.000158	0.000012	0.936000
direct 4	-1.033726	-1.034054	-0.000328	0.000010	0.960000
indirect 1	0.632941	0.636471	0.003530	0.000047	0.914000
indirect 2	-0.632941	-0.636606	-0.003666	0.000048	0.904000
indirect 3	0.632941	0.636524	0.003583	0.000048	0.908000
indirect 4	-0.632941	-0.636629	-0.003688	0.000049	0.894000
total 1	1.666667	1.670270	0.003604	0.000076	0.938000
total 2	-1.666667	-1.670625	-0.003958	0.000077	0.930000
total 3	1.666667	1.670408	0.003741	0.000079	0.920000
total 4	-1.666667	-1.670683	-0.004016	0.000080	0.912000
$A-$ matrix effects					
A direct	0.491569	0.491556	-0.000012	0.000000	0.956000
A indirect	-0.158235	-0.158176	0.000059	0.000000	0.956000
A total	0.333333	0.333381	0.000047	0.000000	0.968000
σ^2	1.000000	0.999901	-0.000099	0.000020	0.952000

Table 20: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	1.000091	0.000091	0.000009	0.960000
β_2	-1.000000	-0.999846	0.000154	0.000010	0.960000
β_3	1.000000	0.999889	-0.000111	0.000010	0.960000
β_4	-1.000000	-1.000029	-0.000029	0.000010	0.954000
ρ	0.400000	0.401406	0.001406	0.000007	0.886000
ϕ	0.500000	0.500164	0.000164	0.000001	0.954000
θ	-0.500000	-0.500390	-0.000390	0.000006	0.950000
Current period effects					
direct 1	1.033726	1.034095	0.000370	0.000010	0.958000
direct 2	-1.033726	-1.033841	-0.000116	0.000010	0.962000
direct 3	1.033726	1.033886	0.000160	0.000011	0.954000
direct 4	-1.033726	-1.034031	-0.000305	0.000011	0.958000
indirect 1	0.632941	0.636684	0.003744	0.000052	0.878000
indirect 2	-0.632941	-0.636527	-0.003587	0.000051	0.892000
indirect 3	0.632941	0.636554	0.003614	0.000051	0.882000
indirect 4	-0.632941	-0.636644	-0.003703	0.000051	0.888000
total 1	1.666667	1.670780	0.004113	0.000082	0.914000
total 2	-1.666667	-1.670369	-0.003702	0.000079	0.908000
total 3	1.666667	1.670440	0.003773	0.000079	0.912000
total 4	-1.666667	-1.670674	-0.004008	0.000080	0.912000
$A-$ matrix effects					
A direct	0.474706	0.474783	0.000077	0.000000	0.964000
A indirect	-0.474706	-0.475163	-0.000458	0.000000	0.954000
A total	0.000000	-0.000381	-0.000381	0.000000	0.960000
σ^2	1.000000	1.000021	0.000021	0.000021	0.954000

Table 21: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	0.999886	-0.000114	0.000010	0.954000
β_2	-1.000000	-0.999884	0.000116	0.000010	0.940000
β_3	1.000000	0.999922	-0.000078	0.000011	0.952000
β_4	-1.000000	-1.000213	-0.000213	0.000009	0.956000
ρ	0.400000	0.401423	0.001423	0.000008	0.886000
ϕ	0.700000	0.700058	0.000058	0.000001	0.944000
θ	-0.300000	-0.300897	-0.000897	0.000007	0.938000
Current period effects					
direct 1	1.033726	1.033893	0.000167	0.000011	0.954000
direct 2	-1.033726	-1.033891	-0.000165	0.000011	0.940000
direct 3	1.033726	1.033930	0.000204	0.000011	0.954000
direct 4	-1.033726	-1.034231	-0.000505	0.000010	0.950000
indirect 1	0.632941	0.636592	0.003651	0.000053	0.876000
indirect 2	-0.632941	-0.636591	-0.003650	0.000053	0.890000
indirect 3	0.632941	0.636615	0.003674	0.000053	0.894000
indirect 4	-0.632941	-0.636800	-0.003859	0.000054	0.890000
total 1	1.666667	1.670486	0.003819	0.000081	0.914000
total 2	-1.666667	-1.670481	-0.003815	0.000080	0.922000
total 3	1.666667	1.670545	0.003878	0.000083	0.924000
total 4	-1.666667	-1.671031	-0.004365	0.000083	0.922000
$A-$ matrix effects					
A direct	0.698314	0.698373	0.000060	0.000000	0.940000
A indirect	-0.031647	-0.031525	0.000122	0.000000	0.970000
A total	0.666667	0.666849	0.000182	0.000000	0.970000
σ^2	1.000000	0.999793	-0.000207	0.000021	0.948000

Table 22: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	0.999834	-0.000166	0.000011	0.946000
β_2	-1.000000	-0.999612	0.000388	0.000011	0.928000
β_3	1.000000	0.999824	-0.000176	0.000011	0.950000
β_4	-1.000000	-0.999809	0.000191	0.000010	0.950000
ρ	0.400000	0.401642	0.001642	0.000008	0.894000
ϕ	0.700000	0.700080	0.000080	0.000001	0.950000
θ	-0.500000	-0.500763	-0.000763	0.000006	0.932000
Current period effects					
direct 1	1.033726	1.033881	0.000156	0.000012	0.944000
direct 2	-1.033726	-1.033652	0.000074	0.000012	0.946000
direct 3	1.033726	1.033872	0.000146	0.000011	0.950000
direct 4	-1.033726	-1.033856	-0.000130	0.000011	0.958000
indirect 1	0.632941	0.637122	0.004181	0.000051	0.924000
indirect 2	-0.632941	-0.636981	-0.004040	0.000051	0.924000
indirect 3	0.632941	0.637117	0.004176	0.000053	0.912000
indirect 4	-0.632941	-0.637108	-0.004167	0.000053	0.914000
total 1	1.666667	1.671003	0.004336	0.000078	0.920000
total 2	-1.666667	-1.670633	-0.003966	0.000077	0.934000
total 3	1.666667	1.670988	0.004322	0.000082	0.928000
total 4	-1.666667	-1.670963	-0.004297	0.000080	0.924000
$A-$ matrix effects					
A direct	0.681451	0.681461	0.000011	0.000000	0.946000
A indirect	-0.348117	-0.348358	-0.000240	0.000000	0.962000
A total	0.333333	0.333104	-0.000230	0.000000	0.962000
σ^2	1.000000	0.999674	-0.000326	0.000019	0.950000

Table 23: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.2, ϕ =0.5, θ =-0.3, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	1.000391	0.000391	0.000086	0.958000
β_2	-1.000000	-1.000463	-0.000463	0.000096	0.958000
β_3	1.000000	0.999842	-0.000158	0.000092	0.952000
β_4	-1.000000	-0.999627	0.000373	0.000090	0.960000
ρ	0.200000	0.200297	0.000297	0.000019	0.946000
ϕ	0.500000	0.499931	-0.000069	0.000005	0.950000
θ	-0.300000	-0.300070	-0.000070	0.000026	0.948000
Current period effects					
direct 1	1.007136	1.007561	0.000424	0.000088	0.958000
direct 2	-1.007136	-1.007633	-0.000497	0.000098	0.958000
direct 3	1.007136	1.007007	-0.000129	0.000093	0.954000
direct 4	-1.007136	-1.006791	0.000346	0.000091	0.958000
indirect 1	0.242864	0.243463	0.000600	0.000048	0.950000
indirect 2	-0.242864	-0.243482	-0.000618	0.000049	0.940000
indirect 3	0.242864	0.243330	0.000467	0.000048	0.952000
indirect 4	-0.242864	-0.243274	-0.000410	0.000046	0.954000
total 1	1.250000	1.251024	0.001024	0.000180	0.960000
total 2	-1.250000	-1.251115	-0.001115	0.000198	0.952000
total 3	1.250000	1.250338	0.000338	0.000190	0.934000
total 4	-1.250000	-1.250064	-0.000064	0.000176	0.958000
$A-$ matrix effects					
A direct	0.492864	0.492785	-0.000078	0.000000	0.954000
A indirect	-0.242864	-0.242869	-0.000005	0.000000	0.954000
A total	0.250000	0.249917	-0.000083	0.000000	0.944000
σ^2	10.000000	10.000329	0.000329	0.002103	0.948000

Table 24: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.2, ϕ =0.5, θ =-0.5, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	1.000110	0.000110	0.000104	0.940000
β_2	-1.000000	-1.000337	-0.000337	0.000102	0.944000
β_3	1.000000	0.999722	-0.000278	0.000110	0.940000
β_4	-1.000000	-1.000467	-0.000467	0.000101	0.952000
ρ	0.200000	0.200385	0.000385	0.000019	0.938000
ϕ	0.500000	0.500113	0.000113	0.000005	0.956000
θ	-0.500000	-0.500364	-0.000364	0.000026	0.956000
Current period effects					
direct 1	1.007136	1.007284	0.000148	0.000105	0.944000
direct 2	-1.007136	-1.007513	-0.000377	0.000103	0.944000
direct 3	1.007136	1.006894	-0.000242	0.000112	0.940000
direct 4	-1.007136	-1.007644	-0.000508	0.000103	0.956000
indirect 1	0.242864	0.243523	0.000659	0.000047	0.932000
indirect 2	-0.242864	-0.243574	-0.000710	0.000045	0.948000
indirect 3	0.242864	0.243428	0.000564	0.000047	0.924000
indirect 4	-0.242864	-0.243611	-0.000747	0.000047	0.928000
total 1	1.250000	1.250807	0.000807	0.000206	0.926000
total 2	-1.250000	-1.251087	-0.001087	0.000192	0.956000
total 3	1.250000	1.250322	0.000322	0.000212	0.938000
total 4	-1.250000	-1.251255	-0.001255	0.000204	0.944000
$A-$ matrix effects					
A direct	0.485727	0.485804	0.000077	0.000000	0.958000
A indirect	-0.485727	-0.486127	-0.000400	0.000000	0.960000
A total	0.000000	-0.000323	-0.000323	0.000000	0.966000
σ^2	10.000000	10.001444	0.001444	0.002206	0.944000

Table 25: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.2, ϕ =0.7, θ =-0.3, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	0.999858	-0.000142	0.000100	0.956000
β_2	-1.000000	-0.999687	0.000313	0.000094	0.944000
β_3	1.000000	0.999785	-0.000215	0.000092	0.960000
β_4	-1.000000	-1.000156	-0.000156	0.000105	0.942000
ρ	0.200000	0.200233	0.000233	0.000017	0.964000
ϕ	0.700000	0.699882	-0.000118	0.000004	0.960000
θ	-0.300000	-0.300158	-0.000158	0.000023	0.964000
Current period effects					
direct 1	1.007136	1.007021	-0.000115	0.000101	0.960000
direct 2	-1.007136	-1.006849	0.000288	0.000095	0.946000
direct 3	1.007136	1.006947	-0.000189	0.000094	0.958000
direct 4	-1.007136	-1.007321	-0.000185	0.000107	0.942000
indirect 1	0.242864	0.243235	0.000372	0.000044	0.956000
indirect 2	-0.242864	-0.243190	-0.000327	0.000042	0.962000
indirect 3	0.242864	0.243214	0.000351	0.000042	0.966000
indirect 4	-0.242864	-0.243304	-0.000440	0.000043	0.962000
total 1	1.250000	1.250257	0.000257	0.000206	0.954000
total 2	-1.250000	-1.250039	-0.000039	0.000187	0.958000
total 3	1.250000	1.250162	0.000162	0.000186	0.954000
total 4	-1.250000	-1.250625	-0.000625	0.000204	0.944000
$A-$ matrix effects					
A direct	0.694291	0.694164	-0.000127	0.000000	0.962000
A indirect	-0.194291	-0.194367	-0.000076	0.000000	0.968000
A total	0.500000	0.499797	-0.000203	0.000000	0.966000
σ^2	10.000000	9.999368	-0.000632	0.001928	0.944000

Table 26: Monte Carlo results ($n = 2,000, T = 50$), all parameters

		ρ =0.2, ϕ =0.7, θ =-0.5, σ^2 =1			
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	1.000080	0.000080	0.000089	0.962000
β_2	-1.000000	-0.999721	0.000279	0.000097	0.948000
β_3	1.000000	1.000039	0.000039	0.000093	0.966000
β_4	-1.000000	-0.999585	0.000415	0.000101	0.950000
ρ	0.200000	0.200131	0.000131	0.000018	0.950000
ϕ	0.700000	0.699929	-0.000071	0.000003	0.950000
θ	-0.500000	-0.499921	0.000079	0.000024	0.964000
Current period effects					
direct 1	1.007136	1.007233	0.000097	0.000090	0.966000
direct 2	-1.007136	-1.006872	0.000264	0.000098	0.950000
direct 3	1.007136	1.007192	0.000056	0.000094	0.962000
direct 4	-1.007136	-1.006735	0.000401	0.000103	0.952000
indirect 1	0.242864	0.243137	0.000273	0.000043	0.954000
indirect 2	-0.242864	-0.243052	-0.000188	0.000044	0.946000
indirect 3	0.242864	0.243134	0.000270	0.000046	0.948000
indirect 4	-0.242864	-0.243022	-0.000158	0.000046	0.946000
total 1	1.250000	1.250370	0.000370	0.000173	0.956000
total 2	-1.250000	-1.249924	0.000076	0.000190	0.960000
total 3	1.250000	1.250326	0.000326	0.000197	0.948000
total 4	-1.250000	-1.249757	0.000243	0.000206	0.954000
$A-$ matrix effects					
A direct	0.687155	0.687079	-0.000075	0.000000	0.952000
A indirect	-0.437155	-0.437039	0.000116	0.000000	0.966000
A total	0.250000	0.250041	0.000041	0.000000	0.958000
σ^2	10.000000	10.000961	0.000961	0.001647	0.976000

Table 27: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.4, ϕ =0.5, θ =-0.3, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	1.000119	0.000119	0.000110	0.930000
β_2	-1.000000	-0.999352	0.000648	0.000095	0.946000
β_3	1.000000	1.000447	0.000447	0.000096	0.962000
β_4	-1.000000	-0.999975	0.000025	0.000094	0.960000
ρ	0.400000	0.403800	0.003800	0.000028	0.814000
ϕ	0.500000	0.500129	0.000129	0.000006	0.938000
θ	-0.300000	-0.301654	-0.001654	0.000024	0.944000
Current period effects					
direct 1	1.033726	1.034632	0.000906	0.000118	0.936000
direct 2	-1.033726	-1.033839	-0.000113	0.000102	0.946000
direct 3	1.033726	1.034972	0.001246	0.000104	0.956000
direct 4	-1.033726	-1.034483	-0.000757	0.000101	0.962000
indirect 1	0.632941	0.642973	0.010032	0.000229	0.868000
indirect 2	-0.632941	-0.642486	-0.009545	0.000220	0.856000
indirect 3	0.632941	0.643188	0.010247	0.000232	0.856000
indirect 4	-0.632941	-0.642881	-0.009940	0.000220	0.878000
total 1	1.666667	1.677605	0.010938	0.000514	0.908000
total 2	-1.666667	-1.676325	-0.009658	0.000465	0.928000
total 3	1.666667	1.678160	0.011493	0.000501	0.910000
total 4	-1.666667	-1.677364	-0.010697	0.000465	0.932000
$A-$ matrix effects					
A direct	0.491569	0.491611	0.000042	0.000000	0.946000
A indirect	-0.158235	-0.158714	-0.000478	0.000000	0.950000
A total	0.333333	0.332897	-0.000436	0.000000	0.948000
σ^2	10.000000	9.992946	-0.007054	0.001986	0.958000

Table 28: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.4, ϕ =0.5, θ =-0.5, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	1.000512	0.000512	0.000111	0.942000
β_2	-1.000000	-0.999367	0.000633	0.000110	0.938000
β_3	1.000000	1.000358	0.000358	0.000087	0.966000
β_4	-1.000000	-0.999759	0.000241	0.000102	0.936000
ρ	0.400000	0.403778	0.003778	0.000028	0.804000
ϕ	0.500000	0.500445	0.000445	0.000006	0.932000
θ	-0.500000	-0.501147	-0.001147	0.000021	0.940000
Current period effects					
direct 1	1.033726	1.035032	0.001306	0.000120	0.940000
direct 2	-1.033726	-1.033847	-0.000121	0.000116	0.936000
direct 3	1.033726	1.034873	0.001147	0.000094	0.958000
direct 4	-1.033726	-1.034254	-0.000528	0.000111	0.944000
indirect 1	0.632941	0.643175	0.010234	0.000241	0.842000
indirect 2	-0.632941	-0.642432	-0.009491	0.000218	0.860000
indirect 3	0.632941	0.643076	0.010135	0.000230	0.850000
indirect 4	-0.632941	-0.642697	-0.009757	0.000237	0.834000
total 1	1.666667	1.678207	0.011540	0.000544	0.888000
total 2	-1.666667	-1.676279	-0.009613	0.000479	0.920000
total 3	1.666667	1.677949	0.011283	0.000474	0.914000
total 4	-1.666667	-1.676952	-0.010285	0.000517	0.892000
$A-$ matrix effects					
A direct	0.474706	0.474894	0.000188	0.000000	0.942000
A indirect	-0.474706	-0.476086	-0.001380	0.000002	0.952000
A total	0.000000	-0.001192	-0.001192	0.000001	0.944000
σ^2	10.000000	9.992204	-0.007796	0.002208	0.948000

Table 29: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.4, ϕ =0.7, θ =-0.3, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	0.999836	-0.000164	0.000091	0.960000
β_2	-1.000000	-1.000305	-0.000305	0.000094	0.952000
β_3	1.000000	0.999529	-0.000471	0.000102	0.964000
β_4	-1.000000	-0.998631	0.001369	0.000102	0.950000
ρ	0.400000	0.403861	0.003861	0.000030	0.808000
ϕ	0.700000	0.699962	-0.000038	0.000004	0.950000
θ	-0.300000	-0.302593	-0.002593	0.000025	0.916000
Current period effects					
direct 1	1.033726	1.034350	0.000624	0.000100	0.958000
direct 2	-1.033726	-1.034834	-0.001109	0.000101	0.950000
direct 3	1.033726	1.034032	0.000307	0.000109	0.966000
direct 4	-1.033726	-1.033104	0.000622	0.000109	0.950000
indirect 1	0.632941	0.642968	0.010027	0.000239	0.844000
indirect 2	-0.632941	-0.643260	-0.010319	0.000233	0.848000
indirect 3	0.632941	0.642766	0.009825	0.000233	0.866000
indirect 4	-0.632941	-0.642194	-0.009253	0.000228	0.860000
total 1	1.666667	1.677318	0.010651	0.000490	0.916000
total 2	-1.666667	-1.678094	-0.011427	0.000478	0.930000
total 3	1.666667	1.676799	0.010132	0.000491	0.920000
total 4	-1.666667	-1.675297	-0.008631	0.000476	0.904000
$A-$ matrix effects					
A direct	0.698314	0.698262	-0.000052	0.000000	0.956000
A indirect	-0.031647	-0.031687	-0.000040	0.000000	0.974000
A total	0.666667	0.666574	-0.000092	0.000000	0.966000
σ^2	10.000000	9.987952	-0.012048	0.002281	0.932000

Table 30: Monte Carlo results ($n = 2,000, T = 50$), all parameters

	ρ =0.4, ϕ =0.7, θ =-0.5, σ^2 =1				
Parameters	truth	estimates	bias	MSE	95% coverage
β_1	1.000000	0.999661	-0.000339	0.000092	0.964000
β_2	-1.000000	-1.000322	-0.000322	0.000098	0.954000
β_3	1.000000	1.000629	0.000629	0.000096	0.966000
β_4	-1.000000	-0.999557	0.000443	0.000097	0.950000
ρ	0.400000	0.404102	0.004102	0.000030	0.786000
ϕ	0.700000	0.700151	0.000151	0.000004	0.954000
θ	-0.500000	-0.502033	-0.002033	0.000023	0.930000
Current period effects					
direct 1	1.033726	1.034221	0.000495	0.000098	0.956000
direct 2	-1.033726	-1.034905	-0.001179	0.000108	0.954000
direct 3	1.033726	1.035223	0.001497	0.000105	0.962000
direct 4	-1.033726	-1.034113	-0.000388	0.000105	0.950000
indirect 1	0.632941	0.643470	0.010529	0.000238	0.858000
indirect 2	-0.632941	-0.643901	-0.010961	0.000258	0.810000
indirect 3	0.632941	0.644097	0.011156	0.000258	0.824000
indirect 4	-0.632941	-0.643409	-0.010468	0.000247	0.842000
total 1	1.666667	1.677691	0.011024	0.000479	0.914000
total 2	-1.666667	-1.678806	-0.012139	0.000544	0.896000
total 3	1.666667	1.679319	0.012653	0.000541	0.890000
total 4	-1.666667	-1.677522	-0.010856	0.000510	0.910000
$A-$ matrix effects					
A direct	0.681451	0.681411	-0.000039	0.000000	0.952000
A indirect	-0.348117	-0.348950	-0.000832	0.000001	0.956000
A total	0.333333	0.332462	-0.000871	0.000001	0.954000
σ^2	10.000000	9.997190	-0.002810	0.002041	0.950000