Fast Two-Stage Variational Bayesian Approach to Estimating Panel Spatial Autoregressive Models with Unrestricted Spatial Weights Matrices

Deborah Gefang[1], Stephen G. Hall[2], and George S. Tavlas[3]

[1]University of Leicester
[2]University of Leicester, Bank of Greece and University of Pretoria
[3]Bank of Greece and Hoover Institution, Stanford University

Abstract

This paper proposes a fast two-stage variational Bayesian algorithm to estimating panel spatial autoregressive models with unknown spatial weights matrices. Using Dirichlet-Laplace global-local shrinkage priors, we are able to uncover the spatial impacts between cross-sectional units without imposing any a priori restrictions. Monte Carlo experiments show that our approach works well for both long and short panels. We are also the first in the literature to develop VB methods to estimate large covariance matrices with unrestricted sparsity patterns. The method is important in itself because of its relevance to other popular large data models such as Bayesian vector autoregressions. Matlab code is provided.

JEL Classification: C11; C33; C55

Keywords: Variational Inference; Spatial Panel Data Models; Simultaneous Equations; Large Data

1 Introduction

The spatial autoregressive (SAR) models, first proposed by Cliff and Ord (1973), have been widely used in the literature to investigate the spatial dependence in cross-sectional units (e.g., Anselin, 1988; Baltagi et al., 2003, 2013; Lee and Yu, 2010). In practice, the spatial weights matrices are usually set a priori based on geographical distances or economic characters (e.g. Cliff and Ord, 1973; Anselin, 1988; Case, 1991). This is not surprising as a spatial weights matrix potentially involves $N^2 - N$ interrelationships between N spatial units, which makes it difficult to estimate, especially when N is large.

In recent years, a number of variable selection and parameter shrinkage methods have been developed to estimate the spatial weights matrices of panel SAR models. Among them, methods resorting to least absolute shrinkage and selection operator (Lasso) of Tibshirani's (1996) and its variants have gained a lot of attention. For example, Basak et al. (2018) propose to estimate a triangular weights matrix under the assumption of recursive ordering. Ahrens and Bhattacharjee (2015) develop a two step Lasso estimator to identify the weights matrix. Lam and Souza (2019) estimate the weights matrix using adaptive Lasso with sparse adjustment in mind. Most of the studies, however, usually impose sometime unrealistic restrictions on the model's coefficients or covariances. Krock et al. (2021) develop a graphical Lasso approach to estimating the unrestricted covariances. Their method, however, does not deal with the impacts of any possible exogenous variables. Moreover, to our knowledge, few of those studies focus on the short panels where N is large while T is small.

This paper contributes to the SAR literature by developing a fast two-stage variational Bayesian (VB) approach to estimating panel SAR models with unknown spatial weights matrices. We do not impose any restrictions on spatial weights matrix or the covariance functions, hence our approach lets the data speak. The prior we used for Bayesian regularization is the Dirichlet–Laplace (D–L) prior of Bhattacharya et al. (2015). With D-L prior, the entire posterior distribution concentrates at the optimal rate. This nice feature remains unchanged when the number of parameters to be estimated is much larger than the number of observations, providing strong theoretical justifications for the two-stage VB's effectiveness in uncovering the spatial dependencies in a short panel.

A secondary contribution of this paper lies in developing VB methods for estimating large covariance matrices with D-L prior. We are among the first in the literature to develop VB estimator for large covariance matrices with unknown sparsity patterns. Our VB methods using D-L prior can be easily extended to allow for other popular priors such as the

graphic Lasso of Wang (2012), the half-Cauchy prior of Makalic and Schimidt (2016) and the graphical horseshoe prior of Li et al. (2019). This is not trivial as VB is a more computationally efficient alternative of Markov Chain Monte Carlo (MCMC), and our approach can be used in estimating other popular models involving large covariance matrices such as large Bayesian vector autoregressions (BVARs).

We have conducted a wide range of simulation studies using a traditional panel SAR model and a panel SAR model that takes account of the simultaneous relationships between cross-sectional groups.[1] Monte Carlo experiments show that two-stage VB is accurate and computationally efficient when $T \gg N$, which usually is more pertinent to macroeconomic and financial data. When $N \gg T$, which tends to be more relevant to microeconomic data, two-stage VB estimates tend to have slightly larger biases and empirical standard deviations. Tighter priors can help lessen that problem.

The rest of the paper is organised as follows. Section 2 extends the traditional panel SAR models to an unrestricted panel SAR. Section 3 develops the two-stage VB. Section 4 conducts Monte Carlo studies. Section 5 concludes with discussions. Detailed VB derivation formulas and more extensive Monte Carlo results are relegated to Online Appendices A to C.

2 Unrestricted Panel SAR Model

In this section, we start from a traditional standard panel SAR and then relax the restrictions imposed upon it in steps, with the aim of giving a flavour of the differences between the traditional model and the unrestricted panel SAR model that we set to estimate using two-stage VB.

Let Y, X and V denote the $T \times N$ matrix of endogenous variables, $T \times (Nm)$ matrix of exogenous variables, and $T \times N$ matrix of disturbances, respectively. A traditional panel SAR model takes the following form:

$$y_t = \lambda W_n y_t + X_t \beta + u_t, \;\; |\lambda| < 1 \tag{1}$$

where $y_t = (Y_{t1}, Y_{t2}, ..., Y_{tN})'$ is the $N \times 1$ vector of observations of the dependent variables,

[1]This research used the ALICE High Performance Computing Facility at the University of Leicester.

W_n is an $N \times N$ known spatial weights matrix with zero diagonal entries, $X_t = \begin{pmatrix} X_{t,1} \\ X_{t,2} \\ ... \\ X_{t,N} \end{pmatrix}$ is the $N \times m$ matrix of exogenous variables, with $X_{t,i}$ denoting the $1 \times m$ row vector of exogenous variables associated with dependent variable y_{ti}, β is an $m \times 1$ vector of parameters, λ is a scalar parameter, and $u_t = (V_{t1}, V_{t2}, ..., V_{tN})'$ is the $N \times 1$ i.i.d error terms with mean zero and diagonal covariance matrix $\begin{pmatrix} \sigma^2 & 0 & ... & 0 \\ 0 & \sigma^2 & ... & 0 \\ .. & .. & ... & .. \\ 0 & 0 & ... & \sigma^2 \end{pmatrix}$.

Model (1) imposes the following unrealistic restrictions on the data generating process: 1) W_N is predetermined, in a fashion that is not related to the variations in the data; 2) λ and β remain the same across equations associated with different dependent variables; and 3) the covariance matrix of u_t is diagonal with the same diagonal entries.

Relaxing those restrictions, model (1) can be written as:

$$y_t = \tilde{\lambda} \odot (\widetilde{W}_N y_t) + \widetilde{X}_t \tilde{\beta} + u_t, \;\; |\lambda_b|_\infty < 1 \tag{2}$$

where $\widetilde{W}_N$ is an $N \times N$ unknown spatial weight matrix with zero diagonal entries, $\tilde{\lambda}$ is a $N \times 1$ parameter vector, $\odot$ is the Hadamard product, $\widetilde{X}_t = \begin{pmatrix} X_{t,1} & 0 & ... & 0 \\ 0 & X_{t,2} & ... & 0 \\ .. & .. & ... & .. \\ 0 & 0 & ... & X_{t,N} \end{pmatrix}$, $\tilde{\beta}$ is a $Nm \times 1$ parameter vector, and u_t is i.i.d with mean zero and diagonal covariance matrix $\tilde{\Sigma}$ with diagonal entries that can be different from each other. In this model, the dimensions of $X_{t,i}$ and $X_{t,j}$ for $i \neq j$ can differ from each other. Let the dimension of $X_{t,i}$ to be $1 \times m_i$. The dimension of the parameter vector $\tilde{\beta}$ is thus $(\sum_{i=1}^{n} m_i) \times 1$.

Note that model (2) is quite flexible. For example, with appropriate restrictions, it can be easily transformed back into the traditional form described in (1) or a panel SAR containing the simultaneous cross-sectional spatial relationship as described in Yang and Lee (2017) and Liu and Saraiva (2019).

Since our main concerns in panel SAR models are the spillover effects, there is therefore little research interest in separately identifying $\tilde{\lambda}$ and $\widetilde{W}_N$. What we care about is the

product $(\tilde{\lambda} \otimes l_N) \odot \widetilde{W}_N$, where l_N is a $N \times 1$ column of ones and $\otimes$ is the Kronecker product, as $(\tilde{\lambda} \otimes l_N) \odot \widetilde{W}_N$ is the $N \times N$ parameter matrix which captures the spillover effects between spatial units.

Let $\Lambda = (\tilde{\lambda} \otimes l_N) \odot \widetilde{W}_N$. Model (2) can be written as:

$$y_t = \Lambda y_t + \widetilde{X}_t \tilde{\beta} + u_t, \tag{3}$$

where $I_N - \Lambda$ is nonsingular and the characteristic roots of $I_N - \Lambda$ lie within the unit circle.

The attractiveness of model (3) is that it turns an unrestricted panel SAR model into a system of simultaneous equations (SEM). As shown in Zellner and Thell (1962) and Fox (1979), the i^{th} equation in model (3) is just identified if $\sum_{i=1}^{N} m_i = N - 1 + m_i$ and over-identified if $\sum_{i=1}^{N} m_i > N - 1 + m_i$. Under these circumstances, a myriad of estimation methods, such as two-stage least squares (2SLS), three-stage least squares (3SLS), maximum likelihoods methods and simultaneous generalized method (GMM), can be used to uncover the structural parameters Λ and $\tilde{\beta}$. Moreover, standard tests can be developed to test the restrictions on $\tilde{\lambda}$, $\widetilde{W}_N$ and $\tilde{\beta}$, if those restrictions are of the researchers' concerns.

This paper proposes to estimate Λ and $\tilde{\beta}$ in two stages as it is computationally simple. To estimate the parameters associated with the i^{th} individual dependent variable, in the first stage, we estimate

$$Y_{/i} = X \Upsilon_i + E_i, \tag{4}$$

where $Y_{/i}$ is the $T \times (N-1)$ matrix of dependent variables except for the i^{th} dependent variable, and E_i is a $T \times (N-1)$ matrix of error terms whose precision matrix might not be diagonal.

Making use of the estimated Υ_i, in the second stage, we estimate

$$y_i = \widehat{Y}_{/i} (\Lambda_{i\bullet})' + X_{\bullet,i} \tilde{\beta}_i + u_i, \tag{5}$$

where y_i is the $T \times 1$ vector of the i^{th} dependent variable, $\widehat{Y}_{/i} = X \Upsilon_i$, $\Lambda_{i\bullet}$ is the $1 \times (n-1)$ vector of the i^{th} row of Λ with Λ_{ii} dropped, and $\tilde{\beta}_i$ is the corresponding coefficients in $\tilde{\beta}$.

Note that in a panel SAR model, we can have $N \gg T$ and $(N - 1 + m_i) \gg T$, which makes it difficult or even impossible to uncover B_i and $\Lambda_{i\bullet}$ using traditional estimation techniques.

3 Two-stage VB

As detailed in Ormerod and Wand (2010) and Blei et al. (2017), the essence of VB is to use appropriate densities from a mean field variational family to approximate the posterior densities through minimizing the Kullback-Leibler divergence, which is equivalent to maximising the evidence lower bound (ELBO). As a more efficient alternative to MCMC, VB has been increasingly used in sophisticated models involving large data where MCMC is too computationally expensive or even untenable (e.g. Gefang et al., 2020, 2022; Loaiza-Maya et al., 2022).

In the two-stage VB, we identify the parameters in model (3) equation by equation. In each stage, we update the parameters using the approximate q densities by iterations. The convergence of the algorithm can be measured by the changes in ELBO across iterations is less than a convergence criteria. When the number of parameters is large, however, calculating ELBO can be time consuming. It is therefore more convenient to check if convergence has occurred by examining if the VB estimates of parameters stop changing across iterations.

3.1 First-stage VB

In the first stage, we estimate model (4) in order to construct the predicted value of $Y_{/i}$.

Let $\gamma = vec(\Upsilon_i)$. We set hierarchical D-L prior for the j^{th}, for $j = 1, ..., np$, element of γ as follows:

$$\gamma_j|\phi, \tau \sim DE(\phi_j\tau), \;\; \phi_j \sim Dir(a, ..., a), \;\; \tau \sim G(npa, 1/2). \tag{6}$$

where $DE(\bullet)$ denotes Double Exponential or Lapalace distribution, $Dir(\bullet)$ denotes Dirichlet distribution, $G(\bullet)$ denotes Gamma distribution, $n = N - 1$, and $p = \sum_{i=1}^{N} m_i$.

Next, we set Exponential priors and D-L priors for the elements of Ω, the precision matrix of E_i, as follows:

$$\begin{aligned}
&\omega_{ii} \sim Exp(\underline{s}), \;\; i = 1, ..., n \\
&\omega_{ij} \sim N(0, \psi_{\omega,ij}\phi^2_{\omega,ij}\tau^2_\omega), \;\; \psi_{\omega,ij} \sim Exp(1/2), \;\; i < j = 2, ..., n, \\
&\phi_{\omega,ij} \sim Dir(a_\omega, ..., a_\omega), \;\; \tau_\omega \sim G(\frac{n^2 - n}{2}a_\omega, 1/2)
\end{aligned} \tag{7}$$

where, with a slight abuse of notations, we use ω_{ii} and ω_{ij} to denote the diagonal and off-diagonal elements of Ω.

Following Wang's (2012) Block Gibbs sampler to update the relevant parameters and hyperparameters, we use the last column and row of Ω as an example on how to update Ω.

Let $S = E_i' E_i$ and H be the $n \times n$ matrix with 0 diagonal elements and the off diagonal element at i^{th} row and j^{th} column be $\psi_{\omega,ij}\phi^2_{\omega,ij}\tau^2_\omega$. Partition Ω, S and H as follows:

$$\Omega = \begin{pmatrix} \Omega_{-n,-n} & \omega_{-n,n} \\ \omega'_{-n,n} & \omega_{nn} \end{pmatrix} \quad S = \begin{pmatrix} S_{-n,-n} & s_{-n,n} \\ s'_{-n,n} & s_{nn} \end{pmatrix} \quad H = \begin{pmatrix} H_{-n,-n} & h_{-n,n} \\ h'_{-n,n} & h_{nn} \end{pmatrix} \tag{8}$$

where $-n$ denotes the set of all indices except for n.

Relegating technical details to Online Appendix A, we outline the VB approximation densities as follows:

3.1.1 $q(\gamma)$

$$q(\gamma) \sim N(\bar{\gamma}, \overline{V}), \tag{9}$$

where

$$\overline{V} = (V^{-1} + \overline{\Omega} \otimes (X'X))^{-1},$$

$$\overline{\gamma} = \overline{V}(\overline{\Omega} \otimes X')vec(Y_{/i}),$$

and

$$V^{-1} = diag(\frac{1}{\overline{\psi_1}\,\overline{\phi_1^2}\,\overline{\tau^2}}, ..., \frac{1}{\overline{\psi_{np}}\,\overline{\phi_{np}^2}\,\overline{\tau^2}})$$

3.1.2 $q(\tau)$

$$q(\tau) \sim giG[npa - np, 1, \sum_{j=1}^{np} 2(\bar{\gamma}_j^2 + \overline{V}_{jj})^{1/2}\frac{1}{\overline{\phi_j}}], \tag{10}$$

Let $\chi = \sum_{j=1}^{np} 2(\bar{b}_j^2 + \overline{V}_{jj})^{1/2}\frac{1}{\overline{\phi_j}}$, we have

$$\bar{\tau} = \frac{\sqrt{\chi}K_{npa-np+1}(\sqrt{\chi})}{K_{npa-np}(\sqrt{\chi})}$$

$$\overline{\tau^2} = \bar{\tau}^2 + \chi[\frac{K_{npa-np+2}(\sqrt{\chi})}{K_{npa-np}(\sqrt{\chi})} - (\frac{K_{npa-np+1}(\sqrt{\chi})}{K_{npa-np}(\sqrt{\chi})})^2]$$

where $K_*[\bullet]$ is the modified Bessel functions of the second kind.

3.1.3 $q(\psi_j)$

$$q(\psi_j^{-1}) \sim iG(\sqrt{\frac{\overline{\phi_j^2}\,\overline{\tau^2}}{\overline{\gamma}_j^2 + \overline{V}_{jj}}}, 1), \tag{11}$$

where $iG(\bullet)$ denotes s Inverse Gaussian distribution.

Let $\rho = \sqrt{\frac{\overline{\phi_j^2}\,\overline{\tau^2}}{\overline{\gamma}_j^2+\overline{V}_{jj}}}$,

$$\overline{\psi_j^{-1}} = \rho$$

and

$$\overline{\psi}_j = 1 + 1/\rho$$

3.1.4 $q(\phi_j)$

$$q(\xi_j) \sim giG(a-1, 1, 2\sqrt{\overline{\gamma}_j^2 + \overline{V}_{jj}}), \tag{12}$$

where $giG(\bullet)$ denotes the generalized inverse Gaussian distribution.

Let $\varpi = 2\sqrt{\overline{\gamma}_j^2 + \overline{V}_{jj}}$, we have

$$\overline{\xi_j} = \frac{\sqrt{\varpi} K_a(\sqrt{\varpi})}{K_{a-1}(\sqrt{\varpi})},$$

and

$$var(\xi_j) = \varpi\{\frac{K_{a+1}(\sqrt{\varpi})}{K_{a-1}(\sqrt{\varpi})} - [\frac{K_a(\sqrt{\varpi})}{K_{a-1}(\sqrt{\varpi})}]^2\},$$

where $var(\bullet)$ denotes the variance.

Scaling ξ_i, we have

$$\overline{\phi_j} = \frac{\overline{\xi_j}}{\sum_{j=1}^{np} \overline{\xi_j}},$$

and

$$\overline{\phi_j^2} = \overline{\phi_j}^2 + \frac{var(\xi_j)}{(\sum_{j=1}^{np} \overline{\xi_j})^2}$$

Thus, the optimal q density of $\phi_{i,j}$ takes the following form:

$$q(\phi_j) \sim giG(a-1, \sum_{j=1}^{np} \overline{\xi_j}, \frac{2\sqrt{\overline{\gamma}_j^2 + (\overline{V}_{jj})^2}}{\sum_{j=1}^{np} \overline{\xi_j}}) \tag{13}$$

3.1.5 $q(b_1)$

Let $b_1 = \omega_{n,n} - \omega'_{-n,n}\Omega^{-1}_{-n,-n}\omega_{-n,n}$.

$$q(b_1) \sim G(\frac{T}{2}, \overline{s}_{n,n}), \tag{14}$$

where

$$\overline{s}_{n,n} = \frac{1}{2}(s_{n,n} + tr(X'X\overline{V_n}) + \underline{s}),$$

and

$$\overline{V_n} = V_{(n-1)\times p+1:n\times p,(n-1)\times p+1:n\times p}.$$

Hence

$$\overline{b_1} = \frac{\frac{T}{2}}{\overline{s}_{n,n}}$$

3.1.6 $q(b_2)$

Here we use b_2 to denote $\omega_{-n,n}$. Let $\overline{s}_{-n,n} = s_{-n,n} + \tilde{s}_{-n,n}$, where $\tilde{s}_{-n,n}$ is a $(n-1)\times 1$ vector with the j^{th} element being $tr(X'XA_j)$ and $A_j = V_{(j-1)\times p+1:j\times p,(j-1)\times p+1:j\times p}$. [2]

$$q(b_2) \sim N(-\overline{C}\overline{s}_{-n,n}, \overline{C}), \tag{15}$$

where

$$\overline{C} = (2\overline{s}_{n,n}\Omega^{-1}_{-n.-n} + \overline{H}^{*-1})^{-1},$$

and

$$\overline{b_2} = (-\overline{C}\overline{s}_{-n,n}).$$

Note that $\overline{H}^* = diag(\overline{h}_{-n,n})$, and j^{th} element of $\overline{h}_{-n,n}$ is $\overline{\psi_{\omega,jn}}\overline{\phi^2_{\omega,jn}}\overline{\tau^2_\omega}$.

3.1.7 $q(\tau_\omega)$

$$q(\tau_\omega) \sim giG[\frac{n^2-n}{2}(a_\omega - 1), 1, \sum_{j<k} 2(\overline{\omega_{jk}}^2 + \overline{C}_{jj})^{1/2}\frac{1}{\overline{\phi_{\omega,jk}}}], \text{ [3]} \tag{16}$$

[2] To calculate $\overline{s}_{-i,i}$, for $i \neq n$, we need to delete the relevant $(i-1)\times p+1$ to $i\times p$ rows and $(i-1)\times p+1$ to $i\times p$ columns of $\overline{V}$ to construct A_j.

[3] Note that $\overline{C}_{jj}$ changes when k changes.

Let $\chi_\omega = \sum_{j<k} 2(\overline{\omega_{jk}}^2 + \overline{V}_{jj})^{1/2}(\overline{\phi_{\omega,jk}})^{-1}$, we have

$$\overline{\tau_\omega} = \frac{\sqrt{\chi_\omega} K_{\frac{n^2-n}{2}(a_\omega-1)+1}(\sqrt{\chi_\omega})}{K_{\frac{n^2-n}{2}(a_\omega-1)}(\sqrt{\chi_\omega})}$$

$$\overline{\tau_\omega^2} = \overline{\tau_\omega}^2 + \chi_\omega[\frac{K_{\frac{n^2-n}{2}(a_\omega-1)+2}(\sqrt{\chi})}{K_{\frac{n^2-n}{2}(a_\omega-1)}(\sqrt{\chi_\omega})} - (\frac{K_{\frac{n^2-n}{2}(a_\omega-1)+1}(\sqrt{\chi_\omega})}{K_{\frac{n^2-n}{2}(a_\omega-1)}(\sqrt{\chi_\omega})})^2]$$

3.1.8 $q(\psi_{\omega,jn})$

$$q(\psi_{\omega,jn}^{-1}) \sim iG(\sqrt{\frac{\overline{\phi_{\omega,jn}^2}\,\overline{\tau_\omega^2}}{\overline{\omega_{jn}^2} + \overline{C}_{jj}}}, 1), \tag{17}$$

Let $\rho_\omega = \sqrt{\frac{\overline{\phi_{\omega,jn}^2}\,\overline{\tau_\omega^2}}{\overline{\omega_{jn}}^2+\overline{C}_{jj}}}$,

$$\overline{\psi_{\omega,jn}^{-1}} = \rho_\omega$$

and

$$\overline{\psi_{w,jn}} = 1 + 1/\rho_\omega$$

3.1.9 $q(\phi_{\omega,jn})$

$$q(\xi_{\omega,jn}) \sim giG(a_\omega - 1, 1, 2\sqrt{\overline{\omega_{jn}}^2 + \overline{C}_{jj}}) \tag{18}$$

Let $\varpi_\omega = 2\sqrt{\overline{\omega_{jn}}^2 + \overline{C}_{jj}}$, we have

$$\overline{\xi_{\omega,jn}} = \frac{\sqrt{\varpi_\omega} K_{a_\omega}(\sqrt{\varpi_\omega})}{K_{a_\omega-1}(\sqrt{\varpi_\omega})},$$

and

$$var(\xi_{\omega,jn}) = \varpi_\omega\{\frac{K_{a_\omega+1}(\sqrt{\varpi_\omega})}{K_{a_\omega-1}(\sqrt{\varpi_\omega})} - [\frac{K_{a_\omega}(\sqrt{\varpi_\omega})}{K_{a_\omega-1}(\sqrt{\varpi_\omega})}]^2\},$$

where $var(\bullet)$ denotes the variance.

Scaling $\xi_{\omega,jn}$, we have

$$\overline{\phi_{\omega,jn}} = \frac{\overline{\xi_{\omega,jn}}}{\sum_{j<k}\overline{\xi_{\omega,jk}}},$$

and

$$\overline{\phi^2_{\omega,jn}} = \overline{\phi_{\omega,jn}}^2 + \frac{var(\xi_{\omega,jn})}{(\sum_{j<k} \overline{\xi_{\omega,jk}})^2}$$

Thus, the optimal q density of $\phi_{\omega,ij}$ takes the following form:

$$q(\phi_{\omega,jn}) \sim giG(a-1, \sum_{j<k} \overline{\xi_{\omega,jk}}, \frac{2\sqrt{\overline{\omega_{jn}}^2 + (\overline{C}_{jj})^2}}{\sum_{j<k} \overline{\xi_{\omega,jk}}}) \tag{19}$$

3.2 Second-stage VB

We explain the technical details of second-stage VB in Online Appendix B. Below we briefly describe the priors of the parameters and hyperparameters and then provide their optimal q densities.

Let $Z = (\widehat{Y}_{/i} \;\; X_{\bullet,i})'$ and $\theta = [(\Lambda_{i\bullet})' \;\; \tilde{\beta}_i]$.

We elicit hierarchical D-L prior for θ as follows:

$$\theta_j|\tilde{\phi}, \tilde{\tau} \sim DE(\tilde{\phi}_j\tilde{\tau}), \;\; \tilde{\phi}_j \sim Dir(\tilde{a}, ..., \tilde{a}), \;\; \tau \sim G(k\tilde{a}, 1/2) \tag{20}$$

where $k = N - 1 + m_i$.

Next we set a Gamma prior for σ^{-2}:

$$\sigma^{-2} \sim G(\nu, \tilde{S}). \tag{21}$$

The VB optimal densities can be found as follows:

3.2.1 $q(\theta)$

$$q(\theta) \sim N(\overline{\theta}, \overline{\tilde{V}}), \tag{22}$$

where

$$\overline{\tilde{V}} = (\frac{\frac{T}{2} + \nu}{\overline{\tilde{S}}} Z'Z + \tilde{V}^{-1})^{-1}$$

$$\overline{\theta} = (\frac{\frac{T}{2} + \nu}{\overline{\tilde{S}}}) \overline{\tilde{V}} Z' y_i$$

$$\tilde{V}^{-1} = diag(\overline{\tilde{\psi}_1^{-1}}\ \overline{\tilde{\phi}_1^{-2}}\overline{\tilde{\tau}^{-2}}, ..., (\overline{\tilde{\psi}_k^{-1}}\ \overline{\tilde{\phi}_k^{-2}}\overline{\tilde{\tau}^{-2}})$$

3.2.2 $q(\sigma^{-2})$

$$q(\sigma^{-2}) \sim G(\frac{T}{2} + \nu, \overline{\tilde{S}}), \tag{23}$$

where

$$\overline{S} = \frac{1}{2}[||y_i - Z\overline{\theta}||^2 + tr(Z'Z\overline{\tilde{V}})] + \tilde{S}$$

3.2.3 $q(\tilde{\tau})$

$$q(\tilde{\tau}) \sim giG[k\tilde{a} - k, 1, \sum_{j=1}^{k} 2(\overline{\theta}_j^2 + \overline{\tilde{V}}_{jj})^{1/2}\frac{1}{\overline{\tilde{\phi}_j}}], \tag{24}$$

Let $\tilde{\chi} = \sum_{j=1}^{k} 2(\overline{\theta}_j^2 + \overline{\tilde{V}}_{jj})^{1/2}\frac{1}{\overline{\tilde{\phi}_j}}$, we have

$$\overline{\tilde{\tau}} = \frac{\sqrt{\tilde{\chi}}K_{k\tilde{a}-k+1}(\sqrt{\tilde{\chi}})}{K_{k\tilde{a}-k}(\sqrt{\tilde{\chi}})}$$

and

$$\overline{\tilde{\tau}^2} = \overline{\tilde{\tau}}^2 + \chi[\frac{K_{k\tilde{a}-k+2}(\sqrt{\tilde{\chi}})}{K_{k\tilde{a}-k}(\sqrt{\tilde{\chi}})} - (\frac{K_{k\tilde{a}-k+1}(\sqrt{\tilde{\chi}})}{K_{k\tilde{a}-k}(\sqrt{\tilde{\chi}})})^2]$$

3.2.4 $q(\tilde{\psi}_j)$

$$q(\frac{1}{\tilde{\psi}_j}) \sim iG(\sqrt{\frac{\overline{\tilde{\phi}_j^2}\,\overline{\tilde{\tau}^2}}{\overline{\theta}_j^2 + \overline{\tilde{V}}^{jj}}}, 1), \tag{25}$$

Let $\tilde{\rho} = \sqrt{\frac{\overline{\tilde{\phi}_j^2}\,\overline{\tilde{\tau}^2}}{\overline{\theta}_j^2 + \overline{\tilde{V}}^{jj}}}$,

$$\overline{\frac{1}{\tilde{\psi}_j}} = \tilde{\rho}$$

and

$$\overline{\tilde{\psi}}_j = 1 + 1/\tilde{\rho}$$

3.2.5 $q(\tilde{\phi}_j)$

$$q(\tilde{\xi}_j) \sim giG(\tilde{a} - 1, 1, 2\sqrt{\overline{\theta}_j^2 + (\overline{\tilde{V}}_{jj})^2}) \tag{26}$$

Let $\tilde{\varpi} = 2\sqrt{\overline{\theta}_j^2 + (\overline{\tilde{V}}_{jj})^2}$, we have

$$\overline{\tilde{\xi}}_j = \frac{\sqrt{\tilde{\varpi}} K_{\tilde{a}}(\sqrt{\tilde{\varpi}})}{K_{\tilde{a}-1}(\sqrt{\tilde{\varpi}})},$$

and

$$var(\tilde{\xi}_j) = \tilde{\varpi}\{\frac{K_{\tilde{a}+1}(\sqrt{\tilde{\varpi}})}{K_{\tilde{a}-1}(\sqrt{\tilde{\varpi}})} - [\frac{K_{\tilde{a}}(\sqrt{\tilde{\varpi}})}{K_{\tilde{a}-1}(\sqrt{\tilde{\varpi}})}]^2\}.$$

Scaling $\tilde{\xi}$, we have

$$\overline{\tilde{\phi}}_j = \frac{\overline{\tilde{\xi}}_j}{\sum^k \overline{\tilde{\xi}}_j},$$

and

$$\overline{\tilde{\phi}_j^2} = \overline{\tilde{\phi}}_j^2 + \frac{var(\tilde{\xi}_j)}{(\sum^k \overline{\tilde{\xi}}_j)^2}$$

Thus, the optimal q density of $\tilde{\phi}_j$ takes the following form:

$$q(\tilde{\phi}_j) \sim giG[\tilde{a} - 1, \sum^k \overline{\tilde{\xi}}_j, (2\sqrt{\overline{\theta}_j^2 + (\overline{\tilde{V}}_{jj})^2})/(\sum^k \overline{\tilde{\xi}}_j)] \tag{27}$$

4 Monte Carlo Studies

In the Monte Carlo studies, we look into two traditional panel SAR models of various sample sizes. The first model is:

$$y_t = 0.6 W_N y_t + 0.9 x_t + u_t, \tag{28}$$

where to specify W_N, we let each cross-sectional unit be connected with the unit ahead of it and the unit behind, and then normalize W by rows. When conducting Monte Carlo, we generate x_t and u_t independently from $N(0,1)$ and $0.1N(0,1)$, respectively.

The second model is:

$$\begin{aligned} y_{t,1} &= 0.5 y_{t,2} + 0.6 W_{N_1} y_{t,1} + 0.4 W_{N_2} y_{t,2} + 0.9 x_{t,1} + u_{t,1} \\ y_{t,2} &= 0.5 y_{t,1} + 0.4 W_{N_1} y_{t,1} + 0.6 W_{N_2} y_{t,2} + 0.9 x_{t,2} + u_{t,2} \end{aligned} \tag{29}$$

where $y_{t,1}$ is a vector of half of the N dependent variables observed at time t, and $y_{t,2}$ is

the other half. When setting W_{N_1} and W_{N_2}, we assume a variable in $y_{t,1}$ is only spatially related with the unit ahead of it and the unit behind, likewise a variable in $y_{t,2}$. Both W_{N_1} and W_{N_2} are normalized by rows. In Monte Carlo, we generate each element of $x_{t,1}$ and $x_{t,2}$ independently from $N(0,1)$, then each element of $u_{t,1}$ and $u_{t,2}$ independently from $0.1N(0,1)$.[4]

For both (28) and (29), the sample sizes considered are $N = 30$ and $T = 20$, $N = 30$ and $T = 80$, $N = 50$ and $T = 30$, $N = 50$ and $T = 100$, $N = 100$ and $T = 50$, and $N = 100$ and $T = 200$. For each case, we conduct 1000 Monte Carlo replications and use the changes in parameters instead of that of ELBO to check whether two-stage VB has converged.

Tables 1-2 report the Monte Carlo results of parameter Λ and $\tilde{\beta}$ for model (28) with $N = 30$ and $T = 80$ as well as $N = 30$ and $T = 20$. To save space, we only report the mean and standard deviations, the latter in parenthesis, of the empirical distributions of the first and last five rows of Λ, which are associated with $y_{t,1}, ..., y_{t,5}$ and $y_{t,26}, ..., y_{t,30}$ and relegate those of the other rows to the Online Appendix C.

Tables 3-4 report the Monte Carlo results of parameter Λ and $\tilde{\beta}$ for model (29) with $N = 30$. Again, we only include the first and last five rows of Λ, but all the elements in $\tilde{\beta}$, leaving the rest of Λ to Online Appendix C.

Results in Tables 1-4 provide strong evidence that the two-stage VB is able to recover the true parameters in the data generating process, especially when $T \gg N$. When $T \ll N$, two-stage VB estimates have larger biases and larger empirical standard deviations, which can be reduced by setting tighter priors. More important, there is clear evidence that the true spatial connections can be identified regardless of the length of the panel, long or short. Additional results presented in Online Appendix C, including those for models with $N = 50$ and $N = 100$, give further support to the effectiveness of two-stage VB.

We use high-performance computing (HPC) services for estimation. All computations are done using 1 compute node and 4 processor core. On average, each Monte Carlo replication for model (28) where $N = 30$ and $T = 80$ takes about 20 seconds, while for the same model, when $N = 30$ and $T = 20$ it takes about 60 seconds. For model (29), it again takes 20 seconds when $N = 30$ and $T = 80$. However, estimating model (29) takes 3 minutes when $N = 30$ and $T = 20$. The same pattern can be observed when estimating models where

[4]We have also experimented on W_N, W_{N_1} and W_{N_2} of other forms. In addition, we have looked into models with different coefficients, $u_{t,1} \sim 0.1N(0,1)$ and $u_{t,2} \sim 0.3N(0,1)$. The Monte Carlo results provide further evidence that two-stage VB method works well. To save space, we do not report them in the paper.

$N = 50$ and $N = 100$, that is: two-stage VB takes much longer to converge when $N \gg T$, and the larger the number of nonzero true parameters, the longer the estimation takes. To give a flavour of how fast two-stage VB converges, we would like to mention that, when allowing for parallel computing, each Monte Carlo replication of model (28) where $N = 100$ and $T = 50$ takes about 20 minutes, and that number reduces to 5 minutes for the same model where $N = 100$ and $T = 200$.

5 Conclusion and Discussions

In applied work, if the spatial weights matrix set a priori were to be far from its true value, the empirical analysis using SAR models would be misleading. In this paper, we have developed a two-stage VB approach to estimating panel SAR models with unknown spatial weights matrices so as to let the data speak. Monte Carlo experiments show that our two-stage VB is rather fast, and it can recover the spatial impacts well for both the long and short panels.

For short panels, two-stage VB with D-L priors tends to bring about larger biases and higher standard deviations. That can be partly addressed by setting tighter priors. Another intuitive way to rectify the problem is to re-run two-stage VB, this time setting more realistic priors based on the results from running two-stage VB with D-L priors. We'll leave that for future research.

The two-stage VB with D-L priors approach can be easily extended by including other popular priors such as the Lasso, Horseshoe and spike and a slab priors. Furthermore, the success of two-stage VB shows the potential of combining VB with more sophisticated methods, such as three-stage least squares, full information likelihood and GMM to estimate panel SAR models, especially those involving $N \gg T$.

References

[1] Ahrens, A., and A. Bhattacharjee (2015). Two-step lasso estimation of the spatial weights matrix. Econometrics, 3, 128–155.

[2] Anselin, L. (1988). Spatial Econometrics: Methods and Models. Kluwer Academic, Dordrecht.

[3] Baltagi, B., S.H. Song and W. Koh (2003). Testing panel data regression models with spatial error correlation. Journal of Econometrics, 117, 123-150

[4] Baltagi, B. H., P. Egger and M. Pfaffermayr (2013). A generalized spatial panel data model with random effects. Econometric reviews, 32(5-6), 650-685.

[5] Basak, G. K., A. Bhattacharjee and S. Das (2018). Causal ordering and inference on acyclic networks. Empirical Economics, 55, 213–232.

[6] Bhattacharya A., D. Pati, N. S. Pillai and D. B. Dunson (2015). Dirichlet–Laplace priors for optimal shrinkage. Journal of the American Statistical Association, 110 (512), 1479-1490

[7] Blei, D.M., A. Kucukelbir and J.D. McAuliffe (2017) Variational inference: A review for statisticians. Journal of the American Statistical Association, 112 (518), 859-877.

[8] Case, A. (1991). Spatial Patterns in Household Demand. Econornetrica, 59, 953-965.

[9] Cliff, A. D. and J. K. Ord. (1973). Spatial Autocorrelation. London: Pion.

[10] Fox, J. (1979). Simultaneous equation models and two-stage least squares. Sociological methodology, 10, 130-150.

[11] Gefang, D., G. Koop and A. Poon (2020). Computationally efficient inference in large Bayesian mixed frequency VARs. Economics Letters, 191, 109120.

[12] Gefang, D., G. Koop and A. Poon (2022). Forecasting using variational Bayesian inference in large vector autoregressions with hierarchical shrinkage. International Journal of Forecasting.

[13] Krock, M., W. Kleiber and S. Becker (2021). Nonstationary modeling with sparsity for spatial data via the basis graphical lasso. Journal of Computational and Graphical Statistics, 30(2), 375-389.

[14] Lam, C., and Souza, P. C. (2019). Estimation and selection of spatial weight matrix in a spatial lag model. Journal of Business and Economic Statistics, 38, 693–710.

[15] Lee, L-F. and J. Yu (2010). Estimation of spatial autoregressive panel data models with fixed effects. Journal of Econometrics, 154, 2, 165-185.

[16] Li, Y., B .A. Craig and A. Bhadra (2019). The graphical horseshoe estimator for inverse covariance matrices. Journal of Computational and Graphical Statistics, 28(3),

747-757.

[17] Liu, X. and P. Saraiva (2019). GMM estimation of spatial autoregressive models in a system of simultaneous equations with heteroskedasticity. Econometric Reviews, 38:4, 359-385.

[18] Loaiza-Maya, R., M. S. Smith, D. J. Nott and P. J. Danaher (2021). Fast and accurate variational inference for models with many latent variables. Journal of Econometrics.

[19] Makalic, E. and D. F. Schmidt (2016). A simple sampler for the horseshoe estimator. IEEE Signal Processing Letters 23, 179–182.

[20] Ormerod, J.T. and M. P. Wand (2010). Explaining variational approximations. The American Statistician, 64 (2), 140-153.

[21] Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society, Series B, 58, 267-288.

[22] Wang, H. (2012). Bayesian graphical lasso models and efficient posterior computation. Bayesian Analysis, 7(4), 867-886.

[23] Yang, K. and L. F. Lee (2017). Identification and QML estimation of multivariate and simultaneous equations spatial autoregressive models. Journal of Econometrics 196:196–214.

[24] Zellner, A., and H. Theil (1962). Three-Stage Least Squares: Simultaneous Estimation of Simultaneous Equations. Econometrica: Journal of the Econometric Society, 54-78.

Table 1: Estimates of the first and last five rows of Λ for model (28), where $N = 30$

T=80																													
	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
0.30		0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(0.02)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
0.00	0.30		0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(0.02)	(0.02)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
0.00	0.00	0.30		0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
0.00	0.00	0.00	0.30		0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(0.02)	(0.02)	(0.02)	(0.01)		(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30		0.30	0.00	0.00	0.00
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)	(0.02)
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30		0.30	0.00	0.00
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30		0.30	0.00
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30		0.00
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)		(0.02)
0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30	
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	
T=20																													
	0.27	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.03	0.27	
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	
0.27		0.27	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	
(0.08)		(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.09)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	
0.04	0.27		0.27	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	-0.01	0.01	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	0.00	
(0.09)	(0.08)		(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	
0.00	0.03	0.27		0.27	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
(0.08)	(0.08)	(0.08)		(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	
0.00	0.00	0.03	0.27		0.27	0.04	0.00	0.00	0.00	0.00	0.00	-0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	
(0.08)	(0.08)	(0.08)	(0.08)		(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.04	0.27		0.27	0.03	0.00	0.00	
(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)		(0.08)	(0.08)	(0.08)	(0.08)	
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.03	0.27		0.27	0.04	0.00			
(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.07)		(0.08)	(0.09)	(0.08)	
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	0.03	0.27		0.27	0.04	0.00		
(0.09)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	(0.08)	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)		(0.08)	(0.09)	(0.08)		
0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.04	0.27		0.27		
(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)			
0.27	0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	-0.01	0.00	0.04	0.27				
(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.09)	(0.09)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)	(0.09)	(0.08)	(0.08)	(0.09)	(0.08)				

Table 2: Estimates of $\tilde{\beta}$ for model (28), where $N = 30$

T=80																													
0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
T=20																													
0.65	0.66	0.65	0.66	0.65	0.65	0.65	0.65	0.65	0.65	0.66	0.65	0.65	0.65	0.65	0.65	0.66	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65	0.65
(0.08)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)	(0.07)	(0.07)	(0.07)	(0.08)	(0.07)	(0.08)	(0.08)

Table 3: Estimates of the first and last five rows of Λ for model (29), where $N = 30$

T=80																													
	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30	0.50	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
0.30		0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.50	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(0.02)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
0.00	0.30		0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.50	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(0.02)	(0.02)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
0.00	0.00	0.30		0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.50	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
0.00	0.00	0.00	0.30		0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.50	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
(0.02)	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
...																													
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.50	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30		0.30	0.00	0.00	0.00
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)	(0.02)
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.50	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30		0.30	0.00	0.00
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)	(0.02)
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.50	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30		0.30	0.00
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)	(0.02)
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.50	0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30		0.30
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)		(0.02)
0.20	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.20	0.50	0.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.30	
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	
T=20																													
	0.25	-0.02	-0.02	0.00	0.00	-0.01	0.00	0.00	0.00	0.01	0.00	-0.02	-0.02	0.26	0.42	0.19	-0.02	-0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	-0.04	0.19
	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
0.25		0.25	-0.02	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	-0.02	-0.02	0.18	0.43	0.19	-0.03	-0.03	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	-0.01	-0.02	-0.03
(0.10)		(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
-0.02	0.25		0.25	-0.03	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	-0.03	-0.02	0.18	0.43	0.19	-0.03	-0.03	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	-0.02
(0.09)	(0.09)		(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)
-0.02	-0.02	0.25		0.25	-0.02	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	-0.03	0.18	0.43	0.18	-0.03	-0.02	-0.01	0.01	0.00	0.00	0.00	0.00	0.00
(0.10)	(0.09)	(0.10)		(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)
0.00	-0.03	-0.03	0.25		0.25	-0.02	-0.03	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	-0.03	0.19	0.42	0.19	-0.02	-0.02	0.00	0.00	0.00	0.00	0.00	0.00
(0.10)	(0.09)	(0.10)	(0.10)		(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)
...																													
0.00	0.00	0.00	0.00	0.00	0.00	-0.01	-0.03	-0.03	0.18	0.42	0.18	-0.03	-0.02	0.00	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	-0.02	-0.02	0.25		0.25	-0.03	-0.02	0.00
(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)		(0.09)	(0.10)	(0.10)	(0.10)
0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	-0.02	0.19	0.43	0.18	-0.03	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	-0.03	0.25		0.25	-0.02	-0.02
(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)		(0.09)	(0.09)	(0.10)
-0.02	-0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	-0.03	0.19	0.43	0.18	-0.03	-0.03	0.00	0.00	0.00	-0.01	0.00	0.00	0.00	0.00	-0.03	-0.02	0.25		0.25	-0.02
(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)		(0.09)	(0.10)
-0.03	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	-0.03	0.18	0.43	0.19	-0.03	-0.02	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	-0.02	-0.02	0.25		0.25
(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)		(0.09)
0.19	-0.03	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.01	-0.02	-0.03	0.19	0.42	0.25	-0.02	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.02	-0.02	0.25	
(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.09)	(0.09)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	

Table 4: Estimates of $\tilde{\beta}$ for model (29), where $N = 30$

T=80																													
0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90	0.90
(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
T=20																													
0.58	0.59	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.58	0.59	0.58	0.58	0.58	0.58	0.58	0.59	0.58	0.58	0.58	0.58	0.59	0.59	0.59	0.58	0.58	0.59	0.59	0.58	0.59
(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.09)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)	(0.10)