# ECDNETOR 

Make Your Publications Visible.

# Working Paper <br> GM estimation of higher order spatial autoregressive processes in panel data error component models 

CESifo Working Paper, No. 2301

## Provided in Cooperation with:

Ifo Institute - Leibniz Institute for Economic Research at the University of Munich

Suggested Citation: Badinger, Harald; Egger, Peter (2008) : GM estimation of higher order spatial autoregressive processes in panel data error component models, CESifo Working Paper, No. 2301, Center for Economic Studies and ifo Institute (CESifo), Munich

This Version is available at:
https://hdl.handle.net/10419/26346

## Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

## Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

# GM Estimation of Higher Order Spatial Autoregressive Processes in Panel Data Error Component Models 

HARALD BADINGER<br>Peter EgGer

CESifo Working Paper No. 2301
Category 10: Empirical and Theoretical Methods
May 2008

An electronic version of the paper may be downloaded

- from the SSRN website: www.SSRN.com
- from the RePEc website: www.RePEc.org
- from the CESifo website: www.CESifo-group.org/wp


# GM Estimation of Higher Order Spatial Autoregressive Processes in Panel Data Error Component Models 


#### Abstract

This paper presents a generalized moments (GM) approach to estimating an $R$-th order spatial regressive process in a panel data error component model. We derive moment conditions to estimate the parameters of the higher order spatial regressive process and the optimal weighting matrix required to achieve asymptotic efficiency. We prove consistency of the proposed GM estimator and provide Monte Carlo evidence that it performs well also in reasonably small samples.


JEL Code: C13, C21, C23.
Keywords: spatial models, panel data models, error component models.

Harald Badinger* Department of Economics<br>Vienna University of Economics and<br>Business Administration<br>Althanstrasse 39-45<br>1090 Vienna<br>Austria<br>harald.badinger@wu-wien.ac.at

May 2008
*This paper was written during my stay as Erwin-Schrödinger Research Fellow at the Center for Economic Studies (CES), University of Munich. Financial support by the Austrian Science Fund (FWF) is gratefully acknowledged.

## I. Introduction

A rapidly growing amount of recent economic research focuses on the empirical modelling of variables with cross-sectional interdependence. Theoretical rationales for such work are almost ubiquitous in economics: game theory for long considered strategic interdependence among agents (firms or individuals) in their behaviour-examples are the quantity setting of Cournot firms, the price setting of Bertrand firms, investment decisions in research and development of oligopolistically competing enterprises, tax competition among regional or national jurisdictions, or group behaviour of individuals; a second reason for (cross-sectional) interdependence are general equilibrium effects and the propagation of 'local' shocks through economic systems that are interrelated, e.g., by trade or factor mobility.

An attractive way of allowing for interdependence between cross-sectional units in empirical models is by means of so-called spatial econometric methods (see Anselin, 1988, for an early treatment, using the maximum-likelihood approach). The latter typically assume that there is some known channel of relations among cross-sectional units, e.g., 'space' in terms of geographical distance or adjacency but also input-output relationships or trade flows. A large class of existing models allows for spatially autoregressive residuals (SAR). There, interdependence occurs among the unobservable variables in the model. Anselin (2003) provides a typology of spatial econometric models. However, existing models seem restrictive from an applied researcher's point of view, since the SAR process is typically assumed to be of first order, referred to as SAR(1) (see Anselin, 1988; Kelejian and Prucha, 1999; or Kapoor, Kelejian, and Prucha, 2007). In the latter case, the researcher may not allow for a flexible decay of interdependence in 'space', but spatial relationships need to be captured by a single parameter, given the assumed channel or matrix of interdependence among crosssectional units.

This paper formulates a GM estimator for the case of a SAR process of order $R$, i.e., $\operatorname{SAR}(R)$, for panel data with a large cross-section that is repeatedly observed over a relatively smaller number of time periods. ${ }^{1}$ In particular, we generalize the existing GM approach to estimating the SAR(1) parameter in panel data error component models by Kapoor, Kelejian, and Prucha (2007) to the case of a spatial regressive error process of arbitrary order $R$. Such a framework allows the applied econometrician to study the strength of interdependence more flexibly than in existing $\operatorname{SAR}(1)$ models. For instance, with the suggested model one may allow first, second, and higher orders of bands of neighbours to exert a different impact on each other, given their 'spatial' distance, for various economic problems (see Kelejian and Robinson, 1992; Bell and Boecksteal, 2000; and Cohen and Morrison Paul, 2007, for applications with

[^0]cross-sectional data). Similarly, one may allow for several alternative channels or concepts of interdependence, e.g. intra-industry and inter-industry spillovers (see Badinger and Egger, 2008, for a cross-section application). Generally, economic, socio-economic, geographical, demographic (e.g., cultural, lingual), or political distance may play a role explicitly and simultaneously.

Using a higher order spatial regressive process allows for a more flexible specification and thus better approximation of the functional form of the decay of interdependence in some predefined space. Moreover, it enables an empirical assessment of the relative importance of alternative channels of interdependence.

The remainder of the paper is organized as follows. Section II briefly summarizes the basic model specification and introduces some notation. Section III derives the moment conditions for the GM estimators of a $\operatorname{SAR}(R)$ process and the optimal weighting matrix. Section IV demonstrates consistency of the GM estimators and provides Monte Carlo evidence to illustrate the small sample performance. The last section concludes with a summary of the key findings.

## II. Basic model specification and notation

The basic set-up of the error components model with spatially correlated error terms represents a generalization of the framework of Kapoor, Kelejian, and Prucha (2007), henceforth referred to as KKP. The model comprises $i=1, \ldots, N$ cross-sectional units and $t=1, \ldots, T$ time periods. For time period $t$, the model reads

$$
\begin{equation*}
y_{N}(t)=X_{N}(t) \beta+u_{N}(t) \tag{1a}
\end{equation*}
$$

where $y_{N}(t)$ is an $N \times 1$ vector with cross-sectional observations of the dependent variable in year $t, X_{N}(t)$ is an $N \times K$ matrix of non-stochastic explanatory variables, with $K$ denoting the number of explanatory variables in the model including the constant, and $u_{N}(t)$ is an $N \times 1$ vector of disturbances which is generated by the following $\operatorname{SAR}(R)$ process:

$$
\begin{align*}
& u_{N}(t)=\sum_{m=1}^{R} \rho_{m} W_{m, N} u_{N}(t)+\varepsilon_{N}(t),  \tag{1b}\\
& \varepsilon_{N}(t)=\mu_{N}+v_{N}(t) \tag{1c}
\end{align*}
$$

where $\rho_{m}$ and $W_{m, N}$ denote the time-invariant, unknown parameter and the known $N \times N$ matrix of spatial interdependence for the $m$-th band or concept of interdependence, respectively. The structure of spatial correlation is determined by the $R$ different, timeinvariant $N \times N$ matrices $W_{m, N}$, whose elements $\mathrm{w}_{i, N}$ are often (but need not be) specified as a
decreasing function of geographical distance between the cross-sectional units $i$ and $j$. Using a higher order process allows the strength of spatial interdependence (reflected in the spatial regressive parameters $\rho_{m}, m=1, \ldots, R$ ) to vary across a fixed number of $R$ subsets of relations between cross-sectional units. Obviously, model (1) nests the specification by KKP as a special case for $R=1$. $\varepsilon_{N}(t)$ is an error term which consists of two components, $\mu_{N}$ and $v_{N}(t)$. As indicated by the notation, $\mu_{N}$ is time-invariant while $v_{N}(t)$ is not. The typical elements of $\varepsilon_{N}(t), \mu_{N}$, and $v_{N}(t)$ are the scalars $\varepsilon_{i t, N}, \mu_{i, N}$, and $v_{i t, N}$, respectively.

Let us now stack the observations for all time periods such that $t$ is the slow index and $i$ is the fast index with all vectors and matrices, respectively. Then, the model reads

$$
\begin{equation*}
y_{N}=X_{N} \beta+u_{N}, \tag{2a}
\end{equation*}
$$

where $y_{N}=\left[y_{N}^{\prime}(1), \ldots, y_{N}^{\prime}(T)\right]^{\prime}$ is the $N T \times 1$ vector of observations on the dependent variable. The regressor matrix $X_{N}=\left[X_{N}^{\prime}(1), \ldots, X_{N}^{\prime}(T)\right]^{\prime}$ is of dimension $N T \times K$. Generalizing the specification in KKP (p. 100), the $N T \times 1$ vector of error terms $u_{N}=\left[u_{N}^{\prime}(1), \ldots, u_{N}^{\prime}(T)\right]$ for a spatial regressive process of order $R$ reads

$$
\begin{equation*}
u_{N}=\sum_{m=1}^{R} \rho_{m}\left(I_{T} \otimes W_{m, N}\right) u_{N}+\varepsilon_{N} \tag{2b}
\end{equation*}
$$

where $I_{T}$ is an identity matrix of dimension $T \times T$. The $N T \times 1$ vector $\varepsilon_{N}=\left[\varepsilon_{N}^{\prime}(1), \ldots, \varepsilon_{N}^{\prime}(T)\right]^{\prime}$ is specified as

$$
\begin{equation*}
\varepsilon_{N}=\left(e_{T} \otimes I_{N}\right) \mu_{N}+v_{N} . \tag{2c}
\end{equation*}
$$

The $N \times 1$ vector of unit specific error components is given by $\mu_{N}=\left[\mu_{1}, \mu_{2}, \ldots, \mu_{N}\right]^{\prime}$. Finally, $I_{N}$ is an identity matrix of dimension $N \times N$ and $e_{T}$ is a unit vector of dimension $T \times 1$. Notice that, in light of (2b), the error term can also be written as

$$
\begin{equation*}
\varepsilon_{N}=u_{N}-\sum_{m=1}^{R} \rho_{m}\left(I_{T} \otimes W_{m, N}\right) u_{N} . \tag{3}
\end{equation*}
$$

It follows that

$$
\begin{equation*}
u_{N}=\left[I_{T} \otimes\left(I_{N}-\sum_{m=1}^{R} \rho_{m} W_{m, N}\right)^{-1}\right] \varepsilon_{N}, \tag{4a}
\end{equation*}
$$

and

$$
\begin{equation*}
y_{N}=X_{N} \beta+\left[I_{T} \otimes\left(I_{N}-\sum_{m=1}^{R} \rho_{m} W_{m, N}\right)^{-1}\right] \varepsilon_{N} \tag{4b}
\end{equation*}
$$

The following assumptions are maintained throughout the analysis.

## Assumption 1.

Let $T$ be a fixed positive integer. (a) For all $1 \leq t \leq T$ and $1 \leq i \leq N, N \geq 1$, the error components $v_{i t, N}$ are identically distributed with zero mean and variance $\sigma_{v}^{2}, 0<\sigma_{v}^{2}<b_{v}<$ $\infty$, and finite fourth moments. In addition, for each $N \geq 1$ and $1 \leq t \leq T, 1 \leq i \leq N$ the error components $v_{i t, N}$ are independently distributed. (b) For all $1 \leq i \leq N, N \geq 1$ the unit specific error components $\mu_{i, N}$ are identically distributed with zero mean and variance $\sigma_{\mu}^{2}, 0<\sigma_{\mu}^{2}<$ $b_{\mu}<\infty$, and finite fourth moments. Moreover, for each $N \geq 1$ and $1 \leq i \leq N$ the unit-specific error components $\mu_{i, N}$ are independently distributed. (c) The processes $\left\{v_{i t, N}\right\}$ and $\left\{\mu_{i, N}\right\}$ are independent of each other. Assumption 1 is exactly identical to the first-order case considered by KKP.

## Assumption 2.

(a) All diagonal elements of $W_{r, N}$ are zero for $r=1, \ldots, R$.
(b) $\sum_{m=1}^{R}\left|\rho_{m}\right|<1$.
(c) The matrix ( $I-\sum_{m=1}^{R} \rho_{m} W_{m, N}$ ) is non-singular.

Assumption (2c) ensures that $u_{N}$ and $y_{N}$ are uniquely identified through equations (4a) and (4b). Assumption (2b) places a restriction on the admissible parameter space. With rownormalized weights matrices typically used in applied work, Assumption (2c) is implied by Assumption (2b). ${ }^{2}$

Assumptions 1 and 2 imply that

$$
\begin{align*}
& E\left(\varepsilon_{i t, N} \varepsilon_{j s, N}\right)=\sigma_{\mu}^{2}+\sigma_{v}^{2} \text { for } i=j \text { and } t=s,  \tag{5a}\\
& E\left(\varepsilon_{i t, N} \varepsilon_{j s, N}\right)=\sigma_{\mu}^{2} \text { for } i=j \text { and } t \neq s \text {, and }  \tag{5b}\\
& E\left(\varepsilon_{i t, N} \varepsilon_{j s, N}\right)=0 \text { otherwise. } \tag{5c}
\end{align*}
$$

[^1]As a consequence, the variance-covariance matrix of the stacked error term $\varepsilon_{N}$ is given by

$$
\begin{equation*}
\Omega_{\varepsilon, N}=E\left[\varepsilon_{N} \varepsilon_{N}^{\prime}\right]=\sigma_{\mu}^{2}\left(J_{T} \otimes I_{N}\right)+\sigma_{v}^{2} I_{N T} \tag{6a}
\end{equation*}
$$

where $J_{T}=e_{T} e_{T}^{\prime}$ is a $T \times T$ matrix with unit elements and $I_{N T}$ is an identity matrix of dimension $N T \times N T$. Equation (6a) can also be written as

$$
\begin{equation*}
\Omega_{\varepsilon, N}=\sigma_{v}^{2} Q_{0, N}+\sigma_{1}^{2} Q_{1, N} \tag{6b}
\end{equation*}
$$

where $\sigma_{1}^{2}=\sigma_{v}^{2}+T \sigma_{\mu}^{2}$. The two matrices $Q_{0, N}$ and $Q_{1, N}$, which are central to the estimation of error component models and the moment conditions of the GM estimator, are defined as

$$
\begin{align*}
& Q_{0, N}=\left(I_{T}-\frac{J_{T}}{T}\right) \otimes I_{N}, \text { and }  \tag{7a}\\
& Q_{1, N}=\frac{J_{T}}{T} \otimes I_{N} . \tag{7b}
\end{align*}
$$

Notice that $Q_{0, N}$ and $Q_{1, N}$ are both of order $N T \times N T$, and they are symmetric, idempotent, orthogonal to each other, and sum up to $I_{N T}$. Pre-multiplying an $N T \times 1$ vector, e.g., $\varepsilon_{N}$, with $Q_{0, N}$ transforms its elements into deviations from cross-section specific sample means taken over time. Pre-multiplying a vector by $Q_{1, N}$ transforms the observations into cross-section specific sample means. The elements of $Q_{0, N} \varepsilon_{N}$ and $Q_{1, N} \varepsilon_{N}$ are then given by $\varepsilon_{i t, N}-1 / T \sum_{t=1}^{T} \varepsilon_{i t, N}$ and $1 / T \sum_{t=1}^{T} \varepsilon_{i t, N}$, respectively. The matrices $Q_{0, N}$ and $Q_{1, N}$ have the following properties, which are repeatedly used in the subsequent derivations (KKP, p. 101):

$$
\begin{align*}
& \operatorname{tr}\left(Q_{0, N}\right)=N(T-1), \operatorname{tr}\left(Q_{1, N}\right)=N, Q_{0, N}\left(e_{T} \otimes I_{N}\right)=0, Q_{1, N}\left(e_{T} \otimes I_{N}\right)=\left(e_{T} \otimes I_{N}\right),  \tag{8}\\
& \left(I_{T} \otimes D_{N}\right) Q_{0, N}=Q_{0, N}\left(I_{T} \otimes D_{N}\right),\left(I_{T} \otimes D_{N}\right) Q_{1, N}=Q_{1, N}\left(I_{T} \otimes D_{N}\right), \\
& \operatorname{tr}\left[\left(I_{T} \otimes D_{N}\right) Q_{0, N}\right]=(T-1) \operatorname{tr}\left(D_{N}\right), \text { and } \operatorname{tr}\left[\left(I_{T} \otimes D_{N}\right) Q_{1, N}\right]=\operatorname{tr}\left(D_{N}\right),
\end{align*}
$$

where $D_{N}$ is an arbitrary $N \times N$ matrix.

Finally, note that the variance-covariance matrix of $u_{N}$ is given by

$$
\begin{align*}
& \Omega_{u, N}=E\left[u_{N} u_{N}^{\prime}\right]=\left[I_{T} \otimes\left(I_{N}-\sum_{m=1}^{R} \rho_{m} W_{m, N}\right)^{-1}\right] \Omega_{\varepsilon, N}\left[I_{T} \otimes\left(I_{N}-\sum_{m=1}^{R} \rho_{m} W_{m, N}^{\prime}\right)^{-1}\right], \text { and }  \tag{9a}\\
& E\left[u_{N}(t) u_{N}^{\prime}(t)\right]=\left(\sigma_{\mu}^{2}+\sigma_{v}^{2}\right)\left(I_{N}-\sum_{m=1}^{R} \rho_{m} W_{m, N}\right)^{-1}\left(I_{T} \otimes\left(I_{N}-\sum_{m=1}^{R} \rho_{m} W_{m, N}^{\prime}\right)^{-1} .\right. \tag{9b}
\end{align*}
$$

## III. GM Estimation of a $\operatorname{SAR}(R)$ model

## 1. Moment conditions

KKP (p. 103) use six moment conditions to derive a generalized moments (GM) estimator for a first-order spatial regressive process $(\operatorname{SAR}(R)$, with $R=1)$. With an $R$-th order process (SAR $(R)$, with $R>1$ ), the GM estimators of the parameters $\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}$, and $\sigma_{1}^{2}$ can be obtained by recognizing that - under Assumptions 1 and 2 - the moment conditions used by KKP hold for each matrix $W_{r, N}, r=1, \ldots, R$. In particular, we define for each $W_{r, N}$

$$
\begin{equation*}
\bar{\varepsilon}_{r, N}=\left(I_{T} \otimes W_{r, N}\right) \varepsilon_{N}=\left(I_{T} \otimes W_{r, N}\right)\left(u_{N}-\sum_{m=1}^{R} \rho_{m}\left(I_{T} \otimes W_{m, N}\right) u_{N}\right) . \tag{10}
\end{equation*}
$$

A word on notation is in order here. In equation (10), subscript $r$ has been introduced together with $m$ to indicate that, with higher order spatial processes, $W_{r, N}$ and $W_{m, N}$ meet in $\bar{\varepsilon}_{r, N}$. While we will use index $r$ to refer to the moment condition involving matrix $W_{r, N}$ in equation (10), index $m$ is required in equation (10) for the summation over the terms $\rho_{m} W_{m, N}$; for summations like that as, e.g., in Assumption 2(b), we use index $m$ throughout. Moreover, index $r$ is used when there is no danger of confusion as in Assumption 2(a), for example. The moment conditions are then given by

$$
\begin{align*}
& \mathrm{M}_{\mathrm{a}} \quad E\left[\frac{1}{N(T-1)} \varepsilon_{N}^{\prime} Q_{0, N} \varepsilon_{N}\right]=\sigma_{v}^{2},  \tag{11}\\
& \mathrm{M}_{1, r} \quad E\left[\frac{1}{N(T-1)} \bar{\varepsilon}_{r, N}^{\prime} Q_{0, N} \bar{\varepsilon}_{r, N}\right]=\sigma_{v}^{2} \frac{1}{N} \operatorname{tr}\left(W_{r, N}^{\prime} W_{r, N}\right), \\
& \mathrm{M}_{2, r} \quad E\left[\frac{1}{N(T-1)} \bar{\varepsilon}_{r, N}^{\prime} Q_{0, N} \varepsilon_{N}\right]=0, \\
& \mathrm{M}_{\mathrm{b}} \quad E\left[\frac{1}{N} \varepsilon_{N}^{\prime} Q_{1, N} \varepsilon_{N}\right]=\sigma_{1}^{2}, \text { where } \sigma_{1}^{2}=\sigma_{v}^{2}+T \sigma_{\mu}^{2}, \\
& \mathrm{M}_{3, r} \quad E\left[\frac{1}{N} \bar{\varepsilon}_{r, N}^{\prime} Q_{1, N} \bar{\varepsilon}_{r, N}\right]=\sigma_{1}^{2} \frac{1}{N} \operatorname{tr}\left(W_{r, N}^{\prime} W_{r, N}\right), \\
& \mathrm{M}_{4, r} \quad E\left[\frac{1}{N} \bar{\varepsilon}_{r, N}^{\prime} Q_{1, N} \varepsilon_{N}\right]=0 .
\end{align*}
$$

The moment conditions associated with matrix $W_{r, N}(r=1, \ldots, R)$ through (10) are indexed with subscripts 1 to 4 . The remaining two moment conditions, which do not depend on $r$, are denoted as $\mathrm{M}_{\mathrm{a}}$ and $\mathrm{M}_{\mathrm{b}}$. For an $R$-th order process, we thus have $(4 R+2)$ moment conditions.

Substituting equations (4), (10), and (1c) into the $4 R+2$ moment conditions (11) yields a ( $4 R$ +2 ) equation system in $\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}$, and $\sigma_{1}^{2}$, which can be written as

$$
\begin{equation*}
\gamma_{N}-\Gamma_{N} \alpha=0 \tag{12}
\end{equation*}
$$

where $\alpha$ is a $[2 R+R(R-1) / 2+2] \times 1$ vector, given by

$$
\alpha=\left(\rho_{1}, \ldots, \rho_{R}, \rho_{1}^{2}, \ldots, \rho_{R}^{2}, \rho_{1} \rho_{2}, \ldots, \rho_{1} \rho_{R}, \ldots, \rho_{R-1} \rho_{R}, \sigma_{v}^{2}, \sigma_{1}^{2}\right)^{\prime},
$$

i.e., $\alpha$ contains $R$ linear terms $\rho_{m}(m=1, \ldots, R), R$ quadratic terms $\rho_{m}^{2}(m=1, \ldots, R)$, $R(R-1) / 2$ cross products $\rho_{m} \rho_{l}(m=1, \ldots, R-1, l=m+1, \ldots, R)$, as well as $\sigma_{v}^{2}$ and $\sigma_{1}^{2}$. For later reference, we define the vector of spatial regressive parameters $\rho=\left(\rho_{1}, \ldots, \rho_{R}\right)^{\prime}$ and the (row) vector of all parameters as $\theta=\left(\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}, \sigma_{1}^{2}\right)$.
$\gamma_{N}$ is a $(4 R+2) \times 1$ vector with elements $\left[\gamma_{i}\right], i=1, \ldots,(4 R+2)$, and $\Gamma_{N}$ is a $(4 R+2) \times[2 R+R(R-1) / 2+2] \quad$ matrix with elements $\quad\left[\gamma_{i j}\right], \quad i=1, \ldots,(4 R+2)$, $j=1, \ldots,[2 R+R(R-1) / 2+2]$, whose elements will be defined below. Subscript $N$ is dropped from the elements of $\gamma_{N}$ and $\Gamma_{N}$ for simplicity of notation here. The row-index of the elements $\gamma_{N}$ and $\Gamma_{N}$ will be chosen such that the equation system (12) has the following order. The first four rows correspond to the moment restrictions $\mathrm{M}_{1,1}$ to $\mathrm{M}_{4,1}$ associated with matrix $W_{1, N}$ through (10); row five to eight correspond to $\mathrm{M}_{1,2}$ to $\mathrm{M}_{4,2}$ associated with matrix $W_{2, N}$, and so fourth; rows ( $R-4$ ) to $4 R$ correspond to the $\mathrm{M}_{1, R}$ to $\mathrm{M}_{4, R}$ associated with the matrix $W_{R, N}$. Finally, rows $4 R+1$ and $4 R+2$ correspond to the moment conditions $\mathrm{M}_{\mathrm{a}}$ and $\mathrm{M}_{\mathrm{b}}$, respectively, which do not depend on $r$.

The sample analogue to equation system (12) is given by

$$
\begin{equation*}
g_{N}-G_{N} \alpha=\vartheta_{N}(\theta) \tag{13}
\end{equation*}
$$

where the elements of $g_{N}$ and $G_{N}$ are equal to those of $\gamma_{N}$ and $\Gamma_{N}$ with the expectations operator suppressed and the disturbances $u_{N}$ replaced by (consistent) estimates $\tilde{u}_{N}$.

GM estimates of the parameters $\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}$ and $\sigma_{1}^{2}$ are then obtained as the solution to

$$
\begin{equation*}
\underset{\rho_{1}, \rho_{2}, \ldots, \rho_{R}, \sigma_{N}^{2}, \sigma_{1}^{2}}{\arg \min }\left[\left(g_{N}-G_{N} \alpha\right)^{\prime} \Xi_{N}^{-1}\left(g_{N}-G_{N} \alpha\right)\right]=\left[\vartheta_{N}(\theta)^{\prime} \tilde{\Xi}_{N}^{-1}\left(\vartheta_{N}(\theta)\right],\right. \tag{14}
\end{equation*}
$$

i.e., the parameter estimates can be obtained from a (weighted) nonlinear least squares regression of $g_{N}$ on the columns of $G_{N} ; \vartheta_{N}(\theta)$ can then be viewed as a vector of regression residuals. The optimal choice of the weighting matrix $\Xi_{N}^{-1}$ will be discussed below.

In the following, we define the elements of $\gamma_{N}$ and $\Gamma_{N}$ grouped by the corresponding moment conditions. For this, let us use the following notation:

$$
\begin{align*}
& \bar{u}_{r, N}=\left(I_{T} \otimes W_{r, N}\right) u_{N}, m=1, \ldots, R, \text { and }  \tag{15a}\\
& \overline{\bar{u}}_{r m, N}=\left(I_{T} \otimes W_{r, N}\right)\left(I_{T} \otimes W_{m, N}\right) u_{N}=\left(I_{T} \otimes W_{r, N} W_{m, N}\right) u_{N}, r=1, \ldots, R, m=1, \ldots R . \tag{15b}
\end{align*}
$$

Moreover, running index $l=1, \ldots, R$ has to be introduced for a proper definition of the elements of $\Gamma_{N}$ and $\gamma_{N}$.

Moment condition $\mathrm{M}_{1, r}$ delivers $r=1, \ldots, R$ lines of equation system (12), appearing in rows $4(r-1)+1$ with the following elements of $\gamma_{N}$ and $\Gamma_{N}$ :

$$
\begin{aligned}
& \gamma_{4(r-1)+1}=\frac{1}{N(T-1)} E\left[\bar{u}_{r, N}^{\prime} Q_{0, N} \bar{u}_{r, N}\right], \\
& \gamma_{4(r-1)+1, m}=\frac{2}{N(T-1)} E\left[\bar{u}_{r, N}^{\prime} Q_{0, N} \overline{\bar{u}}_{r m, N}\right], m=1, \ldots, R, \\
& \gamma_{4(r-1)+1, R+m}=-\frac{1}{N(T-1)} E\left[\overline{\bar{u}}_{r m, N}^{\prime} Q_{0, N} \overline{\bar{u}}_{r m, N}\right], m=1, \ldots, R, \\
& \gamma_{4(r-1)+1, R(m+1)-m(m-1) / 2+l-m}=-\frac{2}{N(T-1)} E\left[\overline{\bar{u}}_{r m, N}^{\prime} Q_{0, N} \overline{\bar{u}}_{r l, N}\right], m=1, \ldots, R-1, l=m+1, \ldots, R, \\
& \gamma_{4(r-1)+1,2 R+R(R-1) / 2+1}=\frac{1}{N} t r\left(W_{r, N}^{\prime} W_{r, N}\right), \\
& \gamma_{4(r-1)+1,2 R+R(R-1) / 2+2}=0 .
\end{aligned}
$$

Moment condition $\mathrm{M}_{2, r}$ consists of $r=1, \ldots R$ lines of equation system (12), appearing in rows $4(r-1)+2$ with the following elements of $\gamma_{N}$ and $\Gamma_{N}$ :

$$
\begin{aligned}
& \gamma_{4(r-1)+2}=\frac{1}{N(T-1)} E\left[\bar{u}_{r, N}^{\prime} Q_{0, N} u_{N}\right], \\
& \gamma_{4(r-1)+2, m}=\frac{1}{N(T-1)} E\left[\overline{\bar{u}}_{r m, N}^{\prime} Q_{0, N} u_{N}+\bar{u}_{r, N}^{\prime} Q_{0, N} \bar{u}_{m, N}\right], m=1, \ldots, R, \\
& \gamma_{4(r-1)+2, R+m}=-\frac{1}{N(T-1)} E\left[\overline{\bar{u}}_{r m, N}^{\prime} Q_{0, N} \bar{u}_{m, N}\right], m=1, \ldots, R, \\
& \gamma_{4(r-1)+2, R(m+1)-m(m-1) / 2+l-m}=-\frac{1}{N(T-1)} E\left[\overline{\bar{u}}_{r l, N}^{\prime} Q_{0, N} \bar{u}_{m, N}+\overline{\bar{u}}_{r m, N}^{\prime} Q_{0, N} \bar{u}_{l, N}\right], m=1, \ldots, R-1, \\
& l=m+1, \ldots, R, \\
& \gamma_{4(r-1)+2,2 R+R(R-1) / 2+1}=0, \\
& \gamma_{4(r-1)+2,2 R+R(R-1) / 2+2}=0 .
\end{aligned}
$$

Moment condition $\mathrm{M}_{3, r}$ corresponds to $r=1, \ldots R$ lines of equation system (12), appearing in rows $4(r-1)+3$ with the following elements of $\gamma_{N}$ and $\Gamma_{N}$ :

$$
\begin{aligned}
& \gamma_{4(r-1)+3}=\frac{1}{N} E\left[\bar{u}_{r, N}^{\prime} Q_{1, N} \bar{u}_{r, N}\right], \\
& \gamma_{4(r-1)+3, m}=\frac{2}{N} E\left[\bar{u}_{r, N}^{\prime} Q_{1, N} \overline{\bar{u}}_{r m, N}\right], m=1, \ldots, R, \\
& \gamma_{4(r-1)+3, R+m}=-\frac{1}{N} E\left[\overline{\bar{u}}_{r m, N}^{\prime} Q_{1, N} \overline{\bar{u}}_{r m, N}\right], m=1, \ldots, R, \\
& \gamma_{4(r-1)+3, R(m+1)-m(m-1) / 2+l-m}=-\frac{2}{N} E\left[\overline{\bar{u}}_{r m, N}^{\prime} Q_{1, N} \overline{\bar{u}}_{r l, N}\right], m=1, \ldots, R-1, l=m+1, \ldots, R, \\
& \gamma_{4(r-1)+3,2 R+R(R-1) / 2+1}=0, \\
& \gamma_{4(r-1)+3,2 R+R(R-1) / 2+2}=\frac{1}{N} t r\left(W_{r, N}^{\prime} W_{r, N}\right) .
\end{aligned}
$$

Moment condition $\mathrm{M}_{4, r}$ represents $r=1, \ldots R$ lines of equation system (12) appearing in rows rows $4(r-1)+4$ with the following elements of $\gamma_{N}$ and $\Gamma_{N}$ :

$$
\begin{aligned}
& \gamma_{4(r-1)+4}=\frac{1}{N} E\left[\bar{u}_{r, N}^{\prime} Q_{1, N} u_{N}\right], \\
& \gamma_{4(r-1)+4, m}=\frac{1}{N} E\left[\overline{\bar{u}}_{r m, N}^{\prime} Q_{1, N} u_{N}+\bar{u}_{r, N}^{\prime} Q_{1, N} \bar{u}_{m, N}\right], m=1, \ldots, R, \\
& \gamma_{4(r-1)+4, R+m}=-\frac{1}{N} E\left[\overline{\bar{u}}_{r m, N}^{\prime} Q_{1, N} \bar{u}_{m, N}\right], m=1, \ldots, R, \\
& \gamma_{4(r-1)+4, R(m+1)-m(m-1) / 2+l-m}=-\frac{1}{N} E\left[\overline{\bar{u}}_{r l, N}^{\prime} Q_{1, N} \bar{u}_{m, N}+\overline{\bar{u}}_{r m, N}^{\prime} Q_{1, N} \bar{u}_{l, N}\right], m=1, \ldots, R-1, \\
& l=m+1, \ldots, R, \\
& \gamma_{4(r-1)+4,2 R+R(R-1) / 2+1}=0, \\
& \gamma_{4(r-1)+4,2 R+R(R-1) / 2+2}=0 .
\end{aligned}
$$

Moment condition $\mathrm{M}_{\mathrm{a}}$ reflects 1 line of equation system (12) appearing in row ( $4 R+1$ ) with the following elements of $\gamma_{N}$ and $\Gamma_{N}$ :

$$
\begin{aligned}
& \gamma_{4 R+1}=\frac{1}{N(T-1)} E\left[u_{N}^{\prime} Q_{0, N} u_{N}\right], \\
& \gamma_{4 R+1, m}=\frac{2}{N(T-1)} E\left[\bar{u}_{m, N}^{\prime} Q_{0, N} u_{N}\right], m=1, \ldots, R, \\
& \gamma_{4 R+1, R+m}=-\frac{1}{N(T-1)} E\left[\bar{u}_{m, N}^{\prime} Q_{0, N} \bar{u}_{m, N}\right], m=1, \ldots, R, \\
& \gamma_{4 R+1, R(m+1)-m(m-1) / 2+l-m}=-\frac{2}{N(T-1)} E\left[\bar{u}_{m, N}^{\prime} Q_{0, N} \bar{u}_{l, N}\right], m=1, \ldots, R-1, l=m+1, \ldots, R,
\end{aligned}
$$

$$
\begin{aligned}
& \gamma_{4 R+1,2 R+R(R-1) / 2+1}=1 \\
& \gamma_{4 R+1,2 R+R(R-1) / 2+2}=0
\end{aligned}
$$

Moment condition $\mathrm{M}_{\mathrm{b}}$ is associated with 1 line of equation system (12) appearing in row (4R $+2)$ with the following elements of $\gamma_{N}$ and $\Gamma_{N}$ :

$$
\begin{aligned}
& \gamma_{4 R+2}=\frac{1}{N} E\left[u_{N}^{\prime} Q_{1, N} u_{N}\right], \\
& \gamma_{4 R+2, m}=\frac{2}{N} E\left[\bar{u}_{m, N}^{\prime} Q_{1, N} u_{N}\right], m=1, \ldots, R, \\
& \gamma_{4 R+2, R+m}=-\frac{1}{N} E\left[\bar{u}_{m, N}^{\prime} Q_{1, N} \bar{u}_{m, N}\right], m=1, \ldots, R, \\
& \gamma_{4 R+2, R(m+1)-m(m-1) / 2+l-m}=-\frac{2}{N} E\left[\bar{u}_{m, N}^{\prime} Q_{1, N} \bar{u}_{l, N}\right], m=1, \ldots, R-1, l=m+1, \ldots, R, \\
& \gamma_{4 R+2,2 R+R(R-1) / 2+1}=0 \\
& \gamma_{4 R+2,2 R+R(R-1) / 2+2}=1
\end{aligned}
$$

This completes the specification of the elements of the matrices $\gamma_{N}$ and $\Gamma_{N}$. The similarity of the structure between the expressions resulting from the moment conditions $\mathrm{M}_{a}, \mathrm{M}_{1, r}$ and $\mathrm{M}_{2, r}$ on the one hand and $\mathrm{M}_{\mathrm{b}}, \mathrm{M}_{3, r}, \mathrm{M}_{4, r}$ on the other hand is apparent: they differ only by the normalization factor and the matrix of the quadratic forms ( $Q_{0, N}$ or $Q_{1, N}$ ). Moreover, note that the rows in (12) resulting from $\mathrm{M}_{\mathrm{a}}, \mathrm{M}_{1, r}$ and $\mathrm{M}_{2, r}(r=1, \ldots R)$ do not depend on $\sigma_{1}^{2}$ whereas the rows resulting from $\mathrm{M}_{\mathrm{b}}, \mathrm{M}_{3, r}$, and $\mathrm{M}_{4, r}(r=1, \ldots R)$ do not depend on $\sigma_{v}^{2}$. This fact will be used to define an initial GM estimator, which is based on a subset of moment conditions $\left(\mathrm{M}_{\mathrm{a}}\right.$, $\mathrm{M}_{1, r}$ and $\mathrm{M}_{2, r}$ ) only, in order to obtain an estimate of the matrix $\Xi_{N}$.

For future reference, we define the $(2 R+1) \times 1$ vector $\gamma_{N}^{0}$ as the sub-vector containing rows $r$ and $(r+1), \quad r=1, \ldots, R$ and row $(2 R+1)$ of $\gamma_{N}$ (corresponding to $\mathrm{M}_{1, r}, \mathrm{M}_{2, r}$ and $\mathrm{M}_{\mathrm{a}}$ ). Moreover, we define the $(2 R+1) \times[2 R+R(R-1) / 2+1]$ matrix $\Gamma_{N}^{0}$ as the sub-matrix containing rows $r$ and $(r+1), r=1, \ldots, R$, and row $(2 R+1)$ of $\Gamma_{N}$ (corresponding to $\mathrm{M}_{1, r}$, $\mathrm{M}_{2, r}$ and $\mathrm{M}_{\mathrm{a}}$ ), with the last column of $\Gamma_{N}$ (associated with $\sigma_{1}^{2}$ ) deleted.

Similarly, we define the $(2 R+1) \times 1$ vector $\gamma_{N}^{1}$ as the sub-vector containing rows $2 r,(2 r+1)$, $r=1, \ldots, R$, and row $(2 R+2)$ of $\gamma_{N}$ (corresponding to $\mathrm{M}_{3, r}, \mathrm{M}_{4, r}$ and $\mathrm{M}_{\mathrm{b}}$ ). Finally, we define the $(2 R+1) \times[2 R+R(R-1) / 2+1]$ matrix $\Gamma_{N}^{1}$ as the sub-matrix containing rows $2 r,(2 r+1)$, $r=1, \ldots, R$, and $(2 R+2)$ and $\Gamma_{N}$ (corresponding to $\mathrm{M}_{3, r}, \mathrm{M}_{4, r}$ and $\mathrm{M}_{\mathrm{b}}$ ), with the second last column of $\Gamma_{N}$ (associated with $\sigma_{v}^{2}$ ) deleted.

## 2. GM estimators of an $\boldsymbol{R}$-th order spatial regressive process

### 2.1. Additional assumptions

Before defining the GM estimators, we make three additional assumptions.

## Assumption 3.

The elements of $X_{N}$ are bounded uniformly in absolute value by $k_{x}<\infty$. Furthermore, for $i=0$, 1 , the matrices

$$
\begin{equation*}
M_{i}^{x x}=\lim _{N \rightarrow \infty} \frac{1}{N T} X_{N}^{*}(\rho)^{\prime} Q_{i, N} X_{N}^{*}(\rho), \tag{16a}
\end{equation*}
$$

with $X_{N}^{*}(\rho)=\left[I_{T} \otimes\left(I_{N}-\sum_{r=1}^{R} \rho_{r} W_{r, N}\right)\right] X_{N}$, and the matrices

$$
\begin{equation*}
\lim _{N \rightarrow \infty} \frac{1}{N T} X_{N}^{\prime} X_{N}, \lim _{N \rightarrow \infty} \frac{1}{N T} X_{N}^{*}(\rho)^{\prime} X_{N}^{*}(\rho), \lim _{N \rightarrow \infty} \frac{1}{N T} X_{N}^{*}(\rho)^{\prime} \Omega_{\varepsilon, N}^{-1} X_{N}^{*}(\rho) \tag{16b}
\end{equation*}
$$

are finite and non-singular.
Assumption 3, which is identical to that in the first order case considered by KKP, is typical in large sample analyses. It is required, since the asymptotic properties of OLS and feasible generalized least-squares estimates (GLS; FGLS for feasible GLS) of $\beta$ in (2a) involve limits of the expressions above.

## Assumption 4.

The row and column sums of $W_{r, N}, r=1, \ldots, R$, and $P_{N}(\rho)=\left(I_{N}-\sum_{r=1}^{R} \rho_{r} W_{r, N}\right)^{-1}$ are bounded uniformly in absolute value by $k_{W}<\infty$ and $k_{P}<\infty$, respectively, where $k_{P}$ may depend on $\rho=$ $\left(\rho_{1}, \ldots, \rho_{R}\right)$. We take $k_{W}$ as largest of the bounds of the weights matrices, i.e., $k_{W}=\max \left(k_{W, 1}, \ldots, k_{W, R}\right) .^{3}$ As KKP (pp. 106) point out, assumption (4) restricts the extent of neighborliness of the cross-sectional units on the one hand, and the degree of cross-sectional correlation between the model disturbances on the other hand. Such restrictions on the degree of permissible correlations are standard in virtually all large sample theory.

## Assumption 5

The smallest eigenvalues of $\left(\Gamma_{N}^{0}\right)^{\prime}\left(\Gamma_{N}^{0}\right)$ and $\left(\Gamma_{N}^{1}\right)^{\prime}\left(\Gamma_{N}^{1}\right)$ are bounded away from zero, i.e., $\lambda_{\min }\left[\left(\Gamma_{N}^{i}\right)^{\prime}\left(\Gamma_{N}^{i}\right)\right] \geq \lambda_{*}>0$ for $i=1$, 2, where $\lambda_{*}$ may depend on $\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}$, and $\sigma_{1}^{2}$. Assumption 5 ensures identifiable uniqueness of the parameters $\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}$, and $\sigma_{1}^{2}$. We show in the Appendix that Assumption 5 also implies that the smallest eigenvalue of $\left(\Gamma_{N}\right)^{\prime}\left(\Gamma_{N}\right)$ is bounded away from zero.

[^2]We may now define three different GM estimators for the case of an $R$-th order spatial regressive process (see KKP for analogous conditions under $\operatorname{SAR}(1)$ estimation).

### 2.2. Initial GM estimation

The initial GM estimator is based on a subset of moment conditions ( $M_{\mathrm{a}}, M_{1, \mathrm{r}}$ and $M_{2, \mathrm{r}}$ ) and thus on the matrices $\Gamma_{N}^{0}$ and $\gamma_{N}^{0}$ only. ${ }^{4}$ Define $\theta^{0}$ as the corresponding parameter vector that excludes $\sigma_{1}^{2}$, i.e. $\theta^{0}=\left(\rho, \sigma_{v}^{2}\right)=\left(\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}\right)$ and accordingly

$$
\alpha^{0}=\left(\rho_{1}, \ldots, \rho_{R}, \rho_{1}^{2}, \ldots, \rho_{R}^{2}, \rho_{1} \rho_{2}, \ldots, \rho_{1} \rho_{R}, \ldots, \rho_{R-1} \rho_{R}, \sigma_{v}^{2}\right)^{\prime} .
$$

The initial GM estimator is then obtained as the solution to

$$
\begin{equation*}
\left(\tilde{\rho}_{1, N}, \ldots, \tilde{\rho}_{R, N}, \tilde{\sigma}_{v, N}^{2}\right)=\arg \min \left\{\vartheta_{N}^{0}\left(\underline{\theta}^{0}\right)^{\prime} \vartheta_{N}^{0}\left(\underline{\theta}^{0}\right), \underline{\rho} \in[-a, a], \underline{\sigma}_{v}^{2} \in[0, b]\right\} \tag{17a}
\end{equation*}
$$

where $a \geq 1, b \geq b_{v}$ and $\vartheta_{N}^{0}\left(\underline{\theta}^{0}\right)=\vartheta_{N}^{0}\left(\underline{\rho}, \underline{\sigma}_{v}^{2}\right)=\left(g_{N}^{0}-G_{N}^{0} \underline{\alpha}^{0}\right)$.

Using these initial estimates of $\rho_{1}, \ldots, \rho_{R}$ and $\sigma_{v}^{2}, \sigma_{1}^{2}$ can be estimated from moment condition $\mathrm{M}_{\mathrm{b}}$ :

$$
\begin{align*}
\tilde{\sigma}_{1, N}^{2}= & \frac{1}{N}\left(\tilde{u}_{N}-\sum_{m=1}^{R} \tilde{\rho}_{m, N} \tilde{\bar{u}}_{m, N}\right)^{\prime} Q_{1, N}\left(\tilde{u}_{N}-\sum_{m=1}^{R} \tilde{\rho}_{m, N} \tilde{\bar{u}}_{m, N}\right)  \tag{17b}\\
=g_{4 R+2} & -g_{4 R+2,1} \tilde{\rho}_{1, N}-\ldots-g_{4 R+2, R} \tilde{\rho}_{R, N}-g_{4 R+2, R+1} \tilde{\rho}_{1, N}^{2} \\
& \quad-g_{4 R+2,2 R} \tilde{\rho}_{R, N}^{2}-g_{4 R+2,2 R+1} \tilde{\rho}_{1, N} \tilde{\rho}_{2, N}-\ldots-g_{4 R+2,2 R+R(R-1) / 2} \tilde{\rho}_{R-1, N} \tilde{\rho}_{R, N} .
\end{align*}
$$

### 2.3. Weighted GM estimation

### 2.3.1 Optimal weighting matrix

It is well known from the literature on generalized method of moments estimation, that it is optimal to use as weights matrix the inverse of the (properly normalized) variance-covariance matrix of the moments, evaluated at true parameter values.

The optimal weights matrix is thus given by $\Xi_{N}^{-1}=\left[N \operatorname{Var}\left(\eta_{N}\right)\right]^{-1}$, where $\eta_{N}$ denotes the $[(4 R+2) \times 1]$ vector of moments, evaluated at the true parameter values. Suppressing the expectations operator (and ignoring the deterministic constants), the elements of $\eta_{N}$ correspond to the expressions on the left hand side in equation (11). Notice, however, that the

[^3]ordering in equation system (12) is different, where the rows corresponding to $\mathrm{M}_{\mathrm{a}}$ and $\mathrm{M}_{\mathrm{b}}$ are placed in the last two rows, such that
\[

\eta_{N}=\left[$$
\begin{array}{c}
\frac{1}{N(T-1)} \bar{\varepsilon}_{r, N}^{\prime} Q_{0, N} \bar{\varepsilon}_{r, N}  \tag{18}\\
\frac{1}{N(T-1)} \bar{\varepsilon}_{r, N}^{\prime} Q_{0, N} \varepsilon_{N} \\
\frac{1}{N} \bar{\varepsilon}_{r, N}^{\prime} Q_{1, N} \bar{\varepsilon}_{r, N} \\
\frac{1}{N} \bar{\varepsilon}_{r, N}^{\prime} Q_{1, N} \varepsilon_{N} \\
\cdot \\
\cdot \\
\frac{1}{N(T-1)} \varepsilon_{N}^{\prime} Q_{0, N} \varepsilon_{N} \\
\frac{1}{N} \varepsilon_{N}^{\prime} Q_{1, N} \varepsilon_{N}
\end{array}
$$\right] .
\]

Notice that the matrix $\Xi_{N}=N \operatorname{Var}\left(\eta_{N}\right)$ is symmetric and of order $(4 R+2) \times(4 R+2)$. Its elements are derived by substituting $\bar{\varepsilon}_{r, N}=\left(I_{T} \otimes W_{r, N}\right) \varepsilon_{N}$ and using the results that $\operatorname{Cov}\left(\varepsilon_{N}^{\prime} A_{N} \varepsilon_{N}, \varepsilon_{N}^{\prime} B_{N} \varepsilon_{N}\right)=2 \operatorname{tr}\left(A_{N} \Omega_{\varepsilon, N} B_{N} \Omega_{\varepsilon, N}\right)$ for two non-negative definite symmetric matrices $A_{N}$ and $B_{N}$ and $\varepsilon_{N} \sim \mathrm{~N}\left(0, \Omega_{\varepsilon, N}\right.$ ), see, e.g., Amemiya (1973, p. 5).

The matrix $\Xi_{N}$ is then given as $\Xi_{N}=\left[\xi_{r, s}\right], r, s=1, \ldots, R+1$, i.e.,

$$
\Xi_{N}=\left[\begin{array}{cccc}
\xi_{1,1} & \cdot & \xi_{1, R} & \xi_{1, R+1}  \tag{19}\\
\cdot & & & \cdot \\
\xi_{R, 1} & & \xi_{R, R} & \xi_{R, R+1} \\
\xi_{R+1,1} & & \xi_{R+1, R} & \xi_{R+1, R+1}
\end{array}\right]
$$

Subscript $N$ has been dropped from the elements $\xi_{r, s}$ here to simplify notation. The matrix $\Xi_{N}$ is made up of three parts.
i) An upper left block of dimension $4 R \times 4 R$, consisting of $R^{2}$ blocks of order $4 \times 4$, which are defined as

[^4]\[

$$
\begin{align*}
& \xi_{r, s}=C \otimes T_{r, s}^{W}, r, s=1, \ldots R, \text { where }  \tag{20}\\
& C=\left[\begin{array}{cc}
\frac{1}{(T-1)} \sigma_{v}^{4} & 0 \\
0 & \sigma_{1}^{4}
\end{array}\right], \text { and }  \tag{21}\\
& T_{r, s}^{W}=\left[\begin{array}{cc}
\frac{2}{N} \operatorname{tr}\left(W_{r, N}^{\prime} W_{r, N} W_{s, N}^{\prime} W_{s, N}\right) & \frac{1}{N} \operatorname{tr}\left(W_{r, N}^{\prime} W_{r, N}\left(W_{s, N}^{\prime}+W_{s, N}\right)\right) \\
\frac{1}{N} \operatorname{tr}\left(W_{s, N}^{\prime} W_{s, N}\left(W_{r, N}^{\prime}+W_{r, N}\right)\right) & \frac{1}{N} \operatorname{tr}\left(W_{r, N} W_{s, N}+W_{r, N}^{\prime} W_{s, N}\right)
\end{array}\right], r, s=1, \ldots, R . \tag{22}
\end{align*}
$$
\]

ii) The last row (and column) block of dimension $2 \times 4 R(4 R \times 2)$, each consisting of $R$ blocks of order $2 \times 4(4 \times 2)$, defined as

$$
\left.\begin{array}{l}
\xi_{R+1, s}=C \otimes t_{R+1, s}^{W}, \text { and } \xi_{s, R+1}=\left(\xi_{R+1, s}\right)^{\prime}, s=1, \ldots, R, \text { with } \\
t_{R+1, s}^{W}=\left[\frac{2}{N} \operatorname{tr}\left(W_{s, N}^{\prime} W_{s, N}\right)\right.  \tag{24}\\
\hline
\end{array}\right], s=1, \ldots, R . ~ \$
$$

iii) The lower right block of order $2 \times 2$, defined as

$$
\begin{equation*}
\xi_{R+1, R+1}=C \otimes t_{R+1, R+1}^{W}=C \otimes 2 . \tag{25}
\end{equation*}
$$

For definiteness, we add that the position of each block $\xi_{r, s}$ is such that its upper left element appears in row $(4 r-3)$ and column ( $4 s-3$ ) of the $(4 R+2) \times(4 R+2)$ matrix $\Xi_{N}$. The position of each block $\xi_{R+1, s}=C \otimes t_{R+1, s}^{W}, s=1, \ldots, R$, is such that its first element appears in row ( $4 R+1$ ) and column ( $4 s-3$ ) of $\Xi_{N}$. Finally, the upper left element of the block ( $\xi_{R+1, R+1}$ ) appears in row $(4 R+1)$ and column $(4 R+1)$ of $\Xi_{N}$.

### 2.3.2 The 'weighted GM estimator'

Using the estimate $\tilde{\Xi}_{N}$, one can proceed with a weighted regression, using all $4 R+2$ moment conditions. The weighted GM estimator is obtained as the solution to

$$
\begin{equation*}
\left(\hat{\rho}_{1, N}, \ldots, \hat{\rho}_{R, N}, \hat{\sigma}_{v, N}^{2}, \hat{\sigma}_{1, N}^{2}\right)=\arg \min \left\{\vartheta_{N}(\underline{\theta})^{\prime} \tilde{\Xi}_{N}^{-1} \vartheta_{N}(\underline{\theta}), \underline{\rho} \in[-a, a], \underline{\sigma}_{v}^{2} \in[0, b], \underline{\sigma}_{1}^{2} \in[0, c]\right\} \tag{26}
\end{equation*}
$$

where $a \geq 1, b \geq b_{v}, c \geq T b_{\mu}+b_{v}$, and $\vartheta_{N}(\underline{\theta})=\vartheta_{N}\left(\underline{\rho}, \underline{\sigma}_{v}^{2}, \underline{\sigma}_{1}^{2}\right)=\left(g_{N}-G_{N} \underline{\alpha}\right)$.

### 2.3.3 The 'partially weighted GM estimator'

KKP suggest using a simplified weighting scheme for computational purposes. This scheme uses the same weight for the first three moment conditions ( $\mathrm{M}_{\mathrm{a}}, \mathrm{M}_{1, r}$ and $\mathrm{M}_{2, r}$ ) and the same weight for the three other moment conditions $\left(\mathrm{M}_{\mathrm{b}}, \mathrm{M}_{3, r}\right.$, and $\mathrm{M}_{4, r}$ ), but the weight used for the
first three moment equations is different from that used for last three moment equations. In case of a higher order process, this simplified weighting matrix $\Xi_{N}^{p}$ is given by

$$
\Xi_{N}^{p}=\left[\begin{array}{ccccc}
C \otimes I_{2} & 0 & . & 0 & 0  \tag{27}\\
0 & C \otimes I_{2} & & & 0 \\
\cdot & & \cdot & & \cdot \\
0 & & & C \otimes I_{2} & 0 \\
0 & 0 & \cdot & \cdot & C \otimes 1
\end{array}\right]
$$

and the partially weighted GM estimator is defined as

$$
\begin{equation*}
\left(\check{\rho}_{1, N}, \ldots, \check{\rho}_{R, N}, \check{\sigma}_{v, N}^{2}, \check{\sigma}_{1, N}^{2}\right)=\arg \min \left\{\vartheta_{N}(\underline{\theta})^{\prime}\left(\widetilde{\Xi}_{N}^{p}\right)^{-1} \vartheta_{N}(\underline{\theta}), \underline{\rho} \in[-a, a], \underline{\sigma}_{v}^{2} \in[0, b], \underline{\sigma}_{1}^{2} \in[0, c]\right\}, \tag{28}
\end{equation*}
$$

where $a \geq 1, b \geq b_{v}, c \geq T b_{\mu}+b_{v}$, and $\vartheta_{N}(\underline{\theta})=\vartheta_{N}\left(\underline{\rho}, \underline{\sigma}_{v}^{2}, \underline{\sigma}_{1}^{2}\right)=\left(g_{N}-G_{N} \underline{\alpha}\right)$.

## 3. Properties of the proposed GM estimators

### 3.1. Large sample results

This section summarizes some important asymptotic properties of the proposed GM estimators. The proofs are relegated to the Appendix.

Theorem 1. Consistency of initial GM estimators
Suppose Assumptions 1-5 hold. Then, if $\widetilde{\beta}_{N}$ is a consistent estimator of $\beta$, the initial GM estimators $\left(\tilde{\rho}_{1}, \ldots, \tilde{\rho}_{R}, \tilde{\sigma}_{v}^{2}, \tilde{\sigma}_{1}^{2}\right)$ defined by (17a) and (17b) are consistent for $\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}, \sigma_{1}^{2}$, i.e., $\left(\tilde{\rho}_{1, N}, \ldots, \tilde{\rho}_{R, N}, \tilde{\sigma}_{v, N}^{2}, \tilde{\sigma}_{1, N}^{2} \xrightarrow{P}\left(\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}, \sigma_{1}^{2}\right)\right.$ as $N \rightarrow \infty$.

## Theorem 2. Consistency of weighted GM estimators

Suppose Assumptions 1-5 hold and that the smallest and largest eigenvalues of the matrices $\Xi_{N}^{-1}$ satisfy $0<\bar{\lambda}_{*} \leq \lambda_{\min }\left(\Xi_{N}^{-1}\right) \leq \lambda_{\max }\left(\Xi_{N}^{-1}\right) \leq \bar{\lambda}_{* *}<\infty$. Suppose furthermore that $\tilde{\beta}_{N}$ and $\widetilde{\Xi}_{N}$ are consistent estimators of $\beta$ and $\Xi_{N}$, respectively. Then the weighted GM estimators ( $\hat{\rho}_{1, N}, \ldots, \hat{\rho}_{R, N}, \hat{\sigma}_{v, N}^{2}, \hat{\sigma}_{1, N}^{2}$ ) defined by (26) are consistent for $\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}, \sigma_{1}^{2}$, i.e., $\left(\hat{\rho}_{1, N}, \ldots, \hat{\rho}_{R, N}, \hat{\sigma}_{v, N}^{2}, \hat{\sigma}_{1, N}^{2}\right) \xrightarrow{P}\left(\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}, \sigma_{1}^{2}\right)$ as $N \rightarrow \infty$.

## Theorem 3. Consistency of partially weighted GM estimators

Suppose Assumptions $1-5$ hold. Suppose furthermore that $\tilde{\beta}$ and $\widetilde{\Xi}_{N}^{p}$ are consistent estimators of $\beta$ and $\Xi_{N}^{p}$, respectively. Then, the partially weighted GM estimators $\left(\stackrel{\vee}{\rho}_{1, N}, \ldots, \stackrel{\vee}{\rho}{ }_{R, N}, \stackrel{\vee}{\sigma} 2{ }_{v, N}, \stackrel{\vee}{\sigma}_{1, N}^{2}\right)$ defined by (28) are consistent for $\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}, \sigma_{1}^{2}$, i.e.,

$$
\left(\check{\rho}_{1, N}, \ldots, \check{\rho}_{R, N}, \stackrel{\rightharpoonup}{\sigma}_{v, N}^{2}, \check{\sigma}_{1, N}^{2}\right) \xrightarrow{P}\left(\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}, \sigma_{1}^{2}\right) \text { as } N \rightarrow \infty .
$$

Since the specification of the main equation is identical to that in KKP, the focus of the present paper is on the spatial regressive error process. But it is readily verified by an inspection of the respective proofs in KKP, ${ }^{6}$ that under the maintained assumptions the following two theorems also hold in case of an $R$-th order spatial regressive error process.

Theorem 4. Consistency of OLS estimator of $\beta$
Suppose Assumptions 1-4 holds. The OLS estimator of $\beta$ based on (2a), which is given by
$\hat{\beta}_{N}^{O L S}=\left(X_{N}^{\prime} X_{N}\right)^{-1} X_{N}^{\prime} y_{N}$, is consistent for $\beta$, i.e., $\hat{\beta}_{N}^{O L S} \xrightarrow{P} \beta$ as $N \rightarrow \infty$.

Theorem 5. Asymptotic distribution of the GLS and FGLS estimators of $\beta$
The true generalized least squares (GLS) estimator of $\beta$ is given by

$$
\begin{equation*}
\hat{\beta}_{N}^{G L S}=\left\{X_{N}^{\prime}\left[\Omega_{u, N}^{-1}\left(\rho, \sigma_{v}^{2}, \sigma_{1}^{2}\right)\right] X_{N}\right\}^{-1} X_{N}^{\prime}\left[\Omega_{u, N}^{-1}\left(\rho, \sigma_{v}^{2}, \sigma_{1}^{2}\right)\right] y_{N} . \tag{29a}
\end{equation*}
$$

Using the expression for $\Omega_{u, N}$ in (9a), this can also be written as

$$
\begin{align*}
& \hat{\beta}_{N}^{G L S}=\left\{X_{N}^{* \prime}(\rho)\left[\Omega_{\varepsilon, N}^{-1}\left(\sigma_{v}^{2}, \sigma_{1}^{2}\right)\right] X_{N}^{*}(\rho)\right\}^{-1} X_{N}^{* \prime}(\rho)\left[\Omega_{\varepsilon, N}^{-1}\left(\sigma_{v}^{2}, \sigma_{1}^{2}\right)\right] y_{N}^{*}(\rho), \text { where }  \tag{29b}\\
& X_{N}^{*}(\rho)=\left[I_{T} \otimes\left(I_{N}-\sum_{m=1}^{R} \rho_{m} W_{m, N}\right)\right] X_{N},  \tag{30a}\\
& y_{N}^{*}(\rho)=\left[I_{T} \otimes\left(I_{N}-\sum_{m=1}^{R} \rho_{m} W_{m, N}\right)\right] y_{N} . \tag{30b}
\end{align*}
$$

The feasible generalized least squares (FGLS) estimator is obtained by replacing the true parameters $\rho, \sigma_{v}^{2}$, and $\sigma_{1}^{2}$ by their respective (initial, weighted, or partially weighted) GM estimates, denoted as $\ddot{\rho}, \ddot{\sigma}_{v}^{2}$, and $\ddot{\sigma}_{1}^{2}$.

Now, suppose that Assumptions 1-4 hold.
(a) Then the GLS estimator is consistent and asymptotically normal, i.e.,

$$
\begin{aligned}
& (N T)^{1 / 2}\left[\hat{\beta}_{N}^{G L S}-\beta\right] \xrightarrow{D} N\{0, \Psi\} \text { as } N \rightarrow \infty, \text { with } \\
& \Psi=\left[\sigma_{v}^{-2} M_{0}^{x x}+\sigma_{1}^{-2} M_{1}^{x x}\right]^{-1} .
\end{aligned}
$$

[^5](b) Let $\ddot{\rho}_{N}, \ddot{\sigma}_{v, N}^{2}$, and $\ddot{\sigma}_{1, N}^{2}$ be (any) consistent estimates of $\rho, \sigma_{v}^{2}$, and $\sigma_{1}^{2}$. Then,
$$
(N T)^{1 / 2}\left[\hat{\beta}_{N}^{G L S}-\hat{\beta}_{N}^{F G L S}\right] \xrightarrow{P} 0 \text { as } N \rightarrow \infty .
$$

This means that GLS and FGLS are asymptotically equivalent and, hence, $\hat{\beta}_{N}^{\text {FGLS }}$ is also consistent and asymptotically normal.
(c) Furthermore,

$$
\begin{aligned}
& \dot{\Psi}_{N}-\Psi \xrightarrow{P} 0 \text { as } N \rightarrow \infty \text {, where } \\
& \dot{\Psi}_{N}=\left\{\frac{1}{N T} X_{N}^{* \prime}\left(\dot{\rho}_{N}\right)\left[\Omega_{\varepsilon, N}^{-1}\left(\dot{\sigma}_{v, N}^{2}, \dot{\sigma}_{1, N}^{2}\right)\right] X_{N}^{*}\left(\dot{\rho}_{N}\right)\right\}^{-1} .
\end{aligned}
$$

This suggests that small sample inference can be based on the approximation $\hat{\beta}_{N}^{F G L S} \dot{\sim} N\left(\beta,(N T)^{-1} \dot{\Psi}_{N}\right)$.

While we demonstrate in the Appendix that the initial, weighted, and partially weighted GM estimators of $\rho, \sigma_{v}^{2}$, and $\sigma_{1}^{2}$ defined in (17), (26), and (28) are consistent, we are also interested in their small sample performance. We thus proceed with a Monte Carlo study.

## IV. Monte Carlo analysis

In this section, we consider a Monte Carlos experiment for the case of a third-order spatial regressive process, i.e.,

$$
\begin{equation*}
u_{N}=\sum_{m=1}^{3} \rho_{m}\left(I_{T} \otimes W_{m}\right) u_{N}+\varepsilon_{N} . \tag{31}
\end{equation*}
$$

In all our Monte Carlo experiments, the time dimension is $T=5$. Concerning the cross-section dimension, we consider three sample sizes: $N=100, N=250$, and $N=500$. For our basic setup of the weights matrix, we follow Kelejian and Prucha (1999) and use a binary "up to 9 ahead and up to 9 behind" contiguity specification. This means that the elements of the timeinvariant, raw weights matrix $W^{0}$ are defined such that the $i$-th cross-section element is related to the 9 elements after it and the 9 elements before it.

The raw $N \times N$ matrix $\mathrm{W}^{0}$ is then split up into three $N \times N$ matrices $W_{1}^{0}, W_{2}^{0}$, and $W_{3}^{0}$, where $W_{1}^{0}+W_{2}^{0}+W_{3}^{0}=W^{0}$. The matrices $W_{1}^{0}, W_{2}^{0}$, and $W_{3}^{0}$ are specified such that they contain the elements of $\mathrm{W}^{0}$ for different bands of neighbours and zeros else: $W_{1}^{0}$ corresponds
to an "up to 3 ahead and up to 3 behind" specification, $W_{2}^{0}$ corresponds to a " 4 to 6 ahead and 4 to 6 behind" specification, and $W_{3}^{0}$ corresponds to a " 7 to 9 ahead and 7 to 9 behind" specification. The final weights matrices $W_{1}, W_{2}$, and $W_{3}$ are obtained by separately rownormalizing $W_{1}^{0}, W_{2}^{0}$, and $W_{3}^{0}$, that is by dividing their elements $w_{1, i j}^{0}, w_{2, i j}^{0}$, and $w_{2, i j}^{0}$ through the corresponding row sums $d_{1, i}, d_{2, i}$, and $d_{3, i}$, respectively.

With three row-normalized matrices $W_{1}, W_{2}$, and $W_{3}$, the parameter space for $\rho_{1}, \rho_{2}$, and $\rho_{3}$ must satisfy $0 \leq\left|\rho_{1}\right|+\left|\rho_{2}\right|+\left|\rho_{3}\right|<1$ for ( $I-\rho_{1} W_{1}-\rho_{2} W_{2}-\rho_{3} W_{3}$ ) to be invertible. We consider 10 parameter constellations, assuming that the parameter values ( $\rho_{1}, \rho_{2}, \rho_{3}$ ) are non-increasing in the order of neighbourhood, i.e., we always have $\rho_{1} \geq \rho_{2} \geq \rho_{3}$.

Table 1. Parameter constellations in the Monte Carlo experiments

| Parameter <br> constellation | $\rho_{1}$ | $\rho_{2}$ | $\rho_{3}$ |
| :---: | :---: | :---: | :---: |
| $(1)$ | 0.4 | 0.4 | 0 |
| $(2)$ | 0.4 | 0.2 | 0.2 |
| $(3)$ | 0.4 | 0.2 | 0.1 |
| $(4)$ | 0.4 | 0.2 | 0 |
| $(5)$ | 0.4 | 0 | 0 |
| $(6)$ | 0.2 | 0.2 | 0.2 |
| $(7)$ | 0.2 | 0.1 | 0 |
| $(8)$ | 0.2 | 0.2 | 0 |
| $(9)$ | 0.2 | 0 | 0 |
| $(10)$ | 0 | 0 | 0 |

Regarding the properties of the error process $\varepsilon_{N}$, we assume that $\sigma_{\mu}^{2}=\sigma_{v}^{2}=1$, i.e., the error components $\mu_{N}$ and $v_{i t}$ are drawn from a standard normal distribution. For each Monte Carlo experiment we consider 2000 draws. To ensure comparability, the same draw of $\mu$ and $v$ is used for each of the 10 combinations of $\rho_{1}, \rho_{2}$, and $\rho_{3}$. Tables 2-4 show the results for the three sample sizes.

The tables are organized as follows. Each column shows the results for one parameter constellation, corresponding to the true parameters values given in the rows $\rho_{1}, \rho_{2}, \rho_{3}, \sigma_{\mu}^{2}$ and $\sigma_{v}^{2}$. Below each parameter, the bias and root mean squared error are listed for each of the three estimators, i.e., the initial GM estimator ( $G^{\text {in }}$ ), the weighted $G M$ estimator $\left(G^{w}\right)$, and the partially weighted GM estimator $\left(\mathrm{GM}^{\mathrm{in}}\right)$.

The results suggest that the proposed GM estimator performs reasonably well, even in small samples. As can be seen from Table 2, which is based on a sample size of 100 observations, the bias over all parameter constellations is fairly small for all three estimators.

Table 2. Monte Carlo Results, $N=100, T=5,2000$ draws

| Parameter Constellation ${ }^{1)}$ |  |  | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | average ${ }^{2)}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\rho_{1}$ |  | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 | 0.2 | 0.2 | 0.2 | 0.2 | 0 | 0.2800 |
| Bias | $\mathrm{GM}^{\text {in }}$ | -0.0049 | -0.0058 | -0.0067 | -0.0073 | -0.0093 | -0.0076 | -0.0097 | -0.0088 | -0.0105 | -0.0115 | 0.0082 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.0126 | 0.0142 | 0.0137 | 0.0133 | 0.0140 | 0.0145 | 0.0144 | 0.0141 | 0.0147 | 0.0147 | 0.0140 |
|  | $\mathrm{GM}^{\text {pw }}$ | 0.0022 | 0.0026 | 0.0026 | 0.0025 | 0.0029 | 0.0024 | 0.0027 | 0.0025 | 0.0028 | 0.0026 | 0.0026 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.0652 | 0.0703 | 0.0691 | 0.0681 | 0.0722 | 0.0789 | 0.0806 | 0.0781 | 0.0834 | 0.0936 | 0.0759 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.0592 | 0.0649 | 0.0634 | 0.0621 | 0.0659 | 0.0722 | 0.0732 | 0.0709 | 0.0758 | 0.0847 | 0.0692 |
|  | $\mathrm{GM}^{\text {pw }}$ | 0.0586 | 0.0632 | 0.0618 | 0.0607 | 0.0641 | 0.0705 | 0.0716 | 0.0694 | 0.0740 | 0.0832 | 0.0677 |
| $\rho_{2}$ |  | 0.4 | 0.2 | 0.2 | 0.2 | 0 | 0.2 | 0.1 | 0.2 | 0 | 0 | 0.1500 |
| Bias | $\mathrm{GM}^{\text {in }}$ | -0.0006 | 0.0005 | -0.0003 | -0.0010 | -0.0012 | -0.0015 | -0.0028 | -0.0028 | -0.0028 | -0.0043 | 0.0018 |
|  | $\mathrm{GM}^{\text {w }}$ | -0.0064 | -0.0078 | -0.0071 | -0.0063 | -0.0058 | -0.0066 | -0.0052 | -0.0052 | -0.0051 | -0.0046 | 0.0060 |
|  | $\mathrm{GM}^{\text {pw }}$ | -0.0009 | -0.0014 | -0.0013 | -0.0012 | -0.0014 | -0.0009 | -0.0008 | -0.0006 | -0.0010 | -0.0007 | 0.0010 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.0814 | 0.0795 | 0.0805 | 0.0813 | 0.0795 | 0.0824 | 0.0850 | 0.0848 | 0.0847 | 0.0904 | 0.0829 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.0754 | 0.0745 | 0.0754 | 0.0761 | 0.0745 | 0.0768 | 0.0795 | 0.0791 | 0.0792 | 0.0846 | 0.0775 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0749 | 0.0732 | 0.0742 | 0.0749 | 0.0733 | 0.0759 | 0.0783 | 0.0781 | 0.0780 | 0.0833 | 0.0764 |
| $\rho_{3}$ |  | 0 | 0.2 | 0.1 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0 | 0.0500 |
| Bias | $\mathrm{GM}^{\text {in }}$ | -0.0012 | -0.0023 | -0.0023 | -0.0023 | -0.0033 | -0.0031 | -0.0033 | -0.0028 | -0.0038 | -0.0043 | 0.0029 |
|  | $\mathrm{GM}^{\text {w }}$ | -0.0051 | -0.0051 | -0.0038 | -0.0026 | -0.0006 | -0.0041 | -0.0017 | -0.0026 | -0.0010 | -0.0013 | 0.0028 |
|  | $\mathrm{GM}^{\text {pw }}$ | -0.0005 | -0.0003 | 0.0000 | 0.0001 | 0.0005 | 0.0001 | 0.0002 | 0.0001 | 0.0003 | 0.0002 | 0.0002 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.0702 | 0.0713 | 0.0725 | 0.0730 | 0.0741 | 0.0792 | 0.0809 | 0.0804 | 0.0812 | 0.0875 | 0.0770 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.0649 | 0.0658 | 0.0671 | 0.0675 | 0.0684 | 0.0735 | 0.0751 | 0.0747 | 0.0753 | 0.0813 | 0.0714 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0646 | 0.0657 | 0.0669 | 0.0675 | 0.0686 | 0.0731 | 0.0750 | 0.0745 | 0.0753 | 0.0813 | 0.0713 |
| $\sigma_{v}^{2}$ |  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.0000 |
| Bias |  | -0.0110 | -0.0111 | -0.0114 | -0.0116 | -0.0120 | -0.0121 | -0.0128 | -0.0126 | -0.0129 | -0.0137 | 0.0121 |
|  | $\mathrm{GM}^{\text {w }}$ | -0.0122 | -0.0120 | -0.0123 | -0.0124 | -0.0126 | -0.0111 | -0.0114 | -0.0113 | -0.0115 | -0.0105 | 0.0117 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | -0.0121 | -0.0121 | -0.0122 | -0.0122 | -0.0123 | -0.0120 | -0.0120 | -0.0120 | -0.0120 | -0.0118 | 0.0121 |
| RMSE |  | 0.0723 | 0.0719 | 0.0717 | 0.0716 | 0.0714 | 0.0713 | 0.0709 | 0.0711 | 0.0709 | 0.0708 | 0.0714 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.0721 | 0.0717 | 0.0716 | 0.0716 | 0.0714 | 0.0710 | 0.0707 | 0.0709 | 0.0707 | 0.0704 | 0.0712 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0721 | 0.0717 | 0.0716 | 0.0715 | 0.0713 | 0.0711 | 0.0707 | 0.0709 | 0.0707 | 0.0705 | 0.0712 |
| $\sigma_{1}^{2}$ |  | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6.0000 |
| Bias |  | 0.0218 | 0.0209 | 0.0179 | 0.0158 | 0.0125 | 0.0135 | 0.0082 | 0.0100 | 0.0068 | 0.0020 | 0.0129 |
|  | $\mathrm{GM}^{\text {w }}$ | -0.0914 | -0.0909 | -0.0923 | -0.0933 | -0.0951 | -0.0847 | -0.0869 | -0.0860 | -0.0879 | -0.0816 | 0.0890 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | -0.0886 | -0.0894 | -0.0894 | -0.0892 | -0.0900 | -0.0878 | -0.0876 | -0.0873 | -0.0880 | -0.0862 | 0.0883 |
| RMSE | GM ${ }^{\text {in }}$ | 0.8745 | 0.8712 | 0.8702 | 0.8700 | 0.8675 | 0.8667 | 0.8647 | 0.8661 | 0.8637 | 0.8616 | 0.8676 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.8641 | 0.8624 | 0.8622 | 0.8621 | 0.8610 | 0.8603 | 0.8588 | 0.8595 | 0.8584 | 0.8571 | 0.8606 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.8649 | 0.8630 | 0.8628 | 0.8627 | 0.8615 | 0.8610 | 0.8596 | 0.8604 | 0.8591 | 0.8580 | 0.8613 |

Note: $\mathrm{GM}^{\text {in }}, \mathrm{GM}^{\mathrm{w}}, \mathrm{GM}^{\mathrm{pw}}$ denote initial, weighted, and partially weighted GM estimator respectively. ${ }^{1 \text { ) }}$ Each column corresponds to one parameter constellation (see Table 1). ${ }^{2)}$ Average of absolute row values.

Table 3. Monte Carlo Results, $N=250, T=5$, 2000 draws

| Parameter Constellation ${ }^{1)}$ |  | ${ }^{1)}$ (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | average ${ }^{2)}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\rho_{1}$ |  | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 | 0.2 | 0.2 | 0.2 | 0.2 | 0 | 0.2800 |
| Bias | $\mathrm{GM}^{\text {in }}$ | -0.0028 | -0.0033 | -0.0037 | -0.0039 | -0.0048 | -0.0043 | -0.0052 | -0.0047 | -0.0055 | -0.0061 | 0.0044 |
|  | GM ${ }^{\text {w }}$ | 0.0060 | 0.0068 | 0.0065 | 0.0063 | 0.0066 | 0.0070 | 0.0069 | 0.0067 | 0.0070 | 0.0070 | 0.0067 |
|  | $\mathrm{GM}^{\text {pw }}$ | 0.0010 | 0.0012 | 0.0012 | 0.0012 | 0.0013 | 0.0010 | 0.0012 | 0.0011 | 0.0012 | 0.0010 | 0.0011 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.0399 | 0.0431 | 0.0419 | 0.0410 | 0.0429 | 0.0478 | 0.0480 | 0.0467 | 0.0495 | 0.0555 | 0.0456 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.0359 | 0.0394 | 0.0383 | 0.0373 | 0.0392 | 0.0435 | 0.0436 | 0.0424 | 0.0450 | 0.0502 | 0.0415 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0361 | 0.0390 | 0.0378 | 0.0369 | 0.0386 | 0.0431 | 0.0432 | 0.0420 | 0.0445 | 0.0499 | 0.0411 |
| $\rho_{2}$ |  | 0.4 | 0.2 | 0.2 | 0.2 | 0 | 0.2 | 0.1 | 0.2 | 0 | 0 | 0.1500 |
| Bias | $\mathrm{GM}^{\text {in }}$ | -0.0005 | 0.0000 | -0.0004 | -0.0007 | -0.0009 | -0.0009 | -0.0015 | -0.0015 | -0.0016 | -0.0023 | 0.0010 |
|  | $\mathrm{GM}^{\text {w }}$ | -0.0036 | -0.0042 | -0.0039 | -0.0035 | -0.0032 | -0.0037 | -0.0029 | -0.0030 | -0.0029 | -0.0027 | 0.0033 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | -0.0012 | -0.0014 | -0.0014 | -0.0013 | -0.0014 | -0.0012 | -0.0012 | -0.0011 | -0.0013 | -0.0012 | 0.0013 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.0497 | 0.0487 | 0.0493 | 0.0498 | 0.0487 | 0.0506 | 0.0522 | 0.0521 | 0.0520 | 0.0556 | 0.0509 |
|  | $\mathrm{GM}^{\mathrm{w}}$ | 0.0443 | 0.0437 | 0.0442 | 0.0446 | 0.0436 | 0.0453 | 0.0467 | 0.0466 | 0.0465 | 0.0498 | 0.0455 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0444 | 0.0435 | 0.0440 | 0.0444 | 0.0434 | 0.0452 | 0.0466 | 0.0465 | 0.0464 | 0.0497 | 0.0454 |
| $\rho_{3}$ |  | 0 | 0.2 | 0.1 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0 | 0.0500 |
| Bias | $\mathrm{GM}^{\text {in }}$ | -0.0005 | 0.0000 | -0.0004 | -0.0007 | -0.0009 | -0.0009 | -0.0015 | -0.0015 | -0.0016 | -0.0023 | 0.0010 |
|  | $\mathrm{GM}^{\text {w }}$ | -0.0036 | -0.0042 | -0.0039 | -0.0035 | -0.0032 | -0.0037 | -0.0029 | -0.0030 | -0.0029 | -0.0027 | 0.0033 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | -0.0012 | -0.0014 | -0.0014 | -0.0013 | -0.0014 | -0.0012 | -0.0012 | -0.0011 | -0.0013 | -0.0012 | 0.0013 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.0497 | 0.0487 | 0.0493 | 0.0498 | 0.0487 | 0.0506 | 0.0522 | 0.0521 | 0.0520 | 0.0556 | 0.0509 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.0443 | 0.0437 | 0.0442 | 0.0446 | 0.0436 | 0.0453 | 0.0467 | 0.0466 | 0.0465 | 0.0498 | 0.0455 |
|  | $\mathrm{GM}^{\text {pw }}$ | 0.0444 | 0.0435 | 0.0440 | 0.0444 | 0.0434 | 0.0452 | 0.0466 | 0.0465 | 0.0464 | 0.0497 | 0.0454 |
| $\sigma_{v}^{2}$ |  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.0000 |
| Bias | $\mathrm{GM}^{\text {in }}$ | -0.0043 | -0.0045 | -0.0046 | -0.0046 | -0.0048 | -0.0050 | -0.0052 | -0.0051 | -0.0053 | -0.0057 | 0.0049 |
|  | $\mathrm{GM}^{\text {w }}$ | -0.0052 | -0.0052 | -0.0053 | -0.0054 | -0.0055 | -0.0047 | -0.0049 | -0.0048 | -0.0050 | -0.0045 | 0.0051 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | -0.0051 | -0.0052 | -0.0052 | -0.0052 | -0.0053 | -0.0052 | -0.0051 | -0.0051 | -0.0052 | -0.0051 | 0.0052 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.0464 | 0.0462 | 0.0461 | 0.0460 | 0.0460 | 0.0458 | 0.0456 | 0.0457 | 0.0456 | 0.0455 | 0.0459 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.0463 | 0.0461 | 0.0460 | 0.0460 | 0.0459 | 0.0458 | 0.0456 | 0.0456 | 0.0456 | 0.0454 | 0.0458 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0463 | 0.0462 | 0.0460 | 0.0460 | 0.0459 | 0.0459 | 0.0456 | 0.0457 | 0.0456 | 0.0455 | 0.0459 |
| $\sigma_{1}^{2}$ |  | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6.0000 |
| Bias |  | 0.0095 | 0.0088 | 0.0083 | 0.0080 | 0.0072 | 0.0059 | 0.0047 | 0.0052 | 0.0043 | 0.0017 | 0.0064 |
|  | GM ${ }^{\text {w }}$ | -0.0352 | -0.0349 | -0.0354 | -0.0357 | -0.0364 | -0.0318 | -0.0326 | -0.0323 | -0.0330 | -0.0300 | 0.0337 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | -0.0336 | -0.0341 | -0.0340 | -0.0338 | -0.0341 | -0.0334 | -0.0331 | -0.0330 | -0.0332 | -0.0325 | 0.0335 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.5331 | 0.5323 | 0.5320 | 0.5320 | 0.5317 | 0.5308 | 0.5301 | 0.5303 | 0.5300 | 0.5293 | 0.5311 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.5290 | 0.5290 | 0.5287 | 0.5284 | 0.5284 | 0.5280 | 0.5273 | 0.5273 | 0.5274 | 0.5269 | 0.5280 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.5297 | 0.5297 | 0.5293 | 0.5290 | 0.5291 | 0.5288 | 0.5281 | 0.5281 | 0.5282 | 0.5280 | 0.5288 |

Note: $\mathrm{GM}^{\text {in }}, \mathrm{GM}^{\mathrm{w}}, \mathrm{GM}^{\mathrm{pw}}$ denote initial, weighted, and partially weighted GM estimator respectively. ${ }^{1}$ ) Each column corresponds to one parameter constellation (see Table 1). ${ }^{2)}$ Average of absolute row values.

Table 4. Monte Carlo Results, $N=500, T=5$, 2000 draws

| Parameter Constellation ${ }^{1)}$ |  | ${ }^{1)}$ (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | average ${ }^{2)}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\rho_{1}$ |  | 0.4 | 0.4 | 0.4 | 0.4 | 0.4 | 0.2 | 0.2 | 0.2 | 0.2 | 0 | 0.2800 |
| Bias | $\mathrm{GM}^{\text {in }}$ | -0.0025 | -0.0028 | -0.0029 | -0.0028 | -0.0031 | -0.0032 | -0.0033 | -0.0032 | -0.0034 | -0.0037 | 0.0031 |
|  | GM ${ }^{\text {w }}$ | 0.0021 | 0.0023 | 0.0023 | 0.0023 | 0.0025 | 0.0024 | 0.0025 | 0.0024 | 0.0026 | 0.0026 | 0.0024 |
|  | $\mathrm{GM}^{\text {pw }}$ | -0.0006 | -0.0006 | -0.0005 | -0.0004 | -0.0002 | -0.0007 | -0.0003 | -0.0004 | -0.0003 | -0.0004 | 0.0004 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.0281 | 0.0304 | 0.0297 | 0.0291 | 0.0306 | 0.0341 | 0.0344 | 0.0334 | 0.0355 | 0.0399 | 0.0325 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.0241 | 0.0265 | 0.0258 | 0.0252 | 0.0267 | 0.0295 | 0.0298 | 0.0288 | 0.0308 | 0.0344 | 0.0282 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0247 | 0.0268 | 0.0261 | 0.0255 | 0.0269 | 0.0299 | 0.0301 | 0.0293 | 0.0311 | 0.0350 | 0.0285 |
| $\rho_{2}$ |  | 0.4 | 0.2 | 0.2 | 0.2 | 0 | 0.2 | 0.1 | 0.2 | 0 | 0 | 0.1500 |
| Bias | $\mathrm{GM}^{\text {in }}$ | 0.0008 | 0.0011 | 0.0009 | 0.0008 | 0.0007 | 0.0008 | 0.0005 | 0.0005 | 0.0005 | 0.0004 | 0.0007 |
|  | $\mathrm{GM}^{\text {w }}$ | -0.0009 | -0.0011 | -0.0010 | -0.0008 | -0.0006 | -0.0008 | -0.0004 | -0.0004 | -0.0004 | -0.0002 | 0.0007 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0006 | 0.0005 | 0.0005 | 0.0005 | 0.0004 | 0.0007 | 0.0007 | 0.0007 | 0.0006 | 0.0007 | 0.0006 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.0342 | 0.0335 | 0.0339 | 0.0343 | 0.0335 | 0.0347 | 0.0359 | 0.0358 | 0.0358 | 0.0383 | 0.0350 |
|  | $\mathrm{GM}^{\mathrm{w}}$ | 0.0304 | 0.0300 | 0.0304 | 0.0308 | 0.0302 | 0.0310 | 0.0322 | 0.0320 | 0.0322 | 0.0344 | 0.0314 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0308 | 0.0301 | 0.0305 | 0.0308 | 0.0301 | 0.0312 | 0.0322 | 0.0321 | 0.0321 | 0.0343 | 0.0314 |
| $\rho_{3}$ |  | 0 | 0.2 | 0.1 | 0 | 0 | 0.2 | 0 | 0 | 0 | 0 | 0.0500 |
| Bias | $\mathrm{GM}^{\text {in }}$ | 0.0008 | 0.0011 | 0.0009 | 0.0008 | 0.0007 | 0.0008 | 0.0005 | 0.0005 | 0.0005 | 0.0004 | 0.0007 |
|  | $\mathrm{GM}^{\text {w }}$ | -0.0009 | -0.0011 | -0.0010 | -0.0008 | -0.0006 | -0.0008 | -0.0004 | -0.0004 | -0.0004 | -0.0002 | 0.0007 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0006 | 0.0005 | 0.0005 | 0.0005 | 0.0004 | 0.0007 | 0.0007 | 0.0007 | 0.0006 | 0.0007 | 0.0006 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.0342 | 0.0335 | 0.0339 | 0.0343 | 0.0335 | 0.0347 | 0.0359 | 0.0358 | 0.0358 | 0.0383 | 0.0350 |
|  | $\mathrm{GM}^{\mathrm{w}}$ | 0.0304 | 0.0300 | 0.0304 | 0.0308 | 0.0302 | 0.0310 | 0.0322 | 0.0320 | 0.0322 | 0.0344 | 0.0314 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0308 | 0.0301 | 0.0305 | 0.0308 | 0.0301 | 0.0312 | 0.0322 | 0.0321 | 0.0321 | 0.0343 | 0.0314 |
| $\sigma_{v}^{2}$ |  | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1 | 1.0000 |
| Bias | $\mathrm{GM}^{\text {in }}$ | -0.0017 | -0.0017 | -0.0017 | -0.0017 | -0.0017 | -0.0020 | -0.0020 | -0.0020 | -0.0020 | -0.0022 | 0.0019 |
|  | $\mathrm{GM}^{\text {w }}$ | -0.0021 | -0.0020 | -0.0021 | -0.0021 | -0.0021 | -0.0018 | -0.0018 | -0.0018 | -0.0018 | -0.0016 | 0.0019 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | -0.0020 | -0.0021 | -0.0020 | -0.0020 | -0.0019 | -0.0021 | -0.0019 | -0.0020 | -0.0019 | -0.0019 | 0.0020 |
| RMSE |  | 0.0330 | 0.0329 | 0.0329 | 0.0328 | 0.0328 | 0.0327 | 0.0326 | 0.0326 | 0.0326 | 0.0327 | 0.0328 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.0328 | 0.0327 | 0.0327 | 0.0327 | 0.0328 | 0.0326 | 0.0326 | 0.0326 | 0.0326 | 0.0326 | 0.0327 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.0329 | 0.0328 | 0.0327 | 0.0327 | 0.0327 | 0.0326 | 0.0326 | 0.0326 | 0.0326 | 0.0326 | 0.0327 |
| $\sigma_{1}^{2}$ |  | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6 | 6.0000 |
| Bias |  | 0.0015 | 0.0011 | 0.0011 | 0.0011 | 0.0012 | -0.0007 | -0.0006 | -0.0006 | -0.0005 | -0.0020 | 0.0010 |
|  | GM ${ }^{\text {w }}$ | -0.0199 | -0.0198 | -0.0199 | -0.0199 | -0.0199 | -0.0186 | -0.0185 | -0.0186 | -0.0186 | -0.0174 | 0.0191 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | -0.0202 | -0.0203 | -0.0200 | -0.0197 | -0.0194 | -0.0203 | -0.0195 | -0.0197 | -0.0194 | -0.0194 | 0.0198 |
| RMSE | $\mathrm{GM}^{\text {in }}$ | 0.3760 | 0.3760 | 0.3756 | 0.3755 | 0.3756 | 0.3749 | 0.3743 | 0.3743 | 0.3745 | 0.3740 | 0.3751 |
|  | $\mathrm{GM}^{\text {w }}$ | 0.3754 | 0.3755 | 0.3751 | 0.3748 | 0.3747 | 0.3746 | 0.3738 | 0.3739 | 0.3738 | 0.3734 | 0.3745 |
|  | $\mathrm{GM}^{\mathrm{pw}}$ | 0.3761 | 0.3762 | 0.3758 | 0.3755 | 0.3754 | 0.3753 | 0.3744 | 0.3745 | 0.3745 | 0.3740 | 0.3752 |

Note: $\mathrm{GM}^{\mathrm{in}}, \mathrm{GM}^{\mathrm{w}}, \mathrm{GM}^{\mathrm{pw}}$ denote initial, weighted, and partially weighted GM estimator respectively. ${ }^{1)}$ Each column corresponds to one parameter constellation (see Table 1). ${ }^{2)}$ Average of absolute row values.

For the initial GM estimator ( $\mathrm{GM}^{\text {in }}$ ), the average absolute bias amounts to 2.9 percent for $\rho_{1}$, to 0.6 percent for $\rho_{2}$, and to 1.0 percent for $\rho_{3}$. The average absolute bias is slightly larger with the weighted GM estimator ( $\mathrm{GM}^{\mathrm{w}}$ ) with a small number of cross-sectional units, $N$. The bias deteriorates quickly as the number of cross-sectional observations grows larger (compare the results in the last column of Table 2 with those in Tables 3 and 4).

Considering the root mean squared error (RMSE) of $\rho_{1}, \rho_{2}$, and $\rho_{3}$, we observe that the partially weighted GM estimator $\left(\mathrm{GM}^{\mathrm{pw}}\right)$ performs as well as the weighted GM estimator in fairly small samples with $N=100$. The RMSE of $\mathrm{GM}^{\text {in }}$ is relatively larger than that of $\mathrm{GM}^{\mathrm{pw}}$ and $\mathrm{GM}^{\mathrm{w}}$. The RMSE of GM ${ }^{\mathrm{w}}$ tends to decline faster with an increase in $N$ than that of $\mathrm{GM}^{\mathrm{pw}}$ and $\mathrm{GM}^{\mathrm{in}}$. To see this, compare the last column of Table 2 with that of Table 5. Overall, the RMSE is fairly small across all considered GM estimators even with $N=100$.

## V. Conclusions

Research on the analysis of interdependent data by means of spatial econometric methods has been evolving quite dynamically in recent years. One reason for this observation lies in the fact that various lines of economic theory provide a rich source of hypotheses that relate to interdependent units—individuals, firms, industries, jurisdictions, or countries.

One limitation of most concurrent econometric work on that matter is that much is known about processes with just a single channel of interdependence, while extensions to generalize the possible number of types or decay segments for spatial interdependence mechanisms are scarce and only available for cross-sectional data-sets.

We contribute to the literature on spatial econometrics by formulating a GM estimation procedure which allows researchers to estimate panel data error component models for short time periods with an $R$-th order spatially autoregressive process. Such a model is useful, if the decay function of a given weights matrix-say, for bands of neighbours-is of unknown degree of non-linearity or even non-monotonic. Also, the approach is applicable if several channels of cross-sectional interdependence in conceptually different dimensions-such as geographical, cultural, institutional, industry, or political 'space'-generate effects at the same time and one wishes to estimate their relative importance on outcome.

We prove that the proposed GM estimators for the spatial autoregressive parameters and error component variances are consistent. Under standard assumptions, generalized least-squares (GLS) and feasible GLS (FGLS) estimates of the slope parameters in the main equation are then asymptotically normal, and the weighted GLS and FGLS estimators are efficient. A Monte Carlo analysis for a third-order spatial autoregressive model illustrates that the estimator is applicable even with panel data of a small to medium-sized cross-sectional dimension and fixed time.

## References

Amemiya, T. (1971). The estimation of the variances in a variance-component model. International Economic Review, 12, 1-13.

Anselin, L. (1988). Spatial Econometrics: Methods and Models. Dordrecht: Kluwer, 1988.
Anselin, L. (2003). Spatial externalities, spatial multipliers, and spatial econometrics. International Regional Science Review, 26(2), 153-166.
Badinger, H. and Egger, P. (2008). Intra- and inter-industry productivity spillovers in OECD manufacturing: A spatial econometric perspective. CESIfo Working Paper, No. 2181, January 2008.
Baltagi, B.H., Song, S.H., Jung, B.C. and Koh, W. (2007). Testing for serial correlation, spatial autocorrelation and random effects using panel data. Journal of Econometrics, 140, 5-51.
Bell, K.P. and Bockstael, N.E. (2000). Applying the generalized-moments estimation apporoach to spatial problems involving microlevel data. The Review of Economics and Statistics, 82(1), 72-82.
Cohen, J.P. and Morrison Paul, C. (2007). The impacts of transportation infrastructure on property values: A higher order spatial econometrics approach. Journal of Regional Science, 47(3), 457-478.

Horn, R.A. and Johnson, C.R. (1985). Matrix Analysis. Cambridge: Cambridge University Press, 1985.

Kapoor, M., Kelejian, H.H., and Prucha, I.R. (2007). Panel data models with spatially correlated error components. Journal of Econometrics, 140, 97-130.
Kelejian, H. and Robinson, D. (1992). Spatial autocorrelation: A new computationally simple test with an application to per capita county police expenditures. Regional Science and Urban Economics, 22, 317-331.

Kelejian, H.H. and Prucha, I.R. (1999). A generalized moments estimator for the autoregresssive parameter in a spatial model. International Economic Review, 40, 509533.

Kelejian, H.H. and Prucha, I.R. (2008). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. Journal of Econometrics, forthcoming.
Lee, L.F. and Xiaodong, L. (2006). Efficient GMM estimation of high order spatial autoregressive models. Working Paper, Department of Economics, Ohio State University.

Lee, L.F. and Yu, J. (2007). A spatial dynamic panel data model with both time and individual fixed effects. Working Paper, Department of Economics, Ohio State University.
Mittelhammer, R.C. (1996). Mathematical Statistics for Economics and Business. New York: Springer.
Pötscher, B.M. and Prucha, I.R. (1997). Dynamic Nonlinear Econometric Models, Asymptotic Theory. New York: Springer.
Rao, C.R. (1973). Linear Statistical Inference and its Applications, $2^{\text {nd }}$ edition. New York: Wiley.

## Appendix

The proof of consistency of the GM estimators for a higher order spatial regressive process is given in full length to the benefit of the reader. It proceeds closely along that for the firstorder case given in KKP. We require that the assumptions regarding the properties of $W_{N}$ in KKP hold for each of the matrices $W_{r, N}$ (Assumptions 2 and 4). Moreover, since we have a vector of autoregressive parameters $\rho=\left(\rho_{1}, \ldots, \rho_{R}\right)^{\prime}$, the admissible parameter space needs to be defined differently (Assumption 2). Finally, it has to be accounted for the higher dimension of $\Gamma_{N}, \gamma_{N}$, and $\Xi_{N}$, when considering the eigenvalues, e.g., of $\Gamma_{N}^{\prime} \Gamma_{N}$ (Assumptions 3 and 5).

## Remark A1. Row and column sum boundedness

Definition (KKP, p. 99). Let $B_{N}, N \geq 1$, be some sequence of $k N \times k N$ matrices with $k$ some fixed positive integer. We will then say that the row and column sums of the (sequence of) matrices $B_{N}$ are bounded uniformly in absolute value, if there exists a constant $c<\infty$, which does not depend on $N$, such that

$$
\begin{equation*}
\max _{1 \leq i \leq k N} \sum_{j=1}^{k N}\left|b_{i j, N}\right| \leq c \text { and } \max _{1 \leq j \leq k N} \sum_{i=1}^{k N}\left|b_{i j, N}\right| \leq c \text { for all } N \geq 1 . \tag{A.1}
\end{equation*}
$$

The following results are repeatedly used in the consistency proof (see KKP, pp. 118).
(i) Let $R_{N}$ be a (sequence of) $N \times N$ matrices whose row and column sums are bounded uniformly in absolute value, and let $S$ be some $k \times k$ matrix (with $k \geq 1$ fixed). Then the row and column sums of $S \otimes R_{N}$ are bounded uniformly in absolute value.
(ii) If $A_{N}$ and $B_{N}$ are (sequences of) $k N \times k N$ matrices (with $k \geq 1$ fixed), whose row and column sums are bounded uniformly in absolute value, then so are the row and column sums of $A_{N} B_{N}$ and $A_{N}+B_{N}$. If $Z_{N}$ is a (sequence of) $k N \times p$ matrices whose elements are uniformly bounded in absolute value, then so are the elements of $A_{N} Z_{N}$ and $(k N)^{-1} Z_{N}^{\prime} A_{N} Z_{N}$.

In the following, we give three Lemmata which will be useful for the consistency proof. ${ }^{7}$

## Lemma A1.

Let $S_{T}$ be some $T \times T$ matrix (with $T$ fixed), and let $R_{N}$ be some $N \times N$ matrix whose row and column sums are bounded in uniformly in absolute value. Let $\varepsilon_{N}=\left(e_{T} \otimes I_{N}\right) \mu_{N}+v_{N}$, where $\mu_{N}$ and $v_{N}$ satisfy Assumption 1. Consider the quadratic form

$$
\begin{equation*}
\varphi_{N}=N^{-1} \varepsilon_{N}^{\prime}\left(S_{T} \otimes R_{N}\right) \varepsilon_{N} \tag{A.2}
\end{equation*}
$$

[^6]Then $E\left(\varphi_{N}\right)=\mathrm{O}(1)$ and $\operatorname{Var}\left(\varphi_{N}\right)=\mathrm{o}(1)$, and as a consequence

$$
\varphi_{N}-E\left(\varphi_{N}\right) \xrightarrow{P} 0 \text { as } N \rightarrow \infty .
$$

See KKP (p. 120) for the proof, which relies on the properties of products and sums of row and column sum bounded matrices summarized in Remark A1.

## Lemma A2.

Let $G_{N}^{*}$ and $g_{N}^{*}$ be identical to $\Gamma_{N}$ and $\gamma_{N}$ in (12) except that the expectations operator is dropped. Suppose Assumptions 1, 2, and 4 hold. Then $\Gamma_{N}=\mathrm{O}(1), \gamma_{N}=\mathrm{O}(1)$, and $G_{N}^{*}-\Gamma_{N} \xrightarrow{P} 0$ and $g_{N}^{*}-\gamma_{N} \xrightarrow{P} 0$ as $N \rightarrow \infty$.

## Proof.

Note from (4a) as well as (15a) and (15b) that

$$
\begin{aligned}
& u_{N}=\left[I_{T} \otimes\left(I_{N}-\sum_{m=1}^{R} \rho_{m} W_{m, N}\right)^{-1}\right] \varepsilon_{N}=\left(I_{T} \otimes P_{N}\right) \varepsilon_{N}, \\
& \bar{u}_{m, N}=\left(I_{T} \otimes W_{m, N}\right) u_{N}=\left(I_{T} \otimes W_{m, N} P_{N}\right) \varepsilon_{N}, m=1, \ldots, R, \text { and } \\
& \overline{\bar{u}}_{m l, N}=\left(I_{T} \otimes W_{m, N} W_{l, N}\right) u_{N}=\left(I_{T} \otimes W_{m, N} W_{l, N} P_{N}\right) \varepsilon_{N}, m, l=1, \ldots, R .
\end{aligned}
$$

Define

$$
\begin{align*}
& S_{0, T}=\frac{1}{T-1}\left(I_{T}-\frac{J_{T}}{T}\right), S_{1, T}=\frac{J_{T}}{T}, \text { such that }  \tag{A.3}\\
& S_{0, T} \otimes I_{N}=\frac{1}{(T-1)} Q_{0, N} \text { and } S_{1, T} \otimes I_{N}=Q_{1, N} \tag{A.4}
\end{align*}
$$

Using these definitions, the elements of $g_{N}^{*}=\left[g_{i, N}^{*}\right], i=1, \ldots,(4 R+2)$, and $G_{N}^{*}=\left[g_{i j, N}^{*}\right], i=$ $1, \ldots,(4 R+2)$ and $j=1, \ldots,[2 R+R(R-1) / 2+2]$ are, apart from a constant, expressible as quadratic forms similar to (A.2), i.e.,

$$
\begin{equation*}
\varphi_{i j, N}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{i j, N}\right) \varepsilon_{N}, \tag{A.5}
\end{equation*}
$$

where $i$ and $j$ refer to the row (and column) of the respective element of $g_{N}^{*}$ and $G_{N}^{*} .{ }^{8}$

[^7]Associated with moment condition $\mathrm{M}_{1}$, for each $r=1, \ldots R$ we have:

$$
\begin{align*}
& \varphi_{4(r-1)+1}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, N} \otimes R_{4(r-1)+1}\right) \varepsilon_{N}, R_{4(r-1)+1}=P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N} P_{N},  \tag{A.6}\\
& \varphi_{4(r-1)+1, m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4(r-1)+1, m}\right) \varepsilon_{N}, R_{4(r-1)+1}=P_{N}^{\prime} W_{r, N}^{\prime} W_{r, N} W_{m, N} P_{N}, m=1, \ldots, R, \\
& \varphi_{4(r-1)+1, R+m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4(r-1)+1, R+m}\right) \varepsilon_{N}, R_{4(r-1)+1, R+m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N}^{\prime} W_{r, N} W_{m, N} P_{N}, \\
& m=1, \ldots, R \\
& \varphi_{4(r-1)+1, R(m+1)-m(m-1) / 2+l-m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4(r-1)+1, R(m+1)-m(m-1) / 2+l-m}\right) \varepsilon_{N}, \\
& R_{4(r-1)+1, R(m+1)-m(m-1) / 2+l-m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N}^{\prime} W_{r, N} W_{l, N} P_{N}, m=1, \ldots, R-1, l=m+1, \ldots R .
\end{align*}
$$

Associated with moment condition $\mathrm{M}_{2}$, for each $r=1, \ldots R$ we have:

$$
\begin{aligned}
& \varphi_{4(r-1)+2}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4(r-1)+2}\right) \varepsilon_{N}, R_{4(r-1)+2}=P_{N}^{\prime} W_{r, N}^{\prime} P_{N}, \\
& \varphi_{4(r-1)+2, m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4(r-1)+2, m}\right) \varepsilon_{N}, R_{4(r-1)+2, m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N}^{\prime} P_{N}+P_{N}^{\prime} W_{r, N}^{\prime} W_{m, N} P_{N}, \\
& m=1, \ldots, R, \\
& \varphi_{4(r-1)+2, R+m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4(r-1)+2, R+m}\right) \varepsilon_{N}, R_{4(r-1)+2, R+m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N}^{\prime} W_{m, N} P_{N}, \\
& m=1, \ldots, R, \\
& \varphi_{4(r-1)+2, R(m+1)-m(m-1) / 2+l-m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4(r-1)+2, R(m+1)-m(m-1) / 2+l-m}\right) \varepsilon_{N}, \\
& R_{4(r-1)+2, R(m+1)-m(m-1) / 2+l-m}=P_{N}^{\prime} W_{l, N}^{\prime} W_{r, N}^{\prime} W_{m, N} P_{N}+P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N}^{\prime} W_{l, N} P_{N}, m=1, \ldots, R-1, l= \\
& m+1, \ldots, R .
\end{aligned}
$$

Associated with moment condition $\mathrm{M}_{3}$, for each $r=1, \ldots R$ we have:

$$
\begin{aligned}
& \varphi_{4(r-1)+3}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4(r-1)+3, N}\right) \varepsilon_{N}, R_{4(r-1)+3, N}=P_{N}^{\prime} W_{r, N}^{\prime} W_{m, N} P_{N}, \\
& \varphi_{4(r-1)+3, m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4(r-1)+3, m}\right) \varepsilon_{N}, R_{4(r-1)+3, m}=P_{N}^{\prime} W_{r, N}^{\prime} W_{r, N} W_{m, N} P_{N}, m=1, \ldots, R, \\
& \varphi_{4(r-1)+3, R+m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4(r-1)+3, R+m}\right) \varepsilon_{N}, R_{4(r-1)+3, R+m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N}^{\prime} W_{r, N} W_{m, N} P_{N}, \\
& \quad m=1, \ldots, R, \\
& \varphi_{4(r-1)+3, R(m+1)-m(m-1) / 2+l-m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4(r-1)+3, R(m+1)-m(m-1) / 2+l-m}\right) \varepsilon_{N}, \\
& \quad R_{4(r-1)+3, R(m+1)-m(m-1) / 2+l-m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N}^{\prime} W_{r, N} W_{l, N} P_{N}, m=1, \ldots, R-1, l=m+1, \ldots R .
\end{aligned}
$$

Associated with moment condition $\mathrm{M}_{4}$, for each $r=1, \ldots R$ we have:

$$
\gamma_{4(r-1)+4}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4(r-1)+4}\right) \varepsilon_{N}, R_{4(r-1)+4}=P_{N}^{\prime} W_{r, N}^{\prime} P_{N}, m=1, \ldots, R,
$$

$$
\begin{aligned}
& \gamma_{4(r-1)+4, m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4(r-1)+4, m}\right) \varepsilon_{N}, R_{4(r-1)+4, m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N}^{\prime} P_{N}+P_{N}^{\prime} W_{r, N}^{\prime} W_{m, N} P_{N}, \\
& m=1, \ldots, R, \\
& \gamma_{4(r-1)+4, R+m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4(r-1)+4, R+m}\right) \varepsilon_{N}, R_{4(r-1)+4, R+m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N}^{\prime} W_{m, N} P_{N}, \\
& m=1, \ldots, R, \\
& \gamma_{4(r-1)+4, R(m+1)-m(m-1) / 2+l-m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4(r-1)+4, R(m+1)-m(m-1) / 2+l-m}\right) \varepsilon_{N}, \\
& R_{4(r-1)+4, R(m+1)-m(m-1) / 2+l-m}=P_{N}^{\prime} W_{l, N}^{\prime} W_{r, N}^{\prime} W_{m, N} P_{N}+P_{N}^{\prime} W_{m, N}^{\prime} W_{r, N}^{\prime} W_{l, N} P_{N}, m=1, \ldots, R-1, \\
& l=m+1, \ldots, R .
\end{aligned}
$$

Associated with moment condition $\mathrm{M}_{\mathrm{a}}$ :

$$
\begin{aligned}
& \varphi_{4 R+1}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4 R+1}\right) \varepsilon_{N}, R_{4 R+1}=P_{N}^{\prime} P_{N}, \\
& \varphi_{4 R+1, m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4 R+1, m}\right) \varepsilon_{N}, R_{4 R+1, m}=P_{N}^{\prime} W_{m, N}^{\prime} P_{N}, m=1, \ldots, R, \\
& \varphi_{4 R+1, R+m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4 R+1, R+m}\right) \varepsilon_{N}, R_{4 R+1, R+m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{m, N} P_{N}, m=1, \ldots, R, \\
& \varphi_{4 R+1, R(m+1)-m(m-1) / 2+l-m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{0, T} \otimes R_{4 R+1, R(m+1)-m(m-1) / 2+l-m}\right) \varepsilon_{N}, \\
& R_{4 R+1, R(m+1)-m(m-1) / 2+l-m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{l, N} P_{N}, m=1, \ldots, R-1, \text { and } l=m+1, \ldots R .
\end{aligned}
$$

Associated with moment condition $\mathrm{M}_{\mathrm{b}}$ we have:

$$
\begin{aligned}
& \varphi_{4 R+2}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4 R+2}\right) \varepsilon_{N}, R_{4 R+2}=P_{N}^{\prime} P_{N}, \\
& \varphi_{4 R+2, m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4 R+2, m}\right) \varepsilon_{N}, R_{4 R+2, m}=P_{N}^{\prime} W_{m, N}^{\prime} P_{N}, m=1, \ldots, R, \\
& \varphi_{4 R+2, R+m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4 R+2, R+m}\right) \varepsilon_{N}, R_{4 R+2, R+m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{m, N} P_{N}, m=1, \ldots, R, \\
& \varphi_{4 R+2, R(m+1)-m(m-1) / 2+l-m}=\frac{1}{N} \varepsilon_{N}^{\prime}\left(S_{1, T} \otimes R_{4 R+2, R(m+1)-m(m-1) / 2+l-m}\right) \varepsilon_{N}, \\
& R_{4 R+2, R(m+1)-m(m-1) / 2+l-m}=P_{N}^{\prime} W_{m, N}^{\prime} W_{l, N} P_{N}, m=1, \ldots, R-1, l=m+1, \ldots R .
\end{aligned}
$$

Since the row and column sums of $W_{N}$ and $P_{N}$ are uniformly bounded in absolute value by Assumption 4, so are the the matrices $R_{i j, N}(i=1, \ldots, 6)$ in light of Remark A1. The other elements of $G_{N}^{*}$ and $g_{N}^{*}$ are 0,1 or of the form $\frac{1}{N} \operatorname{tr}\left(W_{r, N}^{\prime} W_{r, N}\right)$ and thus uniformly bounded in absolute value. Lemma A2 now follows by applying Lemma A1 to each of the quadratic forms in (A.6), which compose $G_{N}^{*}$ and $g_{N}^{*}$.

## Lemma $A 3$.

Let $G_{N}^{*}$ and $g_{N}^{*}$ be defined as in Lemma A2. Then, given Assumptions 1 to 4 $G_{N}-G_{N}^{*} \xrightarrow{P} 0$ and $g_{N}-g_{N}^{*} \xrightarrow{P} 0$ as $N \rightarrow \infty$, provided $\widetilde{\beta}_{N} \xrightarrow{P} \beta$ as $N \rightarrow \infty$.

## Proof.

In (A.16), elements of $g_{N}^{*}=\left[g_{i, N}^{*}\right], i=1, \ldots,(4 R+2)$, and $G_{N}^{*}=\left[g_{i, N}^{*}\right], i=1, \ldots,(4 R+2)$ and $j=1, \ldots,[2 R+R(R-1) / 2+2]$ were shown to be of the form

$$
\begin{equation*}
\varphi_{i j, N}=\frac{1}{N} u_{N}^{\prime} C_{i j, N} u_{N}, \tag{A.7}
\end{equation*}
$$

where $C_{i j, N}$ are nonstochastic $N T \times N T$ matrices. Since the row and column sums of the elements of $W_{r, N}$ and $P_{N}$ are uniformly bounded in absolute value by Assumption 4, this is also true for the row and column sums of the matrices $C_{i j, N}$ in light of Remark A1. The elements of $G_{N}$ and $g_{N}$ defined in (13) are - again apart from a constant - given by

$$
\begin{equation*}
\tilde{\varphi}_{i j, N}=\frac{1}{N} \tilde{u}_{N}^{\prime} C_{i j, N} \tilde{u}_{N} . \tag{A.8}
\end{equation*}
$$

To proof Lemma A3 we have to show that $\tilde{\varphi}_{i j, N}-\varphi_{i j, N} \xrightarrow{P} 0$ as $N \rightarrow \infty$. Note that

$$
\begin{equation*}
\tilde{u}_{N}=y_{N}-X_{N} \tilde{\beta}_{N}=u_{N}-X_{N}\left(\dot{\beta}_{N}-\beta\right) . \tag{A.9}
\end{equation*}
$$

Let $\dot{\beta}_{N}$ be any consistent estimator of $\beta$; in that case $\left(\dot{\beta}_{N}-\beta\right) \xrightarrow{P} 0$ as $N \rightarrow \infty$. Substituting (A.9) into (A.8) yields

$$
\begin{equation*}
\tilde{\varphi}_{i j, N}-\varphi_{i j, N}=\left(\dot{\beta}_{N}-\beta\right)^{\prime}\left(N^{-1} X_{N}^{\prime} C_{i j, N} X_{N}\right)\left(\dot{\beta}_{N}-\beta\right)-2\left(\dot{\beta}_{N}-\beta\right)^{\prime}\left(N^{-1} X_{N}^{\prime} C_{i j, N} u_{N}\right) . \tag{A.10}
\end{equation*}
$$

Regarding the first term on the right hand side of (A.10), the row and column sums of $C_{i j, N}$ are bounded uniformly in absolute value as are the elements of $X_{N}$. Utilizing the results in Remark A1, it follows that all $K^{2}$ elements of $N^{-1} X_{N}^{\prime} C_{i j, N} X_{N}$ are $\mathrm{O}(1)$. Thus, the first term on the right hand side converges in probability to zero since $\left(\dot{\beta}_{N}-\beta\right) \xrightarrow{P} 0$ as $N \rightarrow \infty$.

Regarding the second term on the right hand side of (A.10), consider the vector $\zeta_{N}=N^{-1} X_{N}^{\prime} C_{i j, N} u_{N}$. The mean of $\zeta_{N}$ is zero and its variance covariance matrix is

$$
\begin{equation*}
N^{-1}\left(N^{-1} X_{N}^{\prime} C_{i j, N} \Omega_{u, N} C_{i j, N}^{\prime} X_{N}\right), \tag{A.11}
\end{equation*}
$$

where $\Omega_{u, N}$ is given by (9a) and (9b). Given the maintained assumptions, the row and column sums of $\Omega_{u, N}$ are uniformly bounded in absolute value, and therefore so are those of $C_{i j, N} \Omega_{u, N} C_{i j, N}^{\prime}$. Since the elements of $X_{N}$ are uniformly bounded in absolute value by Assumption A3, it follows that all $K^{2}$ elements of $N^{-1} X_{N}^{\prime} C_{i j, N} \Omega_{u, N} C_{i j, N}^{\prime} X_{N}$ are $\mathrm{O}(1)$ in light of remark A1. As a consequence, the variance covariance matrix of $\zeta_{N}$ converges to zero and hence $\zeta_{N}$ converges to zero in probability. This establishes that also the second term on the right hand side of (A.10) converges to zero in probability.

## Theorem A1.

Combining Lemmata A2 and A3 we have

$$
\begin{equation*}
G_{N}-\Gamma_{N} \xrightarrow{P} 0 \text { and } g_{N}-\gamma_{N} \xrightarrow{P} 0 \text { as } N \rightarrow \infty . \tag{A.12}
\end{equation*}
$$

With these initial results at hand we can now demonstrate the consistency of the GM estimators defined in section III. We first prove consistency of the initial GM estimator (Theorem 1) and then turn to the weighted and partially weighted GM estimators (Theorems 2 and 3). In all three cases the proof proceeds in two steps (based on the assumption that the estimators $\tilde{\theta}_{N}$ exist and are measurable). ${ }^{9}$ We first show that the true parameter vector $\theta$ is identifiable unique using Lemma 4.1 in Pötscher and Prucha (1997). Then we proof consistency by checking the criterion given in Lemma 3.1 in Pötscher and Prucha (1997).

## Proof of Theorem 1. Consistency of initial GM estimator

The objective function of the nonlinear least squares estimator in (17a) and its nonstochastic counterpart are given by

$$
\begin{align*}
& R_{N}^{0}\left(\underline{\theta}^{0}\right)=\left(g_{N}^{0}-G_{N}^{0} \underline{\alpha}^{0}\right]^{\prime}\left[g_{N}^{0}-G_{N}^{0} \underline{\alpha}^{0}\right] \text { and }  \tag{A.13a}\\
& \bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)=\left[\gamma_{N}^{0}-\Gamma_{N}^{0} \underline{\alpha}^{0}\right]^{\prime}\left[\gamma_{N}^{0}-\Gamma_{N}^{0} \underline{\alpha}^{0}\right] . \tag{A.13b}
\end{align*}
$$

Since $\gamma_{N}^{0}-\Gamma_{N}^{0} \alpha^{0}=0$, we have $\bar{R}_{N}^{0}\left(\theta^{0}\right)=0$, i.e., $\bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)=0$ at the true parameter vector $\theta^{0}=\left(\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}\right)$.

Then,

[^8]\[

$$
\begin{equation*}
\bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)-\bar{R}_{N}^{0}\left(\theta^{0}\right)=\left[\underline{\alpha}^{0}-\alpha^{0}\right]^{\prime} \Gamma_{N}^{0^{\prime}} \Gamma_{N}^{0}\left[\underline{\alpha}^{0}-\alpha^{0}\right] . \tag{A.14}
\end{equation*}
$$

\]

In light of Rao (1973, p. 62), it follows that:

$$
\begin{aligned}
& \bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)-\bar{R}_{N}^{0}\left(\theta^{0}\right) \geq \lambda_{\min }\left(\Gamma_{N}^{0^{\prime}} \Gamma_{N}^{0}\right)\left[\underline{\alpha}^{0}-\alpha^{0}\right]^{\prime}\left[\underline{\alpha}^{0}-\alpha^{0}\right] \text { and } \\
& \bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)-\bar{R}_{N}^{0}\left(\theta^{0}\right) \geq \lambda_{*}\left[\underline{\alpha}^{0}-\alpha^{0}\right]^{\prime}\left[\underline{\alpha}^{0}-\alpha^{0}\right] \text { by Assumption } 5 .
\end{aligned}
$$

Using the norm $\|A\|=[\operatorname{tr}(A A)]^{1 / 2}$, we have $\left\|\underline{\theta}^{0}-\theta^{0}\right\|^{2} \leq\left[\underline{\alpha}^{0}-\alpha^{0}\right]^{\prime}\left[\underline{\alpha}^{0}-\alpha^{0}\right]$. It follows that $\bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)-\bar{R}_{N}^{0}\left(\theta^{0}\right) \geq \lambda_{*}\left\|\underline{\theta}^{0}-\theta^{0}\right\|^{2}$. Hence, for every $\varepsilon>0$

$$
\begin{equation*}
\lim _{N \rightarrow \infty} \inf _{\left\{\theta^{0}:\left|\underline{\theta}^{0}-\theta^{0}\right| \geq \varepsilon\right\}}\left[\bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)-\bar{R}_{N}^{0}\left(\theta^{0}\right)\right] \geq \inf _{\left\{\theta^{0}:\left|\underline{\theta}^{0}-\theta^{0}\right| \geq \varepsilon\right\}} \lambda_{*}\left\|\underline{\theta}^{0}-\theta^{0}\right\|^{2}=\lambda_{*} \varepsilon^{2}>0 \tag{A.15}
\end{equation*}
$$

which proves that the true parameter $\theta^{0}$ is identifiable unique.

Next, let $F_{N}^{0}=\left[g_{N}^{0},-G_{N}^{0}\right]$ and $\Phi_{N}^{0}=\left[\gamma_{N}^{0},-\Gamma_{N}^{0}\right]$, then the objective function and its nonstochastic counterpart can be written as

$$
\begin{aligned}
& R_{N}^{0}\left(\underline{\theta}^{0}\right)=\left(1, \underline{\alpha}^{0^{\prime}}\right) F_{N}^{0^{\prime}} F_{N}^{0}\left(1, \underline{\alpha}^{0^{\prime}}\right)^{\prime} \text { and } \\
& \bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)=\left(1, \underline{\alpha}^{0^{\prime}}\right) \Phi_{N}^{0^{\prime}} \Phi_{N}^{0}\left(1, \underline{\alpha}^{0^{\prime}}\right)^{\prime} .
\end{aligned}
$$

Hence for $\rho \in[-a, a]$ and $\sigma_{v}^{2} \in[0, b]$ it holds that

$$
\left\|R_{N}^{0}\left(\underline{\theta}^{0}\right)-\bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)\right\|=\left\|\left(1, \underline{,}^{0^{\prime}}\right)\left[F_{N}^{0^{\prime}} F_{N}^{0}-\Phi_{N}^{0^{\prime}} \underline{\Phi}_{N}^{0}\right]\left(1, \underline{\alpha}^{0^{\prime}}\right)^{\prime}\right\| .
$$

Moreover, since the norm $\|\cdot\|$ is submultiplicative, i.e., $\|A B\| \leq\|A\|\|B\|$, we have

$$
\begin{aligned}
& \left|R_{N}^{0}\left(\underline{\theta}^{0}\right)-\bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)\right| \leq\left\|F_{N}^{o^{\prime}} F_{N}^{0}-\Phi_{N}^{0^{\prime}} \Phi_{N}^{0}\right\|\left\|\left(1, \underline{\underline{\alpha}}^{0^{\prime}}\right)\right\|^{2} \\
& \leq\left\|F_{N}^{0^{\prime}} F_{N}^{0}-\Phi_{N}^{0^{\prime}} \Phi_{N}^{0}\right\|\left[1+R a^{2}+\frac{2 R+R(R-1)}{2} a^{4}+b^{2}\right]
\end{aligned}
$$

In light of Theorem A1, we have $\left\|F_{N}^{0}-\Phi_{N}^{0}\right\| \xrightarrow{P} 0$ as $N \rightarrow \infty$. Observing that, by Lemma A2, the elements of $\Phi_{N}^{0}$ are $\mathrm{O}(1)$ it follows that $\left\|F_{N}^{0^{\prime}} F_{N}^{0}-\Phi_{N}^{0^{\prime}} \Phi_{N}^{0}\right\| \xrightarrow{P} 0$ as $N \rightarrow \infty$. As a consequence, we have (for finite $R$ )

$$
\begin{equation*}
\sup _{\rho \in[-a, a], \sigma_{v}^{2} \in[0, b]}\left|R_{N}^{0}\left(\underline{\theta}^{0}\right)-\bar{R}_{N}^{0}\left(\underline{\theta}^{0}\right)\right| \leq\left\|\left[F_{N}^{0^{\prime}} F_{N}^{0}-\Phi_{N}^{0^{\prime}} \Phi_{N}^{0}\right]\right\|\left[1+R a^{2}+\frac{R(R-1)}{2} a^{4}+b^{2}\right] \xrightarrow{P} 0 \text { as } N \rightarrow \infty \tag{A.16}
\end{equation*}
$$

Together with identifiable uniqueness, the consistency of $\tilde{\theta}_{N}^{0}=\left(\tilde{\rho}_{1, N}, \ldots, \tilde{\rho}_{R, N}, \tilde{\sigma}_{v, N}^{2}\right)$ now follows directly from Lemma 3.1 in Pötscher and Prucha (1997).

Having proved that the estimators $\tilde{\rho}_{1, N} \ldots, \tilde{\rho}_{R, N}, \tilde{\sigma}_{v, N}^{2}$ are consistent for $\rho_{1, \ldots,}, \rho_{R}, \sigma_{v}^{2}$, we now show that $\sigma_{1}^{2}$ can be estimated consistently from the last line $(4 R+2)$ of equation system (12), using

$$
\begin{align*}
\tilde{\sigma}_{1, N}^{2}= & g_{4 R+2}-g_{4 R+2,1} \tilde{\rho}_{1, N}-\ldots-g_{4 R+2, R} \tilde{\rho}_{R, N}-g_{4 R+2, R+1} \tilde{\rho}_{1, N}^{2} \\
& -g_{4 R+2,2 R} \tilde{\rho}_{R, N}^{2}-g_{4 R+2,2 R+1} \tilde{\rho}_{1, N} \tilde{\rho}_{2, N}-\ldots-g_{4 R+2,2 R+R(R-1) / 2} \tilde{\rho}_{R-1, N} \tilde{\rho}_{R, N} . \tag{A.17a}
\end{align*}
$$

Since $\gamma_{N}-\Gamma_{N} \alpha=0$, we have

$$
\begin{align*}
& \tilde{\sigma}_{1}^{2}-\sigma_{1}^{2}=\left(g_{4 R+2}-\gamma_{4 R+2}\right)-\left(g_{4 R+2,1}-\gamma_{4 R+2,1}\right) \tilde{\rho}_{1, N}-\ldots-\left(g_{4 R+2, R}-\gamma_{4 R+2, R}\right) \tilde{\rho}_{R, N} \\
& -\left(g_{4 R+2, R+1}-\gamma_{4 R+2, R+1}\right) \tilde{\rho}_{1, N}^{2}-\ldots-\left(g_{4 R+2,2 R}-\gamma_{4 R+2,2 R}\right) \tilde{\rho}_{R, N}^{2} \\
& -\left(g_{4 R+2,2 R+1}-\gamma_{4 R+2,2 R+1}\right) \tilde{\rho}_{1, N} \tilde{\rho}_{2, N}-\ldots-\left(g_{4 R+2,2 R+R(R-1) / 2}-\gamma_{4 R+2,2 R+R(R-1) / 2}\right) \tilde{\rho}_{R-1, N} \tilde{\rho}_{R, N}  \tag{A.17b}\\
& -\gamma_{4 R+2,1}\left(\tilde{\rho}_{1, N}-\rho_{1}\right)-\ldots-\gamma_{4 R+2, R}\left(\tilde{\rho}_{R, N}-\rho_{R}\right) \\
& -\gamma_{4 R+2, R+1}\left(\tilde{\rho}_{1, N}^{2}-\rho_{1}^{2}\right) \ldots-\gamma_{4 R+2,2 R}\left(\tilde{\rho}_{R, N}^{2}-\rho_{R}^{2}\right) \\
& -\gamma_{4 R+2,2 R+1}\left(\tilde{\rho}_{1, N} \tilde{\rho}_{2, N}^{2}-\rho_{1} \rho_{2}\right) \ldots-\gamma_{4 R+2,2 R+R(R-1) / 2}\left(\tilde{\rho}_{R-1, N} \tilde{\rho}_{R, N}-\rho_{R-1} \rho_{R}\right) .
\end{align*}
$$

Observing by Theorem A1 that $F_{N}-\Phi_{N} \xrightarrow{P} 0$ as $N \rightarrow \infty$ and that the elements of $\Phi_{N}$ are $\mathrm{O}(1)$ it follows from the consistency of $\tilde{\rho}_{1, N}, \ldots, \tilde{\rho}_{R, N}$ that $\tilde{\sigma}_{1, N}^{2}-\sigma_{1}^{2} \xrightarrow{P} 0$ as $N \rightarrow \infty$.

## Proof of Theorem 2. Consistency of the weighted GM estimator

The objective function of the weighted GM estimator and its nonstochastic counterpart are given by

$$
\begin{equation*}
R_{N}(\underline{\theta})=\left(g_{N}-G_{N} \underline{\alpha}\right]^{\prime} \tilde{\Xi}_{N}^{-1}\left[g_{N}-G_{N} \underline{\alpha}\right] \text { and } \tag{A.18a}
\end{equation*}
$$

$$
\begin{equation*}
\bar{R}_{N}(\underline{\theta})=\left[\gamma_{N}-\Gamma_{N} \underline{\alpha}\right]^{\prime} \Xi_{N}^{-1}\left[\gamma_{N}-\Gamma_{N} \underline{\alpha}\right] \tag{A.18b}
\end{equation*}
$$

First, in order to ensure identifiable uniqueness, we show that Assumption 5 also implies that the smallest eigenvalue of $\left(\Gamma_{N}\right)^{\prime} \Xi_{N}^{-1}\left(\Gamma_{N}\right)$ is bounded away from zero, i.e.,

$$
\begin{equation*}
\lambda_{\min }\left(\Gamma_{N}^{\prime} \Xi_{N}^{-1} \Gamma_{N}\right) \geq \lambda_{0} \text { for some } \lambda_{0}>0 . \tag{A.19}
\end{equation*}
$$

Let $A=\left(a_{i j}\right)=\Gamma_{N}^{0^{\prime}} \Gamma_{N}^{0}$ and $B=\left(b_{i j}\right)=\Gamma_{N}^{1^{\prime}} \Gamma_{N}^{1}$. Note that $\Gamma_{N}^{0}$ and $\Gamma_{N}^{1}$ are of dimension $(2 R+1) \times[2 R+R(R-1) / 2+1]$ (i.e., they have half the rows and one column less than than $\Gamma_{N}$ ). $A$ and $B$ are of order $[2 R+R(R-1) / 2+1] \times[2 R+R(R-1) / 2+1]$ (i.e., they have one row and column less than $\Gamma_{N}^{\prime} \Gamma_{N}$ ).

Now define $\breve{\Gamma}_{N}$ as

$$
\breve{\Gamma}_{N}=\left[\begin{array}{c}
\breve{\Gamma}_{N}^{0}  \tag{A.20a}\\
\breve{\Gamma}_{N}^{1}
\end{array}\right],
$$

which differs from $\Gamma_{N}$ only by the ordering of the rows.
$\breve{\Gamma}_{N}^{0}$ corresponds to $\Gamma_{N}^{0}$ with a zero column appended as last column, i.e., $\breve{\Gamma}_{N}^{0}=\left[\Gamma_{N}^{0}, 0\right]$, such that

$$
\breve{\Gamma}_{N}^{0^{\prime}} \breve{\Gamma}_{N}^{0}=\left[\begin{array}{cc}
\breve{\Gamma}_{N}^{0^{\prime}} \breve{\Gamma}_{N}^{0} & 0  \tag{A.21a}\\
0 & 0
\end{array}\right]=\left[\begin{array}{ccc}
a_{1,1} & a_{1,2 R+R(R-1) / 2+1} & 0 \\
\cdot & & 0 \\
a_{2 R+R(R-1) / 2+1,1} & a_{2 R+R(R-1) / 2+1,2 R+R(R-1) / 2+1} & 0 \\
0 & 0 & 0
\end{array}\right] .
$$

( $\breve{\Gamma}_{N}^{0^{\prime}} \breve{\Gamma}_{N}^{0}$ is of the same dimension as $\Gamma_{N}^{\prime} \Gamma_{N}$, i.e., $[2 R+R(R-1) / 2+2] \times[2 R+R(R-1) / 2+2]$.)
$\breve{\Gamma}_{N}^{1}$ is a modified version of $\Gamma_{N}^{1}$, with a zero column included as second last column, such that

$$
\breve{\Gamma}_{N}^{1} \breve{\Gamma}_{N}^{1}=\left[\begin{array}{cccc}
b_{1,1} & & 0 & b_{1,2 R+R(R-1) / 2+1}  \tag{A.21b}\\
\cdot & & 0 & \cdot \\
0 & 0 & 0 & 0 \\
b_{2 R+R(R-1) / 2+1,1} & \cdot & 0 & b_{2 R+R(R-1) / 2+1,2 R+R(R-1) / 2+1}
\end{array}\right] .
$$

$\left(\breve{\Gamma}_{N}^{\prime} \breve{\Gamma}_{N}^{1}\right.$ is of the same dimension as $\Gamma_{N}^{\prime} \Gamma_{N}$, i.e., $\left.[2 R+R(R-1) / 2+2] \times[2 R+R(R-1) / 2+2].\right)$

Since $\breve{\Gamma}_{N}=\left[\begin{array}{c}\breve{\Gamma}_{N}^{0} \\ \breve{\Gamma}_{N}^{1}\end{array}\right]$ differs from $\Gamma_{N}$ only by the ordering of the rows, it follows that

$$
\begin{align*}
& \Gamma_{N}^{\prime} \Gamma_{N}=\breve{\Gamma}_{N}^{\prime} \breve{\Gamma}_{N}=\left[\begin{array}{ll}
\breve{\Gamma}_{N}^{0^{\prime}} & \breve{\Gamma}_{N}^{1}
\end{array}\right]\left[\begin{array}{c}
\breve{\Gamma}_{N}^{0} \\
\breve{\Gamma}_{N}^{1}
\end{array}\right]=\breve{\Gamma}_{N}^{0^{\prime}} \breve{\Gamma}_{N}^{0}+\breve{\Gamma}_{N}^{\prime} \breve{\Gamma}_{N}^{1} \text {, i.e., }  \tag{A.22}\\
& \Gamma_{N}^{\prime} \Gamma_{N}=\left[\begin{array}{cccc}
a_{1,1} & & a_{1,2 R+R(R-1) / 2+1} & 0 \\
\cdot & & 0 \\
a_{2 R+R(R-1) / 2+1,1} & & a_{2 R+R(R-1) / 2+1,2 R+R(R-1) / 2+1} & 0 \\
0 & 0 & 0 & 0
\end{array}\right] \\
& +\left[\begin{array}{cccc}
b_{1,1} & & 0 & b_{1,2 R+R(R-1) / 2+1} \\
\cdot & & 0 & \cdot \\
0 & 0 & 0 & 0 \\
b_{2 R+R(R-1) / 2+1,1} & \cdot & 0 & b_{2 R+R(R-1) / 2+1,2 R+R(R-1) / 2+1}
\end{array}\right] \text {. }
\end{align*}
$$

Utilizing Assumption A5 we have

$$
\begin{equation*}
x^{\prime} \Gamma_{N}^{\prime} \Gamma_{N} x=x^{\prime} \breve{\Gamma}_{N}^{0} \breve{\Gamma}_{N}^{0} x+x^{\prime} \breve{\Gamma}_{N}^{1} \breve{\Gamma}_{N}^{1} x=x_{A}^{\prime} A x_{A}+x_{B}^{\prime} B x_{B} \tag{A.23}
\end{equation*}
$$

The vector $x$ is of dimension [2R $+R(R-1) / 2+2] \times 1$ (corresponding to the number of columns of $\Gamma_{N}$ ), wheras $x_{A}$ and $x_{B}$ are of dimension [ $2 R+R(R-1) / 2+1$ ], i.e. both have one row less: $x_{A}$ excludes the last element of $x$, i.e., $x_{2 R+R(R-1)+2,} x_{B}$ excludes the second-last element of $x$, i.e. $x_{2 R+R(R-1)+1}$.

Again, we invoke Rao (1973, p. 62) for each quadratic form. It follows

$$
\begin{equation*}
x_{A}^{\prime} A x_{A}+x_{B}^{\prime} B x_{B} \geq \lambda_{\text {min }}(A) x_{A}^{\prime} x_{A}+\lambda_{\text {min }}(B) x_{B}^{\prime} x_{B} \geq \lambda^{*}\left(x_{A}^{\prime} x_{A}+x_{B}^{\prime} x_{B}\right) \geq \lambda^{*} x^{\prime} x \tag{A.24}
\end{equation*}
$$

for any $x=\left[x_{1}, x_{2}, \ldots, x_{2 R+2}\right]$.

Hence, we have shown that

$$
x^{\prime} \Gamma_{N}^{\prime} \Gamma_{N} x \geq \lambda^{*} x^{\prime} x
$$

or, equivalently,

$$
\begin{equation*}
\frac{x^{\prime} \Gamma_{N}^{\prime} \Gamma_{N} x}{x^{\prime} x} \geq \lambda^{*} \text { for } x \neq 0 . \tag{A.25}
\end{equation*}
$$

Next, note that in light of Rao (1973, p. 62),

$$
\begin{equation*}
\lambda_{\min }\left(\Gamma_{N}^{\prime} \Gamma_{N}\right)=\inf _{x} \frac{x^{\prime} \Gamma_{N}^{\prime} \Gamma_{N} x}{x^{\prime} x} \geq \lambda_{*}>0 . \tag{A.26}
\end{equation*}
$$

Using Mittelhammer (1996, p. 254) we have

$$
\begin{align*}
& \lambda_{\min }\left(\Gamma_{N}^{\prime} \Xi_{N}^{-1} \Gamma_{N}\right)=\inf _{x} \frac{x^{\prime} \Gamma_{N}^{\prime} \Xi_{N}^{-1} \Gamma_{N} x}{x^{\prime} x} \geq \lambda_{\min }\left(\Xi_{N}^{-1}\right) \inf _{x} \frac{x^{\prime} \Gamma_{N}^{\prime} \Gamma_{N} x}{x^{\prime} x} \\
& \quad=\lambda_{\min }\left(\Xi_{N}^{-1}\right) \lambda_{\min }\left(\Gamma_{N}^{\prime} \Gamma_{N}\right) \geq \lambda_{0}>0, \tag{A.27}
\end{align*}
$$

with $\lambda_{0}=\bar{\lambda}_{*} \lambda_{*}$ since $=\lambda_{\text {min }}\left(\Xi_{N}^{-1}\right) \geq \bar{\lambda}_{*}>0$ by assumption (see Theorem 2).

This ensures that the true parameter vector $\theta=\left(\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}, \sigma_{1}^{2}\right)$ is identifiable unique.

Next note that in light of the assumptions in Theorem $2, \Xi_{N}^{-1}$ is $\mathrm{O}(1)$ by the equivalence of matrix norms.

Analogous to the prove of theorem 1, observe that $\bar{R}_{N}(\theta)=0$, i.e., $\bar{R}_{N}(\underline{\theta})=0$ at the true parameter vector $\theta=\left(\rho_{1}, \ldots, \rho_{R}, \sigma_{v}^{2}, \sigma_{1}^{2}\right)$. It follows that

$$
\begin{equation*}
\bar{R}_{N}(\underline{\theta})-\bar{R}_{N}(\theta)=[\underline{\alpha}-\alpha]^{\prime} \Gamma_{N}^{\prime} \Xi_{N}^{-1} \Gamma_{N}[\underline{\alpha}-\alpha] . \tag{A.30}
\end{equation*}
$$

Moreover, let $F_{N}=\left[g_{N},-G_{N}\right]$ and $\Phi_{N}=\left[\gamma_{N},-\Gamma_{N}\right]$, then,

$$
\begin{align*}
& R_{N}(\underline{\theta})=\left(1, \underline{\alpha}^{\prime}\right) F_{N}^{\prime} \tilde{\Xi}_{N}^{-1} F_{N}\left(1, \underline{\alpha}^{\prime}\right)^{\prime} \text { and }  \tag{A.31a}\\
& \bar{R}_{N}(\underline{\theta})=\left(1, \underline{\alpha}^{\prime}\right) \Phi_{N}^{\prime} \Xi_{N}^{-1} \Phi_{N}\left(1, \underline{\alpha}^{\prime}\right)^{\prime} . \tag{A.31b}
\end{align*}
$$

The remainder of the proof is now analogous to that of Theorem 1.

## Proof of Theorem 3. Consistency of partially weighted GM estimator

Let $\bar{\lambda}_{*}=\min \left[(T-1) \sigma_{v}^{-4}, \sigma_{1}^{-4}\right]$ and $\bar{\lambda}_{* *}=\max \left[(T-1) \sigma_{v}^{-4}, \sigma_{1}^{-4}\right]$.
Then $0<\bar{\lambda}_{*} \leq \lambda_{\text {min }}\left(\Xi_{N}^{p}\right)^{-1} \leq \bar{\lambda}_{* *}<\infty$. The proof of Theorem 3 is now analogous to that of Theorem 2 with $\Xi_{N}$ and $\widetilde{\Xi}_{N}$ replaced by $\Xi_{N}^{p}$ and $\widetilde{\Xi}_{N}^{p}$.

# CESifo Working Paper Series 

for full list see www.cesifo-group.org/wp
(address: Poschingerstr. 5, 81679 Munich, Germany, office@cesifo.de)

2241 Gary E. Bolton and Axel Ockenfels, Does Laboratory Trading Mirror Behavior in Real World Markets? Fair Bargaining and Competitive Bidding on EBay, March 2008

2242 Atsushi Oshima, B. Ravikumar and Raymond Riezman, Entrepreneurship, Organization Capital and the Evolution of the Firm, March 2008

2243 Walter Krämer and Sebastian Schich, Large-Scale Disasters and the Insurance Industry, March 2008

2244 Leif Danziger, Adjustment Costs, Inventories and Output, March 2008
2245 Anne van Aaken, Lars P. Feld and Stefan Voigt, Power over Prosecutors Corrupts Politicians: Cross Country Evidence Using a New Indicator, March 2008

2246 Hans-Christian Heinemeyer, Max-Stephan Schulze and Nikolaus Wolf, Endogenous Borders? The Effects of New Borders on Trade in Central Europe 1885-1933, March 2008

2247 Johannes Becker and Clemens Fuest, Tax Competition - Greenfield Investment versus Mergers and Acquisitions, March 2008

2248 Giorgio Bellettini and Hubert Kempf, Why not in your Backyard? On the Location and Size of a Public Facility, March 2008

2249 Jose Luis Evia, Roberto Laserna and Stergios Skaperdas, Socio-Political Conflict and Economic Performance in Bolivia, March 2008

2250 Bas Jacobs and A. Lans Bovenberg, Optimal Taxation of Human Capital and the Earnings Function, March 2008

2251 Jan-Egbert Sturm and Timo Wollmershäuser, The Stress of Having a Single Monetary Policy in Europe, March 2008

2252 Guido Schwerdt, Labor Turnover before Plant Closure: ‘Leaving the Sinking Ship’ vs. ‘Captain Throwing Ballast Overboard’, March 2008

2253 Keith E. Maskus and Shuichiro Nishioka, Development-Related Biases in Factor Productivities and the HOV Model of Trade, March 2008

2254 Jeremy Edwards and Sheilagh Ogilvie, Contract Enforcement, Institutions and Social Capital: the Maghribi Traders Reappraised, March 2008

2255 Imed Drine and Christophe Rault, Purchasing Power Parity for Developing and Developed Countries. What can we Learn from Non-Stationary Panel Data Models?, March 2008

2256 Scott Alan Carson, Health, Wealth and Inequality: a Contribution to the Debate about the Relationship between Inequality and Health, March 2008

2257 C.A.E. Goodhart, The Regulatory Response to the Financial Crisis, March 2008
2258 Stefan Bauernschuster, Oliver Falck and Stephan Heblich, The Impact of Continuous Training on a Firm’s Innovations, March 2008

2259 Michael Grimm and Stephan Klasen, Geography vs. Institutions at the Village Level, March 2008

2260 Fwu-Ranq Chang, Property Insurance, Portfolio Selection and their Interdependence, March 2008

2261 J. Atsu Amegashie and Marco Runkel, The Paradoxes of Revenge in Conflicts, March 2008

2262 Hans Jarle Kind, Marko Koethenbuerger and Guttorm Schjelderup, Efficiency Enhancing Taxation in Two-sided Markets, March 2008

2263 M. Hashem Pesaran, Til Schuermann and L. Vanessa Smith, Forecasting Economic and Financial Variables with Global VARs, March 2008

2264 Volker Grossmann, Entrepreneurial Innovation and Sustained Long-run Growth without Weak or Strong Scale Effects, March 2008

2265 Robert S. Chirinko and Huntley Schaller, The Irreversibility Premium, March 2008
2266 Andrea Galeotti and José Luis Moraga-González, Platform Intermediation in a Market for Differentiated Products, April 2008

2267 Torben M. Andersen and Michael Svarer, The Role of Workfare in Striking a Balance between Incentives and Insurance in the Labour Market, April 2008

2268 Harald Badinger, Cyclical Fiscal Policy, Output Volatility, and Economic Growth, April 2008

2269 Thomas Aronsson and Erkki Koskela, Outsourcing and Optimal Nonlinear Taxation: A Note, April 2008

2270 Gary E. Bolton, Claudia Loebbecke and Axel Ockenfels, How Social Reputation Networks Interact with Competition in Anonymous Online Trading: An Experimental Study, April 2008

2271 Nikolaus Wolf, Scylla and Charybdis. Explaining Europe’s Exit from Gold, January 1928 - December 1936, April 2008

2272 Michael Funke and Marc Gronwald, The Undisclosed Renminbi Basket: Are the Markets Telling us something about where the Renminbi - US Dollar Exchange Rate is Going?, April 2008

2273 Thor Olav Thoresen and Annette Alstadsæter, Shifts in Organizational Form under a Dual Income Tax System, April 2008

2274 Helge Berger and Volker Nitsch, Too many Cooks? Committees in Monetary Policy, April 2008

2275 Yin-Wong Cheung and Eiji Fujii, Deviations from the Law of One Price in Japan, April 2008

2276 Michael S. Michael, Sajal Lahiri and Panos Hatzipanayotou, Integrated Reforms of Indirect Taxes in the Presence of Pollution, April 2008

2277 Bas Jacobs, Is Prescott Right? Welfare State Policies and the Incentives to Work, Learn and Retire, April 2008

2278 Burkhard Heer and Alfred Maußner, Value Function Iteration as a Solution Method for the Ramsey Model, April 2008

2279 Jarko Fidrmuc and Christa Hainz, Integrating with their Feet: Cross-Border Lending at the German-Austrian Border, April 2008

2280 Kristof Dascher and Alexander Haupt, The Political Economy of Regional Integration Projects at Borders where Rich and Poor Meet: The Role of Cross-Border Shopping and Community Sorting, April 2008

2281 Katrin Assenmacher-Wesche and M. Hashem Pesaran, A VECX* Model of the Swiss Economy, April 2008

2282 Christophe Rault, Robert Sova and Ana Maria Sova, Modeling International Trade Flows between CEEC and OECD Countries, April 2008

2283 Timo Boppart, Josef Falkinger, Volker Grossmann, Ulrich Woitek and Gabriela Wüthrich, Qualifying Religion: The Role of Plural Identities for Educational Production, April 2008

2284 Armin Falk, David Huffman and W. Bentley MacLeod, Institutions and Contract Enforcement, April 2008

2285 Axel Dreher and Stefan Voigt, Does Membership in International Organizations Increase Governments' Credibility? Testing the Effects of Delegating Powers, April 2008

2286 Xavier Freixas and Bruno M. Parigi, Lender of Last Resort and Bank Closure Policy, April 2008

2287 Regina Dionisius, Samuel Muehlemann, Harald Pfeifer, Günter Walden, Felix Wenzelmann and Stefan C. Wolter, Cost and Benefit of Apprenticeship Training - A Comparison of Germany and Switzerland, April 2008

2288 Francesco Daveri and Cecilia Jona-Lasinio, Off-Shoring and Productivity Growth in the Italian Manufacturing Industries, April 2008

2289 Mikael Priks, Do Surveillance Cameras Affect Unruly Behavior? A Close Look at Grandstands, April 2008

2290 Marianna Belloc and Daniela Federici, A Two-Country NATREX Model for the Euro/Dollar, April 2008

2291 Nicolas Treich, The Value of a Statistical Life under Ambiguity Aversion, April 2008
2292 J. Atsu Amegashie, Socially-Tolerable Discrimination, April 2008
2293 M. Hashem Pesaran and Andreas Pick, Forecasting Random Walks Under Drift Instability, April 2008

2294 Steven Brakman, Gus Garita, Harry Garretsen and Charles van Marrewijk, Unlocking the Value of Cross-Border Mergers and Acquisitions, May 2008

2295 Eric O’N. Fisher and Kathryn G. Marshall, The Structure of the American Economy, May 2008

2296 Claudia M. Buch and Martin Schlotter, Regional Origins of Employment Volatility: Evidence from German States, May 2008

2297 Helmuth Cremer, Philippe De Donder, Dario Maldonado and Pierre Pestieau, Taxing Sin Goods and Subsidizing Health Care, May 2008

2298 Reinhilde Veugelers and Frederick van der Ploeg, Reforming European Universities: Scope for an Evidence-Based Process, May 2008

2299 Jon H. Fiva and Lars J. Kirkebøen, Does the Housing Market React to New Information on School Quality?, May 2008

2300 Tina Klautke and Alfons J. Weichenrieder, Interest Income Tax Evasion, the EU Savings Directive, and Capital Market Effects, May 2008

2301 Harald Badinger and Peter Egger, GM Estimation of Higher Order Spatial Autoregressive Processes in Panel Data Error Component Models, May 2008


[^0]:    ${ }^{1}$ Apart from Kapoor, Kelejian, and Prucha (2007), panel data models for SAR processes have been suggested, for instance, by Anselin (1988), Baltagi, Song, Jung, and Koh (2007), and Lee and Yu (2007). While this list is not comprehensive, our approach of allowing for a $\operatorname{SAR}(R)$ process in a panel data error components model is novel, to the best of our knowledge. Previous work on higher order spatial processes, focussing on a cross-sectional model however, includes Lee and Xiadong (2006).

[^1]:    ${ }^{2}$ If the weights matrices $W_{m, N}$ are not row-normalized, assumption (c) would be implied by taking the permissible parameter space to be $\sum_{m=1}^{R}\left|\rho_{m}\right|<\left(\max _{m=1, \ldots, R}\left\|\mid W_{m, N}\right\|\right)^{-1}$, where $\|\|$ is any matrix norm (see Horn and Johnson, 1985, p. 301).

[^2]:    ${ }^{3}$ See Appendix A1 for a definition of row (column) sum boundedness.

[^3]:    ${ }^{4}$ An alternative would be to use an unweighted GM estimator based on all six moment conditions; however, as KKP (2007) show in their Monte Carlos study for a SAR(1) model, this estimator may perform much worse than the initial GM estimator, based on just three of their six moment conditions corresponding to our $\Gamma_{N}^{0}$ and $\gamma_{N}^{0}$. Hence, in the following, the initial GM estimator will be used to obtain an estimate of $\Xi_{N}^{-1}$.

[^4]:    ${ }^{5}$ For quadratic forms in non-symmetric matrices $A_{N}$ (or $B_{N}$ ) we use the fact that $\varepsilon_{N}^{\prime} A_{N} \varepsilon_{N}=\varepsilon_{N}^{\prime} A^{\prime}{ }_{N} \varepsilon_{N}=\varepsilon_{N}^{\prime}\left(A_{N}+A_{N}^{\prime}\right) \varepsilon_{N} / 2$, which is a quadratic form in the symmetric matrix $\left(A_{N}+A_{N}^{\prime}\right) / 2$.

[^5]:    ${ }^{6}$ In particular, the proofs of consistency of $\operatorname{OLS}$ in $\operatorname{KKP}$ (2007, p. 124) and the proof of Theorem 4 in KKP (2007, p. 126).

[^6]:    ${ }^{7}$ The lemmata are derived in KKP (2007) and adapted here to the higher order setting.

[^7]:    ${ }^{8}$ For notational simplicity, we drop the subscript $N$ from $\varphi_{i j, N}$ and $R_{i j, N}$. As in section III, the elements are grouped by moment conditions.

[^8]:    ${ }^{9}$ This is ensured, for example, by Lemma 2 in Jennrich (1969) or Lemma 3.4 in Pötscher and Prucha (1997).

