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

HARALD BADINGER
PETER EGGER

CESIFO WORKING PAPER NO. 2301
CATEGORY 10: EMPIRICAL AND THEORETICAL METHODS
MAY 2008

CESifo *Working Paper No. 2301*

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
Vienna University of Economics and Business Administration
Althanstrasse 39 - 45
1090 Vienna
Austria
harald.badinger@wu-wien.ac.at

Peter Egger
Ifo Institute for Economic Research at the University of Munich
Poschingerstrasse 5
81679 Munich
Germany
egger@ifo.de

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 cross-sectional units.

This paper formulates a GM estimator for the case of a SAR process of order *R*, i.e., 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 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

[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 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).

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 pre-defined 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 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 $i = 1,...,N$ cross-sectional units and $t = 1,...,T$ time periods. For time period *t*, the model reads

$$y_N(t) = X_N(t)\beta + u_N(t), \tag{1a}$$

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 SAR(*R*) process:

$$u_N(t) = \sum_{m=1}^{R} \rho_m W_{m,N} u_N(t) + \varepsilon_N(t), \tag{1b}$$

$$\varepsilon_N(t) = \mu_N + v_N(t), \tag{1c}$$

where ρ_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, time-invariant $N \times N$ matrices $W_{m,N}$, whose elements $w_{ij,N}$ 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 ρ_m, $m = 1,...,R$) 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, μ_N and $v_N(t)$. As indicated by the notation, μ_N is time-invariant while $v_N(t)$ is not. The typical elements of $\varepsilon_N(t)$, μ_N, and $v_N(t)$ are the scalars $\varepsilon_{it,N}$, $\mu_{i,N}$, and $v_{it,N}$, 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

$$y_N = X_N \beta + u_N , \tag{2a}$$

where $y_N = [y'_N(1),..., y'_N(T)]'$ is the $NT \times 1$ vector of observations on the dependent variable. The regressor matrix $X_N = [X'_N(1),..., X'_N(T)]'$ is of dimension $NT \times K$. Generalizing the specification in KKP (p. 100), the $NT \times 1$ vector of error terms $u_N = [u'_N(1),...,u'_N(T)]$ for a spatial regressive process of order R reads

$$u_N = \sum_{m=1}^{R} \rho_m (I_T \otimes W_{m,N}) u_N + \varepsilon_N , \tag{2b}$$

where I_T is an identity matrix of dimension $T \times T$. The $NT \times 1$ vector $\varepsilon_N = [\varepsilon'_N(1),...,\varepsilon'_N(T)]'$ is specified as

$$\varepsilon_N = (e_T \otimes I_N)\mu_N + v_N . \tag{2c}$$

The $N \times 1$ vector of unit specific error components is given by $\mu_N = [\mu_1, \mu_2,..., \mu_N]'$. 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

$$\varepsilon_N = u_N - \sum_{m=1}^{R} \rho_m (I_T \otimes W_{m,N}) u_N . \tag{3}$$

It follows that

$$u_N = [I_T \otimes (I_N - \sum_{m=1}^{R} \rho_m W_{m,N})^{-1}]\varepsilon_N , \tag{4a}$$

and

$$y_N = X_N\beta + [I_T \otimes (I_N - \sum_{m=1}^{R}\rho_m W_{m,N})^{-1}]\varepsilon_N \ . \tag{4b}$$

The following assumptions are maintained throughout the analysis.

Assumption 1.
Let T be a fixed positive integer. (a) For all $1 \le t \le T$ and $1 \le i \le N$, $N \ge 1$, the error components $v_{it,N}$ are identically distributed with zero mean and variance σ_v^2 , $0 < \sigma_v^2 < b_v < \infty$, and finite fourth moments. In addition, for each $N \ge 1$ and $1 \le t \le T$, $1 \le i \le N$ the error components $v_{it,N}$ are independently distributed. (b) For all $1 \le i \le N$, $N \ge 1$ the unit specific error components $\mu_{i,N}$ are identically distributed with zero mean and variance σ_μ^2, $0 < \sigma_\mu^2 < b_\mu < \infty$, and finite fourth moments. Moreover, for each $N \ge 1$ and $1 \le i \le N$ the unit-specific error components $\mu_{i,N}$ are independently distributed. (c) The processes $\{v_{it,N}\}$ and $\{\mu_{i,N}\}$ 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) $\sum_{m=1}^{R}|\rho_m| < 1$.

(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 row-normalized weights matrices typically used in applied work, Assumption (2c) is implied by Assumption (2b).[2]

Assumptions 1 and 2 imply that

$$E(\varepsilon_{it,N}\varepsilon_{js,N}) = \sigma_\mu^2 + \sigma_v^2 \text{ for } i = j \text{ and } t = s, \tag{5a}$$

$$E(\varepsilon_{it,N}\varepsilon_{js,N}) = \sigma_\mu^2 \text{ for } i = j \text{ and } t \neq s, \text{ and} \tag{5b}$$

$$E(\varepsilon_{it,N}\varepsilon_{js,N}) = 0 \text{ otherwise.} \tag{5c}$$

[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 $\sum_{m=1}^{R}|\rho_m| < \left(\max_{m=1,\ldots,R}\|W_{m,N}\|\right)^{-1}$, where $\|\ \|$ is any matrix norm (see Horn and Johnson, 1985, p. 301).

As a consequence, the variance-covariance matrix of the stacked error term ε_N is given by

$$\Omega_{\varepsilon,N} = E[\varepsilon_N \varepsilon_N'] = \sigma_\mu^2 (J_T \otimes I_N) + \sigma_v^2 I_{NT}, \tag{6a}$$

where $J_T = e_T e_T'$ is a $T \times T$ matrix with unit elements and I_{NT} is an identity matrix of dimension $NT \times NT$. Equation (6a) can also be written as

$$\Omega_{\varepsilon,N} = \sigma_v^2 Q_{0,N} + \sigma_1^2 Q_{1,N}, \tag{6b}$$

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

$$Q_{0,N} = (I_T - \frac{J_T}{T}) \otimes I_N \text{, and} \tag{7a}$$

$$Q_{1,N} = \frac{J_T}{T} \otimes I_N . \tag{7b}$$

Notice that $Q_{0,N}$ and $Q_{1,N}$ are both of order $NT \times NT$, and they are symmetric, idempotent, orthogonal to each other, and sum up to I_{NT}. Pre-multiplying an $NT \times 1$ vector, e.g., ε_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 $\varepsilon_{it,N} - 1/T\sum_{t=1}^{T}\varepsilon_{it,N}$ and $1/T\sum_{t=1}^{T}\varepsilon_{it,N}$, 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):

$$tr(Q_{0,N}) = N(T-1),\ tr(Q_{1,N}) = N,\ Q_{0,N}(e_T \otimes I_N) = 0,\ Q_{1,N}(e_T \otimes I_N) = (e_T \otimes I_N), \tag{8}$$
$$(I_T \otimes D_N)Q_{0,N} = Q_{0,N}(I_T \otimes D_N),\ (I_T \otimes D_N)Q_{1,N} = Q_{1,N}(I_T \otimes D_N),$$
$$tr[(I_T \otimes D_N)Q_{0,N}] = (T-1)tr(D_N),\ \text{and}\ tr[(I_T \otimes D_N)Q_{1,N}] = tr(D_N),$$

where D_N is an arbitrary $N \times N$ matrix.

Finally, note that the variance-covariance matrix of u_N is given by

$$\Omega_{u,N} = E[u_N u_N'] = [I_T \otimes (I_N - \sum_{m=1}^{R}\rho_m W_{m,N})^{-1}]\Omega_{\varepsilon,N}[I_T \otimes (I_N - \sum_{m=1}^{R}\rho_m W_{m,N}')^{-1}] \text{, and} \tag{9a}$$

$$E[u_N(t)u_N'(t)] = (\sigma_\mu^2 + \sigma_v^2)(I_N - \sum_{m=1}^{R}\rho_m W_{m,N})^{-1}(I_T \otimes (I_N - \sum_{m=1}^{R}\rho_m W_{m,N}')^{-1} . \tag{9b}$$

III. GM Estimation of a 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 (SAR(R), with R = 1). With an R-th order process (SAR(R), with $R > 1$), the GM estimators of the parameters $\rho_1,...,\rho_R$, σ_v^2, and σ_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 $W_{r,N}$

$$\bar{\varepsilon}_{r,N} = (I_T \otimes W_{r,N})\varepsilon_N = (I_T \otimes W_{r,N})(u_N - \sum_{m=1}^{R} \rho_m (I_T \otimes W_{m,N}) u_N) . \tag{10}$$

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

$$\begin{aligned}
&\mathrm{M}_{a} \quad E[\frac{1}{N(T-1)} \varepsilon_N' Q_{0,N} \varepsilon_N] = \sigma_v^2 , \\
&\mathrm{M}_{1,r} \quad E[\frac{1}{N(T-1)} \bar{\varepsilon}_{r,N}' Q_{0,N} \bar{\varepsilon}_{r,N}] = \sigma_v^2 \frac{1}{N} tr(W_{r,N}' W_{r,N}) , \\
&\mathrm{M}_{2,r} \quad E[\frac{1}{N(T-1)} \bar{\varepsilon}_{r,N}' Q_{0,N} \varepsilon_N] = 0 , \\
&\mathrm{M}_{b} \quad E[\frac{1}{N} \varepsilon_N' Q_{1,N} \varepsilon_N] = \sigma_1^2 \text{ , where } \sigma_1^2 = \sigma_v^2 + T\sigma_\mu^2 , \\
&\mathrm{M}_{3,r} \quad E[\frac{1}{N} \bar{\varepsilon}_{r,N}' Q_{1,N} \bar{\varepsilon}_{r,N}] = \sigma_1^2 \frac{1}{N} tr(W_{r,N}' W_{r,N}) , \\
&\mathrm{M}_{4,r} \quad E[\frac{1}{N} \bar{\varepsilon}_{r,N}' Q_{1,N} \varepsilon_N] = 0 .
\end{aligned} \tag{11}$$

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

Substituting equations (4), (10), and (1c) into the $4R$ +2 moment conditions (11) yields a ($4R$ + 2) equation system in $\rho_1, \ldots, \rho_R$, σ_v^2, and σ_1^2, which can be written as

$$\gamma_N - \Gamma_N \alpha = 0 , \tag{12}$$

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

$$\alpha = (\rho_1,...,\rho_R, \rho_1^2,...,\rho_R^2, \rho_1\rho_2,...,\rho_1\rho_R,...,\rho_{R-1}\rho_R, \sigma_v^2, \sigma_1^2)',$$

i.e., α contains R linear terms ρ_m ($m = 1,...,R$), R quadratic terms ρ_m^2 ($m = 1,...,R$), $R(R-1)/2$ cross products $\rho_m \rho_l$ ($m = 1,...,R-1, l = m+1,...,R$), as well as σ_v^2 and σ_1^2. For later reference, we define the vector of spatial regressive parameters $\rho = (\rho_1,...,\rho_R)'$ and the (row) vector of all parameters as $\theta = (\rho_1,...,\rho_R, \sigma_v^2, \sigma_1^2)$.

γ_N is a $(4R+2)\times 1$ vector with elements $[\gamma_i]$, i = 1, …, $(4R+2)$, and Γ_N is a $(4R+2)\times[2R + R(R-1)/2 + 2]$ matrix with elements $[\gamma_{ij}]$, $i = 1,...,(4R+2)$, $j = 1,...,[2R + R(R-1)/2 + 2]$, whose elements will be defined below. Subscript N is dropped from the elements of γ_N and Γ_N for simplicity of notation here. The row-index of the elements γ_N and Γ_N will be chosen such that the equation system (12) has the following order. The first four rows correspond to the moment restrictions $M_{1,1}$ to $M_{4,1}$ associated with matrix $W_{1,N}$ through (10); row five to eight correspond to $M_{1,2}$ to $M_{4,2}$ associated with matrix $W_{2,N}$, and so fourth; rows $(R-4)$ to $4R$ correspond to the $M_{1,R}$ to $M_{4,R}$ associated with the matrix $W_{R,N}$. Finally, rows $4R$+1 and $4R$+2 correspond to the moment conditions M_a and M_b, respectively, which do not depend on r.

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

$$g_N - G_N\alpha = \vartheta_N(\theta), \tag{13}$$

where the elements of g_N and G_N are equal to those of γ_N and Γ_N with the expectations operator suppressed and the disturbances u_N replaced by (consistent) estimates $\tilde{u}_N$.

GM estimates of the parameters ρ_1, …, ρ_R , σ_v^2 and σ_1^2 are then obtained as the solution to

$$\underset{\rho_1,\rho_2,...,\rho_R,\sigma_v^2,\sigma_1^2}{\arg\min} [(g_N - G_N\alpha)' \Xi_N^{-1}(g_N - G_N\alpha)] = [\vartheta_N(\theta)' \tilde{\Xi}_N^{-1}(\vartheta_N(\theta)], \tag{14}$$

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 Ξ_N^{-1} will be discussed below.

In the following, we define the elements of γ_N and Γ_N grouped by the corresponding moment conditions. For this, let us use the following notation:

$$\bar{u}_{r,N} = (I_T \otimes W_{r,N})u_N \text{ , } m = 1, \ldots, R, \text{ and} \tag{15a}$$

$$\bar{\bar{u}}_{rm,N} = (I_T \otimes W_{r,N})(I_T \otimes W_{m,N})u_N = (I_T \otimes W_{r,N}W_{m,N})u_N \text{ , } r = 1, \ldots, R, \; m = 1, \ldots R. \tag{15b}$$

Moreover, running index $l = 1,...,R$ has to be introduced for a proper definition of the elements of Γ_N and γ_N.

Moment condition $M_{1,r}$ delivers $r = 1,...,R$ lines of equation system (12), appearing in rows $4(r-1)+1$ with the following elements of γ_N and Γ_N:

$$\gamma_{4(r-1)+1} = \frac{1}{N(T-1)} E[\bar{u}'_{r,N} Q_{0,N} \bar{u}_{r,N}] ,$$

$$\gamma_{4(r-1)+1,m} = \frac{2}{N(T-1)} E[\bar{u}'_{r,N} Q_{0,N} \bar{\bar{u}}_{rm,N}] , \; m = 1,...,R ,$$

$$\gamma_{4(r-1)+1,R+m} = -\frac{1}{N(T-1)} E[\bar{\bar{u}}'_{rm,N} Q_{0,N} \bar{\bar{u}}_{rm,N}] , \; m = 1,...,R ,$$

$$\gamma_{4(r-1)+1,R(m+1)-m(m-1)/2+l-m} = -\frac{2}{N(T-1)} E[\bar{\bar{u}}'_{rm,N} Q_{0,N} \bar{\bar{u}}_{rl,N}] , \; m = 1,...,R-1 , \; l = m+1,...,R ,$$

$$\gamma_{4(r-1)+1,2R+R(R-1)/2+1} = \frac{1}{N} tr(W'_{r,N} W_{r,N}) ,$$

$$\gamma_{4(r-1)+1,2R+R(R-1)/2+2} = 0 .$$

Moment condition $M_{2,r}$ consists of $r = 1, \ldots R$ lines of equation system (12), appearing in rows $4(r-1)+2$ with the following elements of γ_N and Γ_N:

$$\gamma_{4(r-1)+2} = \frac{1}{N(T-1)} E[\bar{u}'_{r,N} Q_{0,N} u_N] ,$$

$$\gamma_{4(r-1)+2,m} = \frac{1}{N(T-1)} E[\bar{\bar{u}}'_{rm,N} Q_{0,N} u_N + \bar{u}'_{r,N} Q_{0,N} \bar{u}_{m,N}] , \; m = 1,...,R ,$$

$$\gamma_{4(r-1)+2,R+m} = -\frac{1}{N(T-1)} E[\bar{\bar{u}}'_{rm,N} Q_{0,N} \bar{u}_{m,N}] , \; m = 1,...,R ,$$

$$\gamma_{4(r-1)+2,R(m+1)-m(m-1)/2+l-m} = -\frac{1}{N(T-1)} E[\bar{\bar{u}}'_{rl,N} Q_{0,N} \bar{u}_{m,N} + \bar{\bar{u}}'_{rm,N} Q_{0,N} \bar{u}_{l,N}] , \; m = 1,...,R-1 ,$$

$l = m+1,...,R$,

$$\gamma_{4(r-1)+2,2R+R(R-1)/2+1} = 0 ,$$

$$\gamma_{4(r-1)+2,2R+R(R-1)/2+2} = 0 .$$

Moment condition $M_{3,r}$ corresponds to r = 1, … R lines of equation system (12), appearing in rows 4(r-1)+3 with the following elements of γ_N and Γ_N:

$$\gamma_{4(r-1)+3} = \frac{1}{N} E[\bar{u}'_{r,N} Q_{1,N} \bar{u}_{r,N}],$$

$$\gamma_{4(r-1)+3,m} = \frac{2}{N} E[\bar{u}'_{r,N} Q_{1,N} \bar{\bar{u}}_{rm,N}],\ m = 1,\dots,R,$$

$$\gamma_{4(r-1)+3,R+m} = -\frac{1}{N} E[\bar{\bar{u}}'_{rm,N} Q_{1,N} \bar{\bar{u}}_{rm,N}],\ m = 1,\dots,R,$$

$$\gamma_{4(r-1)+3,R(m+1)-m(m-1)/2+l-m} = -\frac{2}{N} E[\bar{\bar{u}}'_{rm,N} Q_{1,N} \bar{\bar{u}}_{rl,N}],\ m = 1,\dots,R-1,\ l = m+1,\dots,R,$$

$$\gamma_{4(r-1)+3,2R+R(R-1)/2+1} = 0,$$

$$\gamma_{4(r-1)+3,2R+R(R-1)/2+2} = \frac{1}{N} tr(W'_{r,N} W_{r,N}).$$

Moment condition $M_{4,r}$ represents r = 1, … R lines of equation system (12) appearing in rows rows 4(r-1)+4 with the following elements of γ_N and Γ_N:

$$\gamma_{4(r-1)+4} = \frac{1}{N} E[\bar{u}'_{r,N} Q_{1,N} u_N],$$

$$\gamma_{4(r-1)+4,m} = \frac{1}{N} E[\bar{\bar{u}}'_{rm,N} Q_{1,N} u_N + \bar{u}'_{r,N} Q_{1,N} \bar{u}_{m,N}],\ m = 1,\dots,R,$$

$$\gamma_{4(r-1)+4,R+m} = -\frac{1}{N} E[\bar{\bar{u}}'_{rm,N} Q_{1,N} \bar{u}_{m,N}],\ m = 1,\dots,R,$$

$$\gamma_{4(r-1)+4,R(m+1)-m(m-1)/2+l-m} = -\frac{1}{N} E[\bar{\bar{u}}'_{rl,N} Q_{1,N} \bar{u}_{m,N} + \bar{\bar{u}}'_{rm,N} Q_{1,N} \bar{u}_{l,N}],\ m = 1,\dots,R-1,$$
$$l = m+1,\dots,R,$$

$$\gamma_{4(r-1)+4,2R+R(R-1)/2+1} = 0,$$

$$\gamma_{4(r-1)+4,2R+R(R-1)/2+2} = 0.$$

Moment condition M_a reflects 1 line of equation system (12) appearing in row (4R + 1) with the following elements of γ_N and Γ_N:

$$\gamma_{4R+1} = \frac{1}{N(T-1)} E[u'_N Q_{0,N} u_N],$$

$$\gamma_{4R+1,m} = \frac{2}{N(T-1)} E[\bar{u}'_{m,N} Q_{0,N} u_N],\ m = 1,\dots,R,$$

$$\gamma_{4R+1,R+m} = -\frac{1}{N(T-1)} E[\bar{u}'_{m,N} Q_{0,N} \bar{u}_{m,N}],\ m = 1,\dots,R,$$

$$\gamma_{4R+1,R(m+1)-m(m-1)/2+l-m} = -\frac{2}{N(T-1)} E[\bar{u}'_{m,N} Q_{0,N} \bar{u}_{l,N}],\ m = 1,\dots,R-1,\ l = m+1,\dots,R,$$

$\gamma_{4R+1,2R+R(R-1)/2+1} = 1$,

$\gamma_{4R+1,2R+R(R-1)/2+2} = 0$.

Moment condition M_b is associated with 1 line of equation system (12) appearing in row (4R + 2) with the following elements of γ_N and Γ_N:

$$\gamma_{4R+2} = \frac{1}{N} E[u'_N Q_{1,N} u_N],$$

$$\gamma_{4R+2,m} = \frac{2}{N} E[\bar{u}'_{m,N} Q_{1,N} u_N],\ m = 1,...,R,$$

$$\gamma_{4R+2,R+m} = -\frac{1}{N} E[\bar{u}'_{m,N} Q_{1,N} \bar{u}_{m,N}],\ m = 1,...,R,$$

$$\gamma_{4R+2,R(m+1)-m(m-1)/2+l-m} = -\frac{2}{N} E[\bar{u}'_{m,N} Q_{1,N} \bar{u}_{l,N}],\ m = 1,...,R-1,\ l = m+1,...,R,$$

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

$$\gamma_{4R+2,2R+R(R-1)/2+2} = 1.$$

This completes the specification of the elements of the matrices γ_N and Γ_N. The similarity of the structure between the expressions resulting from the moment conditions M_a, $M_{1,r}$ and $M_{2,r}$ on the one hand and M_b, $M_{3,r}$, $M_{4,r}$ 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 M_a, $M_{1,r}$ and $M_{2,r}$ (r = 1, …R) do not depend on σ_1^2 whereas the rows resulting from M_b, $M_{3,r}$, and $M_{4,r}$ (r = 1, …R) do not depend on σ_v^2. This fact will be used to define an initial GM estimator, which is based on a subset of moment conditions (M_a, $M_{1,r}$ and $M_{2,r}$) only, in order to obtain an estimate of the matrix Ξ_N.

For future reference, we define the $(2R+1)\times 1$ vector γ_N^0 as the sub-vector containing rows r and $(r+1)$, $r = 1,...,R$ and row $(2R+1)$ of γ_N (corresponding to $M_{1,r}$, $M_{2,r}$ and M_a). Moreover, we define the $(2R+1)\times[2R+R(R-1)/2+1]$ matrix Γ_N^0 as the sub-matrix containing rows r and $(r+1)$, $r = 1,...,R$, and row $(2R+1)$ of Γ_N (corresponding to $M_{1,r}$, $M_{2,r}$ and M_a), with the last column of Γ_N (associated with σ_1^2) deleted.

Similarly, we define the $(2R+1)\times 1$ vector γ_N^1 as the sub-vector containing rows 2r, $(2r+1)$, $r = 1,...,R$, and row $(2R+2)$ of γ_N (corresponding to $M_{3,r}$, $M_{4,r}$ and M_b). Finally, we define the $(2R+1)\times[2R+R(R-1)/2+1]$ matrix Γ_N^1 as the sub-matrix containing rows 2r, $(2r+1)$, $r = 1,...,R$, and $(2R+2)$ and Γ_N (corresponding to $M_{3,r}$, $M_{4,r}$ and M_b), with the second last column of Γ_N (associated with σ_v^2) deleted.

2. GM estimators of an *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

$$M_i^{xx} = \lim_{N\to\infty} \frac{1}{NT} X_N^*(\rho)' Q_{i,N} X_N^*(\rho) \ , \tag{16a}$$

with $X_N^*(\rho) = [I_T \otimes (I_N - \sum_{r=1}^{R} \rho_r W_{r,N})] X_N$, and the matrices

$$\lim_{N\to\infty} \frac{1}{NT} X_N' X_N \ , \ \lim_{N\to\infty} \frac{1}{NT} X_N^*(\rho)' X_N^*(\rho) , \ \lim_{N\to\infty} \frac{1}{NT} X_N^*(\rho)' \Omega_{\varepsilon,N}^{-1} X_N^*(\rho) \tag{16b}$$

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 β in (2a) involve limits of the expressions above.

Assumption 4.

The row and column sums of $W_{r,N}$, $r = 1, \dots, R$, and $P_N(\rho) = (I_N - \sum_{r=1}^{R} \rho_r W_{r,N})^{-1}$ are bounded uniformly in absolute value by $k_W < \infty$ and $k_P < \infty$, respectively, where k_P may depend on $\rho = (\rho_1, \dots, \rho_R)$. We take k_W as largest of the bounds of the weights matrices, i.e., $k_W = \max(k_{W,1}, \dots, k_{W,R})$.[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 $(\Gamma_N^0)'(\Gamma_N^0)$ and $(\Gamma_N^1)'(\Gamma_N^1)$ are bounded away from zero, i.e., $\lambda_{\min}[(\Gamma_N^i)'(\Gamma_N^i)] \geq \lambda_* > 0$ for $i = 1, 2$, where λ_* may depend on $\rho_1, \dots, \rho_R$, σ_v^2, and σ_1^2. Assumption 5 ensures identifiable uniqueness of the parameters $\rho_1, \dots, \rho_R$, σ_v^2, and σ_1^2. We show in the Appendix that Assumption 5 also implies that the smallest eigenvalue of $(\Gamma_N)'(\Gamma_N)$ is bounded away from zero.

[3] See Appendix A1 for a definition of row (column) sum boundedness.

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 SAR(1) estimation).

2.2. Initial GM estimation

The initial GM estimator is based on a subset of moment conditions (M_a, $M_{1,r}$ and $M_{2,r}$) and thus on the matrices Γ_N^0 and γ_N^0 only.[4] Define θ^0 as the corresponding parameter vector that excludes σ_1^2, i.e. $\theta^0 = (\rho, \sigma_v^2) = (\rho_1, \ldots, \rho_R, \sigma_v^2)$ and accordingly $\alpha^0 = (\rho_1, \ldots, \rho_R, \rho_1^2, \ldots, \rho_R^2, \rho_1\rho_2, \ldots, \rho_1\rho_R, \ldots, \rho_{R-1}\rho_R, \sigma_v^2)'$.

The initial GM estimator is then obtained as the solution to

$$(\tilde{\rho}_{1,N}, \ldots, \tilde{\rho}_{R,N}, \tilde{\sigma}_{v,N}^2) = \arg\min\{\vartheta_N^0(\underline{\theta}^0)' \vartheta_N^0(\underline{\theta}^0), \underline{\rho} \in [-a, a], \underline{\sigma}_v^2 \in [0, b]\}, \tag{17a}$$

where $a \geq 1$, $b \geq b_v$ and $\vartheta_N^0(\underline{\theta}^0) = \vartheta_N^0(\underline{\rho}, \underline{\sigma}_v^2) = (g_N^0 - G_N^0 \underline{\alpha}^0)$.

Using these initial estimates of ρ_1, …, ρ_R and σ_v^2, σ_1^2 can be estimated from moment condition M_b:

$$\tilde{\sigma}_{1,N}^2 = \frac{1}{N}(\tilde{u}_N - \sum_{m=1}^{R} \tilde{\rho}_{m,N} \tilde{\bar{u}}_{m,N})' Q_{1,N} (\tilde{u}_N - \sum_{m=1}^{R} \tilde{\rho}_{m,N} \tilde{\bar{u}}_{m,N}) \tag{17b}$$

$$= g_{4R+2} - g_{4R+2,1}\tilde{\rho}_{1,N} - \ldots - g_{4R+2,R}\tilde{\rho}_{R,N} - g_{4R+2,R+1}\tilde{\rho}_{1,N}^2$$
$$- g_{4R+2,2R}\tilde{\rho}_{R,N}^2 - g_{4R+2,2R+1}\tilde{\rho}_{1,N}\tilde{\rho}_{2,N} - \ldots - g_{4R+2,2R+R(R-1)/2}\tilde{\rho}_{R-1,N}\tilde{\rho}_{R,N}.$$

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 $\Xi_N^{-1} = [NVar(\eta_N)]^{-1}$, where η_N denotes the $[(4R+2) \times 1]$ vector of moments, evaluated at the true parameter values. Suppressing the expectations operator (and ignoring the deterministic constants), the elements of η_N correspond to the expressions on the left hand side in equation (11). Notice, however, that the

[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 Γ_N^0 and γ_N^0. Hence, in the following, the initial GM estimator will be used to obtain an estimate of Ξ_N^{-1}.

ordering in equation system (12) is different, where the rows corresponding to M_a and M_b are placed in the last two rows, such that

$$\eta_N = \begin{bmatrix} \frac{1}{N(T-1)} \bar{\varepsilon}'_{r,N} Q_{0,N} \bar{\varepsilon}_{r,N} \\ \frac{1}{N(T-1)} \bar{\varepsilon}'_{r,N} Q_{0,N} \varepsilon_N \\ \frac{1}{N} \bar{\varepsilon}'_{r,N} Q_{1,N} \bar{\varepsilon}_{r,N} \\ \frac{1}{N} \bar{\varepsilon}'_{r,N} Q_{1,N} \varepsilon_N \\ . \\ . \\ \frac{1}{N(T-1)} \varepsilon'_N Q_{0,N} \varepsilon_N \\ \frac{1}{N} \varepsilon'_N Q_{1,N} \varepsilon_N \end{bmatrix}. \tag{18}$$

Notice that the matrix $\Xi_N = NVar(\eta_N)$ is symmetric and of order $(4R + 2) \times (4R + 2)$. Its elements are derived by substituting $\bar{\varepsilon}_{r,N} = (I_T \otimes W_{r,N})\varepsilon_N$ and using the results that $Cov(\varepsilon'_N A_N \varepsilon_N, \varepsilon'_N B_N \varepsilon_N) = 2tr(A_N \Omega_{\varepsilon,N} B_N \Omega_{\varepsilon,N})$ for two non-negative definite symmetric matrices A_N and B_N and $\varepsilon_N \sim N(0, \Omega_{\varepsilon,N})$, see, e.g., Amemiya (1973, p. 5).[5]

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

$$\Xi_N = \begin{bmatrix} \xi_{1,1} & . & \xi_{1,R} & \xi_{1,R+1} \\ . & & & . \\ \xi_{R,1} & & \xi_{R,R} & \xi_{R,R+1} \\ \xi_{R+1,1} & & \xi_{R+1,R} & \xi_{R+1,R+1} \end{bmatrix}. \tag{19}$$

Subscript N has been dropped from the elements $\xi_{r,s}$ here to simplify notation. The matrix Ξ_N is made up of three parts.

i) An upper left block of dimension $4R \times 4R$, consisting of R^2 blocks of order 4×4, which are defined as

[5] For quadratic forms in non-symmetric matrices A_N (or B_N) we use the fact that $\varepsilon'_N A_N \varepsilon_N = \varepsilon'_N A'_N \varepsilon_N = \varepsilon'_N (A_N + A'_N)\varepsilon_N / 2$, which is a quadratic form in the symmetric matrix $(A_N + A'_N)/2$.

$$\xi_{r,s} = C \otimes T^{W}_{r,s}\text{, } r, s = 1, \ldots R\text{, where} \tag{20}$$

$$C = \begin{bmatrix} \frac{1}{(T-1)}\sigma_v^4 & 0 \\ 0 & \sigma_1^4 \end{bmatrix}\text{, and} \tag{21}$$

$$T^{W}_{r,s} = \begin{bmatrix} \frac{2}{N} tr(W'_{r,N} W_{r,N} W'_{s,N} W_{s,N}) & \frac{1}{N} tr(W'_{r,N} W_{r,N} (W'_{s,N} + W_{s,N})) \\ \frac{1}{N} tr(W'_{s,N} W_{s,N} (W'_{r,N} + W_{r,N})) & \frac{1}{N} tr(W_{r,N} W_{s,N} + W'_{r,N} W_{s,N}) \end{bmatrix}\text{, } r, s = 1, \ldots, R. \tag{22}$$

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

$$\xi_{R+1,s} = C \otimes t^{W}_{R+1,s}\text{, and } \xi_{s,R+1} = (\xi_{R+1,s})'\text{, } s = 1, \ldots, R\text{, with} \tag{23}$$

$$t^{W}_{R+1,s} = \begin{bmatrix} \frac{2}{N} tr(W'_{s,N} W_{s,N}) & 0 \end{bmatrix}\text{, } s = 1, \ldots, R. \tag{24}$$

iii) The lower right block of order 2×2, defined as

$$\xi_{R+1,R+1} = C \otimes t^{W}_{R+1,R+1} = C \otimes 2\,. \tag{25}$$

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

2.3.2 The 'weighted GM estimator'

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

$$(\hat{\rho}_{1,N}, \ldots, \hat{\rho}_{R,N}, \hat{\sigma}^2_{v,N}, \hat{\sigma}^2_{1,N}) = \arg\min \{\vartheta_N(\underline{\theta})' \tilde{\Xi}_N^{-1} \vartheta_N(\underline{\theta}), \underline{\rho} \in [-a, a], \underline{\sigma}_v^2 \in [0, b], \underline{\sigma}_1^2 \in [0, c]\}, \tag{26}$$

where $a \geq 1$, $b \geq b_v$, $c \geq T\, b_\mu + b_v$, and $\vartheta_N(\underline{\theta}) = \vartheta_N(\underline{\rho}, \underline{\sigma}_v^2, \underline{\sigma}_1^2) = (g_N - G_N \underline{\alpha})$.

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 (M_a, $M_{1,r}$ and $M_{2,r}$) and the same weight for the three other moment conditions (M_b, $M_{3,r}$, and $M_{4,r}$), 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 Ξ_N^p is given by

$$\Xi_N^p = \begin{bmatrix} C\otimes I_2 & 0 & . & 0 & 0 \\ 0 & C\otimes I_2 & & & 0 \\ . & & . & & . \\ 0 & & & C\otimes I_2 & 0 \\ 0 & 0 & . & . & C\otimes 1 \end{bmatrix}, \quad (27)$$

and the partially weighted GM estimator is defined as

$$(\check{\rho}_{1,N},...,\check{\rho}_{R,N},\check{\sigma}^2_{v,N},\check{\sigma}^2_{1,N}) = \arg\min\{\vartheta_N(\underline{\theta})'(\tilde{\Xi}_N^p)^{-1}\,\vartheta_N(\underline{\theta}), \underline{\rho}\in[-a,a], \underline{\sigma}_v^2\in[0,b], \underline{\sigma}_1^2\in[0,c]\}, \quad (28)$$

where $a \geq 1$, $b \geq b_v$, $c \geq T\,b_\mu + b_v$, and $\vartheta_N(\underline{\theta}) = \vartheta_N(\underline{\rho},\underline{\sigma}_v^2,\underline{\sigma}_1^2) = (g_N - G_N\underline{\alpha})$.

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 $\tilde{\beta}_N$ is a consistent estimator of β, the initial GM estimators $(\tilde{\rho}_1,...,\tilde{\rho}_R,\tilde{\sigma}_v^2,\tilde{\sigma}_1^2)$ defined by (17a) and (17b) are consistent for $\rho_1,...,\rho_R,\sigma_v^2,\sigma_1^2$, i.e., $(\tilde{\rho}_{1,N},...,\tilde{\rho}_{R,N},\tilde{\sigma}^2_{v,N},\tilde{\sigma}^2_{1,N}) \xrightarrow{P} (\rho_1,...,\rho_R,\sigma_v^2,\sigma_1^2)$ as $N\to\infty$.

Theorem 2. *Consistency of weighted GM estimators*

Suppose Assumptions 1-5 hold and that the smallest and largest eigenvalues of the matrices Ξ_N^{-1} satisfy $0 < \underline{\lambda}_* \leq \lambda_{\min}(\Xi_N^{-1}) \leq \lambda_{\max}(\Xi_N^{-1}) \leq \overline{\lambda}_{**} < \infty$. Suppose furthermore that $\tilde{\beta}_N$ and $\tilde{\Xi}_N$ are consistent estimators of β and Ξ_N, respectively. Then the weighted GM estimators $(\hat{\rho}_{1,N},...,\hat{\rho}_{R,N},\hat{\sigma}^2_{v,N},\hat{\sigma}^2_{1,N})$ defined by (26) are consistent for $\rho_1,...,\rho_R,\sigma_v^2,\sigma_1^2$, i.e., $(\hat{\rho}_{1,N},...,\hat{\rho}_{R,N},\hat{\sigma}^2_{v,N},\hat{\sigma}^2_{1,N}) \xrightarrow{P} (\rho_1,...,\rho_R,\sigma_v^2,\sigma_1^2)$ as $N\to\infty$.

Theorem 3. *Consistency of partially weighted GM estimators*

Suppose Assumptions 1-5 hold. Suppose furthermore that $\tilde{\beta}$ and $\tilde{\Xi}_N^p$ are consistent estimators of β and Ξ_N^p, respectively. Then, the partially weighted GM estimators $(\check{\rho}_{1,N},...,\check{\rho}_{R,N},\check{\sigma}^2_{v,N},\check{\sigma}^2_{1,N})$ defined by (28) are consistent for $\rho_1,...,\rho_R,\sigma_v^2,\sigma_1^2$, i.e.,

$(\check{\rho}_{1,N},...,\check{\rho}_{R,N},\check{\sigma}^2_{v,N},\check{\sigma}^2_{1,N}) \xrightarrow{P} (\rho_1,...,\rho_R,\sigma_v^2,\sigma_1^2)$ as $N \to \infty$.

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* β
Suppose Assumptions 1-4 holds. The OLS estimator of β based on (2a), which is given by $\hat{\beta}_N^{OLS} = (X_N' X_N)^{-1} X_N' y_N$, is consistent for β, i.e., $\hat{\beta}_N^{OLS} \xrightarrow{P} \beta$ as $N \to \infty$.

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

$$\hat{\beta}_N^{GLS} = \{X_N'[\Omega_{u,N}^{-1}(\rho,\sigma_v^2,\sigma_1^2)]X_N\}^{-1} X_N'[\Omega_{u,N}^{-1}(\rho,\sigma_v^2,\sigma_1^2)]y_N . \quad (29a)$$

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

$$\hat{\beta}_N^{GLS} = \{X_N^{*\prime}(\rho)[\Omega_{\varepsilon,N}^{-1}(\sigma_v^2,\sigma_1^2)]X_N^*(\rho)\}^{-1} X_N^{*\prime}(\rho)[\Omega_{\varepsilon,N}^{-1}(\sigma_v^2,\sigma_1^2)]y_N^*(\rho), \text{ where} \quad (29b)$$

$$X_N^*(\rho) = [I_T \otimes (I_N - \sum_{m=1}^{R} \rho_m W_{m,N})]X_N , \quad (30a)$$

$$y_N^*(\rho) = [I_T \otimes (I_N - \sum_{m=1}^{R} \rho_m W_{m,N})]y_N . \quad (30b)$$

The feasible generalized least squares (FGLS) estimator is obtained by replacing the true parameters ρ, σ_v^2, and σ_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.,

$$(NT)^{1/2}[\hat{\beta}_N^{GLS} - \beta] \xrightarrow{D} N\{0,\Psi\} \text{ as } N \to \infty, \text{ with}$$
$$\Psi = [\sigma_v^{-2} M_0^{xx} + \sigma_1^{-2} M_1^{xx}]^{-1}.$$

[6] In particular, the proofs of consistency of OLS in KKP (2007, p. 124) and the proof of Theorem 4 in KKP (2007, p. 126).

(b) Let $\ddot{\rho}_N, \ddot{\sigma}^2_{v,N}$, and $\ddot{\sigma}^2_{1,N}$ be (any) consistent estimates of ρ, σ^2_v, and σ^2_1. Then,

$$(NT)^{1/2}[\hat{\beta}_N^{GLS} - \hat{\beta}_N^{FGLS}] \xrightarrow{P} 0 \text{ as } N \to \infty .$$

This means that GLS and FGLS are asymptotically equivalent and, hence, $\hat{\beta}_N^{FGLS}$ is also consistent and asymptotically normal.

(c) Furthermore,

$$\dot{\Psi}_N - \Psi \xrightarrow{P} 0 \text{ as } N \to \infty \text{, where}$$
$$\dot{\Psi}_N = \left\{ \frac{1}{NT} X_N^{*\prime}(\dot{\rho}_N)[\Omega^{-1}_{\varepsilon,N}(\dot{\sigma}^2_{v,N}, \dot{\sigma}^2_{1,N})] X_N^*(\dot{\rho}_N) \right\}^{-1} .$$

This suggests that small sample inference can be based on the approximation $\hat{\beta}_N^{FGLS} \sim N(\beta, (NT)^{-1}\dot{\Psi}_N)$.

While we demonstrate in the Appendix that the initial, weighted, and partially weighted GM estimators of ρ, σ^2_v, and σ^2_1 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.,

$$u_N = \sum_{m=1}^{3} \rho_m (I_T \otimes W_m) u_N + \varepsilon_N . \tag{31}$$

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 time-invariant, 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 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 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 row-normalizing W_1^0, W_2^0, and W_3^0, that is by dividing their elements $w_{1,ij}^0$, $w_{2,ij}^0$, and $w_{2,ij}^0$ 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 ρ_1, ρ_2, and ρ_3 must satisfy $0 \leq |\rho_1| + |\rho_2| + |\rho_3| < 1$ 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 (ρ_1, ρ_2, ρ_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 constellation	ρ_1	ρ_2	ρ_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 ε_N, we assume that $\sigma_\mu^2 = \sigma_v^2 = 1$, i.e., the error components μ_N and v_{it} are drawn from a standard normal distribution. For each Monte Carlo experiment we consider 2000 draws. To ensure comparability, the same draw of μ and v is used for each of the 10 combinations of ρ_1, ρ_2, and ρ_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 ρ_1, ρ_2, ρ_3, σ_μ^2 and σ_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 (GM$^{\text{in}}$), the weighted GM estimator (GM$^{\text{w}}$), and the partially weighted GM estimator (GM$^{\text{in}}$).

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]		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	average [2]
ρ_1		0.4	0.4	0.4	0.4	0.4	0.2	0.2	0.2	0.2	0	0.2800
Bias	GM^{in}	-0.0049	-0.0058	-0.0067	-0.0073	-0.0093	-0.0076	-0.0097	-0.0088	-0.0105	-0.0115	0.0082
	GM^{w}	0.0126	0.0142	0.0137	0.0133	0.0140	0.0145	0.0144	0.0141	0.0147	0.0147	0.0140
	GM^{pw}	0.0022	0.0026	0.0026	0.0025	0.0029	0.0024	0.0027	0.0025	0.0028	0.0026	0.0026
RMSE	GM^{in}	0.0652	0.0703	0.0691	0.0681	0.0722	0.0789	0.0806	0.0781	0.0834	0.0936	0.0759
	GM^{w}	0.0592	0.0649	0.0634	0.0621	0.0659	0.0722	0.0732	0.0709	0.0758	0.0847	0.0692
	GM^{pw}	0.0586	0.0632	0.0618	0.0607	0.0641	0.0705	0.0716	0.0694	0.0740	0.0832	0.0677
ρ_2		0.4	0.2	0.2	0.2	0	0.2	0.1	0.2	0	0	0.1500
Bias	GM^{in}	-0.0006	0.0005	-0.0003	-0.0010	-0.0012	-0.0015	-0.0028	-0.0028	-0.0028	-0.0043	0.0018
	GM^{w}	-0.0064	-0.0078	-0.0071	-0.0063	-0.0058	-0.0066	-0.0052	-0.0052	-0.0051	-0.0046	0.0060
	GM^{pw}	-0.0009	-0.0014	-0.0013	-0.0012	-0.0014	-0.0009	-0.0008	-0.0006	-0.0010	-0.0007	0.0010
RMSE	GM^{in}	0.0814	0.0795	0.0805	0.0813	0.0795	0.0824	0.0850	0.0848	0.0847	0.0904	0.0829
	GM^{w}	0.0754	0.0745	0.0754	0.0761	0.0745	0.0768	0.0795	0.0791	0.0792	0.0846	0.0775
	GM^{pw}	0.0749	0.0732	0.0742	0.0749	0.0733	0.0759	0.0783	0.0781	0.0780	0.0833	0.0764
ρ_3		0	0.2	0.1	0	0	0.2	0	0	0	0	0.0500
Bias	GM^{in}	-0.0012	-0.0023	-0.0023	-0.0023	-0.0033	-0.0031	-0.0033	-0.0028	-0.0038	-0.0043	0.0029
	GM^{w}	-0.0051	-0.0051	-0.0038	-0.0026	-0.0006	-0.0041	-0.0017	-0.0026	-0.0010	-0.0013	0.0028
	GM^{pw}	-0.0005	-0.0003	0.0000	0.0001	0.0005	0.0001	0.0002	0.0001	0.0003	0.0002	0.0002
RMSE	GM^{in}	0.0702	0.0713	0.0725	0.0730	0.0741	0.0792	0.0809	0.0804	0.0812	0.0875	0.0770
	GM^{w}	0.0649	0.0658	0.0671	0.0675	0.0684	0.0735	0.0751	0.0747	0.0753	0.0813	0.0714
	GM^{pw}	0.0646	0.0657	0.0669	0.0675	0.0686	0.0731	0.0750	0.0745	0.0753	0.0813	0.0713
σ_v^2		1	1	1	1	1	1	1	1	1	1	1.0000
Bias	GM^{in}	-0.0110	-0.0111	-0.0114	-0.0116	-0.0120	-0.0121	-0.0128	-0.0126	-0.0129	-0.0137	0.0121
	GM^{w}	-0.0122	-0.0120	-0.0123	-0.0124	-0.0126	-0.0111	-0.0114	-0.0113	-0.0115	-0.0105	0.0117
	GM^{pw}	-0.0121	-0.0121	-0.0122	-0.0122	-0.0123	-0.0120	-0.0120	-0.0120	-0.0120	-0.0118	0.0121
RMSE	GM^{in}	0.0723	0.0719	0.0717	0.0716	0.0714	0.0713	0.0709	0.0711	0.0709	0.0708	0.0714
	GM^{w}	0.0721	0.0717	0.0716	0.0716	0.0714	0.0710	0.0707	0.0709	0.0707	0.0704	0.0712
	GM^{pw}	0.0721	0.0717	0.0716	0.0715	0.0713	0.0711	0.0707	0.0709	0.0707	0.0705	0.0712
σ_1^2		6	6	6	6	6	6	6	6	6	6	6.0000
Bias	GM^{in}	0.0218	0.0209	0.0179	0.0158	0.0125	0.0135	0.0082	0.0100	0.0068	0.0020	0.0129
	GM^{w}	-0.0914	-0.0909	-0.0923	-0.0933	-0.0951	-0.0847	-0.0869	-0.0860	-0.0879	-0.0816	0.0890
	GM^{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^{in}	0.8745	0.8712	0.8702	0.8700	0.8675	0.8667	0.8647	0.8661	0.8637	0.8616	0.8676
	GM^{w}	0.8641	0.8624	0.8622	0.8621	0.8610	0.8603	0.8588	0.8595	0.8584	0.8571	0.8606
	GM^{pw}	0.8649	0.8630	0.8628	0.8627	0.8615	0.8610	0.8596	0.8604	0.8591	0.8580	0.8613

Note: GM^{in}, GM^{w}, GM^{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 3. Monte Carlo Results, N = 250, T = 5, 2000 draws

Parameter Constellation [1]		(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	average [2]
ρ_1		0.4	0.4	0.4	0.4	0.4	0.2	0.2	0.2	0.2	0	0.2800
Bias	GM^{in}	-0.0028	-0.0033	-0.0037	-0.0039	-0.0048	-0.0043	-0.0052	-0.0047	-0.0055	-0.0061	0.0044
	GM^{w}	0.0060	0.0068	0.0065	0.0063	0.0066	0.0070	0.0069	0.0067	0.0070	0.0070	0.0067
	GM^{pw}	0.0010	0.0012	0.0012	0.0012	0.0013	0.0010	0.0012	0.0011	0.0012	0.0010	0.0011
RMSE	GM^{in}	0.0399	0.0431	0.0419	0.0410	0.0429	0.0478	0.0480	0.0467	0.0495	0.0555	0.0456
	GM^{w}	0.0359	0.0394	0.0383	0.0373	0.0392	0.0435	0.0436	0.0424	0.0450	0.0502	0.0415
	GM^{pw}	0.0361	0.0390	0.0378	0.0369	0.0386	0.0431	0.0432	0.0420	0.0445	0.0499	0.0411
ρ_2		0.4	0.2	0.2	0.2	0	0.2	0.1	0.2	0	0	0.1500
Bias	GM^{in}	-0.0005	0.0000	-0.0004	-0.0007	-0.0009	-0.0009	-0.0015	-0.0015	-0.0016	-0.0023	0.0010
	GM^{w}	-0.0036	-0.0042	-0.0039	-0.0035	-0.0032	-0.0037	-0.0029	-0.0030	-0.0029	-0.0027	0.0033
	GM^{pw}	-0.0012	-0.0014	-0.0014	-0.0013	-0.0014	-0.0012	-0.0012	-0.0011	-0.0013	-0.0012	0.0013
RMSE	GM^{in}	0.0497	0.0487	0.0493	0.0498	0.0487	0.0506	0.0522	0.0521	0.0520	0.0556	0.0509
	GM^{w}	0.0443	0.0437	0.0442	0.0446	0.0436	0.0453	0.0467	0.0466	0.0465	0.0498	0.0455
	GM^{pw}	0.0444	0.0435	0.0440	0.0444	0.0434	0.0452	0.0466	0.0465	0.0464	0.0497	0.0454
ρ_3		0	0.2	0.1	0	0	0.2	0	0	0	0	0.0500
Bias	GM^{in}	-0.0005	0.0000	-0.0004	-0.0007	-0.0009	-0.0009	-0.0015	-0.0015	-0.0016	-0.0023	0.0010
	GM^{w}	-0.0036	-0.0042	-0.0039	-0.0035	-0.0032	-0.0037	-0.0029	-0.0030	-0.0029	-0.0027	0.0033
	GM^{pw}	-0.0012	-0.0014	-0.0014	-0.0013	-0.0014	-0.0012	-0.0012	-0.0011	-0.0013	-0.0012	0.0013
RMSE	GM^{in}	0.0497	0.0487	0.0493	0.0498	0.0487	0.0506	0.0522	0.0521	0.0520	0.0556	0.0509
	GM^{w}	0.0443	0.0437	0.0442	0.0446	0.0436	0.0453	0.0467	0.0466	0.0465	0.0498	0.0455
	GM^{pw}	0.0444	0.0435	0.0440	0.0444	0.0434	0.0452	0.0466	0.0465	0.0464	0.0497	0.0454
σ_v^2		1	1	1	1	1	1	1	1	1	1	1.0000
Bias	GM^{in}	-0.0043	-0.0045	-0.0046	-0.0046	-0.0048	-0.0050	-0.0052	-0.0051	-0.0053	-0.0057	0.0049
	GM^{w}	-0.0052	-0.0052	-0.0053	-0.0054	-0.0055	-0.0047	-0.0049	-0.0048	-0.0050	-0.0045	0.0051
	GM^{pw}	-0.0051	-0.0052	-0.0052	-0.0052	-0.0053	-0.0052	-0.0051	-0.0051	-0.0052	-0.0051	0.0052
RMSE	GM^{in}	0.0464	0.0462	0.0461	0.0460	0.0460	0.0458	0.0456	0.0457	0.0456	0.0455	0.0459
	GM^{w}	0.0463	0.0461	0.0460	0.0460	0.0459	0.0458	0.0456	0.0456	0.0456	0.0454	0.0458
	GM^{pw}	0.0463	0.0462	0.0460	0.0460	0.0459	0.0459	0.0456	0.0457	0.0456	0.0455	0.0459
σ_1^2		6	6	6	6	6	6	6	6	6	6	6.0000
Bias	GM^{in}	0.0095	0.0088	0.0083	0.0080	0.0072	0.0059	0.0047	0.0052	0.0043	0.0017	0.0064
	GM^{w}	-0.0352	-0.0349	-0.0354	-0.0357	-0.0364	-0.0318	-0.0326	-0.0323	-0.0330	-0.0300	0.0337
	GM^{pw}	-0.0336	-0.0341	-0.0340	-0.0338	-0.0341	-0.0334	-0.0331	-0.0330	-0.0332	-0.0325	0.0335
RMSE	GM^{in}	0.5331	0.5323	0.5320	0.5320	0.5317	0.5308	0.5301	0.5303	0.5300	0.5293	0.5311
	GM^{w}	0.5290	0.5290	0.5287	0.5284	0.5284	0.5280	0.5273	0.5273	0.5274	0.5269	0.5280
	GM^{pw}	0.5297	0.5297	0.5293	0.5290	0.5291	0.5288	0.5281	0.5281	0.5282	0.5280	0.5288

Note: GM^{in}, GM^{w}, GM^{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)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	average [2]
ρ_1		0.4	0.4	0.4	0.4	0.4	0.2	0.2	0.2	0.2	0	0.2800
Bias	GM^{in}	-0.0025	-0.0028	-0.0029	-0.0028	-0.0031	-0.0032	-0.0033	-0.0032	-0.0034	-0.0037	0.0031
	GM^{w}	0.0021	0.0023	0.0023	0.0023	0.0025	0.0024	0.0025	0.0024	0.0026	0.0026	0.0024
	GM^{pw}	-0.0006	-0.0006	-0.0005	-0.0004	-0.0002	-0.0007	-0.0003	-0.0004	-0.0003	-0.0004	0.0004
RMSE	GM^{in}	0.0281	0.0304	0.0297	0.0291	0.0306	0.0341	0.0344	0.0334	0.0355	0.0399	0.0325
	GM^{w}	0.0241	0.0265	0.0258	0.0252	0.0267	0.0295	0.0298	0.0288	0.0308	0.0344	0.0282
	GM^{pw}	0.0247	0.0268	0.0261	0.0255	0.0269	0.0299	0.0301	0.0293	0.0311	0.0350	0.0285
ρ_2		0.4	0.2	0.2	0.2	0	0.2	0.1	0.2	0	0	0.1500
Bias	GM^{in}	0.0008	0.0011	0.0009	0.0008	0.0007	0.0008	0.0005	0.0005	0.0005	0.0004	0.0007
	GM^{w}	-0.0009	-0.0011	-0.0010	-0.0008	-0.0006	-0.0008	-0.0004	-0.0004	-0.0004	-0.0002	0.0007
	GM^{pw}	0.0006	0.0005	0.0005	0.0005	0.0004	0.0007	0.0007	0.0007	0.0006	0.0007	0.0006
RMSE	GM^{in}	0.0342	0.0335	0.0339	0.0343	0.0335	0.0347	0.0359	0.0358	0.0358	0.0383	0.0350
	GM^{w}	0.0304	0.0300	0.0304	0.0308	0.0302	0.0310	0.0322	0.0320	0.0322	0.0344	0.0314
	GM^{pw}	0.0308	0.0301	0.0305	0.0308	0.0301	0.0312	0.0322	0.0321	0.0321	0.0343	0.0314
ρ_3		0	0.2	0.1	0	0	0.2	0	0	0	0	0.0500
Bias	GM^{in}	0.0008	0.0011	0.0009	0.0008	0.0007	0.0008	0.0005	0.0005	0.0005	0.0004	0.0007
	GM^{w}	-0.0009	-0.0011	-0.0010	-0.0008	-0.0006	-0.0008	-0.0004	-0.0004	-0.0004	-0.0002	0.0007
	GM^{pw}	0.0006	0.0005	0.0005	0.0005	0.0004	0.0007	0.0007	0.0007	0.0006	0.0007	0.0006
RMSE	GM^{in}	0.0342	0.0335	0.0339	0.0343	0.0335	0.0347	0.0359	0.0358	0.0358	0.0383	0.0350
	GM^{w}	0.0304	0.0300	0.0304	0.0308	0.0302	0.0310	0.0322	0.0320	0.0322	0.0344	0.0314
	GM^{pw}	0.0308	0.0301	0.0305	0.0308	0.0301	0.0312	0.0322	0.0321	0.0321	0.0343	0.0314
σ_v^2		1	1	1	1	1	1	1	1	1	1	1.0000
Bias	GM^{in}	-0.0017	-0.0017	-0.0017	-0.0017	-0.0017	-0.0020	-0.0020	-0.0020	-0.0020	-0.0022	0.0019
	GM^{w}	-0.0021	-0.0020	-0.0021	-0.0021	-0.0021	-0.0018	-0.0018	-0.0018	-0.0018	-0.0016	0.0019
	GM^{pw}	-0.0020	-0.0021	-0.0020	-0.0020	-0.0019	-0.0021	-0.0019	-0.0020	-0.0019	-0.0019	0.0020
RMSE	GM^{in}	0.0330	0.0329	0.0329	0.0328	0.0328	0.0327	0.0326	0.0326	0.0326	0.0327	0.0328
	GM^{w}	0.0328	0.0327	0.0327	0.0327	0.0328	0.0326	0.0326	0.0326	0.0326	0.0326	0.0327
	GM^{pw}	0.0329	0.0328	0.0327	0.0327	0.0327	0.0326	0.0326	0.0326	0.0326	0.0326	0.0327
σ_1^2		6	6	6	6	6	6	6	6	6	6	6.0000
Bias	GM^{in}	0.0015	0.0011	0.0011	0.0011	0.0012	-0.0007	-0.0006	-0.0006	-0.0005	-0.0020	0.0010
	GM^{w}	-0.0199	-0.0198	-0.0199	-0.0199	-0.0199	-0.0186	-0.0185	-0.0186	-0.0186	-0.0174	0.0191
	GM^{pw}	-0.0202	-0.0203	-0.0200	-0.0197	-0.0194	-0.0203	-0.0195	-0.0197	-0.0194	-0.0194	0.0198
RMSE	GM^{in}	0.3760	0.3760	0.3756	0.3755	0.3756	0.3749	0.3743	0.3743	0.3745	0.3740	0.3751
	GM^{w}	0.3754	0.3755	0.3751	0.3748	0.3747	0.3746	0.3738	0.3739	0.3738	0.3734	0.3745
	GM^{pw}	0.3761	0.3762	0.3758	0.3755	0.3754	0.3753	0.3744	0.3745	0.3745	0.3740	0.3752

Note: GM^{in}, GM^{w}, GM^{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 (GM^{in}), the average absolute bias amounts to 2.9 percent for ρ_1, to 0.6 percent for ρ_2, and to 1.0 percent for ρ_3. The average absolute bias is slightly larger with the weighted GM estimator (GM^{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 ρ_1, ρ_2, and ρ_3, we observe that the partially weighted GM estimator (GM^{pw}) performs as well as the weighted GM estimator in fairly small samples with $N=100$. The RMSE of GM^{in} is relatively larger than that of GM^{pw} and GM^{w}. The RMSE of GM^{w} tends to decline faster with an increase in N than that of GM^{pw} and GM^{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, 509-533.

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,* 2nd 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 first-order 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 $\rho = (\rho_1,...,\rho_R)'$, the admissible parameter space needs to be defined differently (Assumption 2). Finally, it has to be accounted for the higher dimension of Γ_N, γ_N, and Ξ_N, when considering the eigenvalues, e.g., of $\Gamma'_N \Gamma_N$ (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 $kN \times kN$ 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

$$\max_{1\leq i\leq kN} \sum_{j=1}^{kN} \left| b_{ij,N} \right| \leq c \text{ and } \max_{1\leq j\leq kN} \sum_{i=1}^{kN} \left| b_{ij,N} \right| \leq c \text{ for all } N \geq 1. \tag{A.1}$$

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) $kN \times kN$ 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) $kN \times p$ matrices whose elements are uniformly bounded in absolute value, then so are the elements of $A_N Z_N$ and $(kN)^{-1} Z'_N A_N Z_N$.

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 $\varepsilon_N = (e_T \otimes I_N)\mu_N + \nu_N$, where μ_N and ν_N satisfy Assumption 1. Consider the quadratic form

$$\varphi_N = N^{-1} \varepsilon'_N (S_T \otimes R_N) \varepsilon_N . \tag{A.2}$$

[7] The lemmata are derived in KKP (2007) and adapted here to the higher order setting.

Then $E(\varphi_N) = \mathrm{O}(1)$ and $Var(\varphi_N) = \mathrm{o}(1)$, and as a consequence

$$\varphi_N - E(\varphi_N) \xrightarrow{P} 0 \text{ as } N \to \infty.$$

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 Γ_N and γ_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

$$G_N^* - \Gamma_N \xrightarrow{P} 0 \text{ and } g_N^* - \gamma_N \xrightarrow{P} 0 \text{ as } N \to \infty.$$

Proof.

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

$$u_N = [I_T \otimes (I_N - \sum_{m=1}^{R} \rho_m W_{m,N})^{-1}]\varepsilon_N = (I_T \otimes P_N)\varepsilon_N,$$

$$\bar{u}_{m,N} = (I_T \otimes W_{m,N})u_N = (I_T \otimes W_{m,N}P_N)\varepsilon_N, \; m = 1, \dots, R, \text{ and}$$

$$\bar{\bar{u}}_{ml,N} = (I_T \otimes W_{m,N}W_{l,N})u_N = (I_T \otimes W_{m,N}W_{l,N}P_N)\varepsilon_N, \; m, l = 1, \dots, R.$$

Define

$$S_{0,T} = \frac{1}{T-1}(I_T - \frac{J_T}{T}), \; S_{1,T} = \frac{J_T}{T}, \text{ such that} \tag{A.3}$$

$$S_{0,T} \otimes I_N = \frac{1}{(T-1)} Q_{0,N} \text{ and } S_{1,T} \otimes I_N = Q_{1,N}. \tag{A.4}$$

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

$$\varphi_{ij,N} = \frac{1}{N}\varepsilon_N'(S_{0,T} \otimes R_{ij,N})\varepsilon_N, \tag{A.5}$$

where i and j refer to the row (and column) of the respective element of g_N^* and G_N^*.[8]

---

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

Associated with moment condition M_1, for each r = 1, …R we have:

$$\varphi_{4(r-1)+1} = \frac{1}{N}\varepsilon'_N(S_{0,N} \otimes R_{4(r-1)+1})\varepsilon_N\,,\ R_{4(r-1)+1} = P'_N W'_{m,N} W_{r,N} P_N\,, \tag{A.6}$$

$$\varphi_{4(r-1)+1,m} = \frac{1}{N}\varepsilon'_N(S_{0,T} \otimes R_{4(r-1)+1,m})\varepsilon_N\,,\ R_{4(r-1)+1} = P'_N W'_{r,N} W_{r,N} W_{m,N} P_N\,,\ m = 1, \ldots, R,$$

$$\varphi_{4(r-1)+1,R+m} = \frac{1}{N}\varepsilon'_N(S_{0,T} \otimes R_{4(r-1)+1,R+m})\varepsilon_N\,,\ R_{4(r-1)+1,R+m} = P'_N W'_{m,N} W'_{r,N} W_{r,N} W_{m,N} P_N\,,$$

m = 1, …, R

$$\varphi_{4(r-1)+1,R(m+1)-m(m-1)/2+l-m} = \frac{1}{N}\varepsilon'_N(S_{0,T} \otimes R_{4(r-1)+1,R(m+1)-m(m-1)/2+l-m})\varepsilon_N\,,$$

$R_{4(r-1)+1,R(m+1)-m(m-1)/2+l-m} = P'_N W'_{m,N} W'_{r,N} W_{r,N} W_{l,N} P_N$, m = 1, …, R-1, l = m +1, … R.

Associated with moment condition M_2, for each r = 1, …R we have:

$$\varphi_{4(r-1)+2} = \frac{1}{N}\varepsilon'_N(S_{0,T} \otimes R_{4(r-1)+2})\varepsilon_N\,,\ R_{4(r-1)+2} = P'_N W'_{r,N} P_N\,,$$

$$\varphi_{4(r-1)+2,m} = \frac{1}{N}\varepsilon'_N(S_{0,T} \otimes R_{4(r-1)+2,m})\varepsilon_N\,,\ R_{4(r-1)+2,m} = P'_N W'_{m,N} W'_{r,N} P_N + P'_N W'_{r,N} W_{m,N} P_N\,,$$

m = 1, …, R,

$$\varphi_{4(r-1)+2,R+m} = \frac{1}{N}\varepsilon'_N(S_{0,T} \otimes R_{4(r-1)+2,R+m})\varepsilon_N\,,\ R_{4(r-1)+2,R+m} = P'_N W'_{m,N} W'_{r,N} W_{m,N} P_N\,,$$

m = 1, …, R,

$$\varphi_{4(r-1)+2,R(m+1)-m(m-1)/2+l-m} = \frac{1}{N}\varepsilon'_N(S_{0,T} \otimes R_{4(r-1)+2,R(m+1)-m(m-1)/2+l-m})\varepsilon_N\,,$$

$R_{4(r-1)+2,R(m+1)-m(m-1)/2+l-m} = P'_N W'_{l,N} W'_{r,N} W_{m,N} P_N + P'_N W'_{m,N} W'_{r,N} W_{l,N} P_N$, m = 1, …, R-1, l = m +1, …, R.

Associated with moment condition M_3, for each r = 1, …R we have:

$$\varphi_{4(r-1)+3} = \frac{1}{N}\varepsilon'_N(S_{1,T} \otimes R_{4(r-1)+3,N})\varepsilon_N\,, R_{4(r-1)+3,N} = P'_N W'_{r,N} W_{m,N} P_N\,,$$

$$\varphi_{4(r-1)+3,m} = \frac{1}{N}\varepsilon'_N(S_{1,T} \otimes R_{4(r-1)+3,m})\varepsilon_N\,, R_{4(r-1)+3,m} = P'_N W'_{r,N} W_{r,N} W_{m,N} P_N\,,\ m = 1, \ldots, R,$$

$$\varphi_{4(r-1)+3,R+m} = \frac{1}{N}\varepsilon'_N(S_{1,T} \otimes R_{4(r-1)+3,R+m})\varepsilon_N\,, R_{4(r-1)+3,R+m} = P'_N W'_{m,N} W'_{r,N} W_{r,N} W_{m,N} P_N\,,$$

m = 1, …, R,

$$\varphi_{4(r-1)+3,R(m+1)-m(m-1)/2+l-m} = \frac{1}{N}\varepsilon'_N(S_{1,T} \otimes R_{4(r-1)+3,R(m+1)-m(m-1)/2+l-m})\varepsilon_N\,,$$

$R_{4(r-1)+3,R(m+1)-m(m-1)/2+l-m} = P'_N W'_{m,N} W'_{r,N} W_{r,N} W_{l,N} P_N$, m = 1, …, R-1, l = m +1, … R.

Associated with moment condition M_4, for each r = 1, …R we have:

$$\gamma_{4(r-1)+4} = \frac{1}{N}\varepsilon'_N(S_{1,T} \otimes R_{4(r-1)+4})\varepsilon_N\,, R_{4(r-1)+4} = P'_N W'_{r,N} P_N\,,\ m = 1, \ldots, R,$$

$$\gamma_{4(r-1)+4,m} = \frac{1}{N}\varepsilon_N'(S_{1,T} \otimes R_{4(r-1)+4,m})\varepsilon_N\,, R_{4(r-1)+4,m} = P_N'W_{m,N}'W_{r,N}'P_N + P_N'W_{r,N}'W_{m,N}P_N\,,$$

m = 1, …, R,

$$\gamma_{4(r-1)+4,R+m} = \frac{1}{N}\varepsilon_N'(S_{1,T} \otimes R_{4(r-1)+4,R+m})\varepsilon_N\,, R_{4(r-1)+4,R+m} = P_N'W_{m,N}'W_{r,N}'W_{m,N}P_N\,,$$

m = 1, …, R,

$$\gamma_{4(r-1)+4,R(m+1)-m(m-1)/2+l-m} = \frac{1}{N}\varepsilon_N'(S_{1,T} \otimes R_{4(r-1)+4,R(m+1)-m(m-1)/2+l-m})\varepsilon_N\,,$$

$R_{4(r-1)+4,R(m+1)-m(m-1)/2+l-m} = P_N'W_{l,N}'W_{r,N}'W_{m,N}P_N + P_N'W_{m,N}'W_{r,N}'W_{l,N}P_N$, m = 1, …, R-1, l = m+1, …, R.

Associated with moment condition M_a:

$$\varphi_{4R+1} = \frac{1}{N}\varepsilon_N'(S_{0,T} \otimes R_{4R+1})\varepsilon_N\,,\ R_{4R+1} = P_N'P_N\,,$$

$$\varphi_{4R+1,m} = \frac{1}{N}\varepsilon_N'(S_{0,T} \otimes R_{4R+1,m})\varepsilon_N\,,\ R_{4R+1,m} = P_N'W_{m,N}'P_N\,,\ m = 1, \dots, R,$$

$$\varphi_{4R+1,R+m} = \frac{1}{N}\varepsilon_N'(S_{0,T} \otimes R_{4R+1,R+m})\varepsilon_N\,,\ R_{4R+1,R+m} = P_N'W_{m,N}'W_{m,N}P_N\,,\ m = 1, \dots, R,$$

$$\varphi_{4R+1,R(m+1)-m(m-1)/2+l-m} = \frac{1}{N}\varepsilon_N'(S_{0,T} \otimes R_{4R+1,R(m+1)-m(m-1)/2+l-m})\varepsilon_N\,,$$

$R_{4R+1,R(m+1)-m(m-1)/2+l-m} = P_N'W_{m,N}'W_{l,N}P_N$, m = 1, …, R-1, and l = m +1, … R.

Associated with moment condition M_b we have:

$$\varphi_{4R+2} = \frac{1}{N}\varepsilon_N'(S_{1,T} \otimes R_{4R+2})\varepsilon_N\,,\ R_{4R+2} = P_N'P_N\,,$$

$$\varphi_{4R+2,m} = \frac{1}{N}\varepsilon_N'(S_{1,T} \otimes R_{4R+2,m})\varepsilon_N\,,\ R_{4R+2,m} = P_N'W_{m,N}'P_N\,,\ m = 1, \dots, R,$$

$$\varphi_{4R+2,R+m} = \frac{1}{N}\varepsilon_N'(S_{1,T} \otimes R_{4R+2,R+m})\varepsilon_N\,,\ R_{4R+2,R+m} = P_N'W_{m,N}'W_{m,N}P_N\,,\ m = 1, \dots, R,$$

$$\varphi_{4R+2,R(m+1)-m(m-1)/2+l-m} = \frac{1}{N}\varepsilon_N'(S_{1,T} \otimes R_{4R+2,R(m+1)-m(m-1)/2+l-m})\varepsilon_N\,,$$

$R_{4R+2,R(m+1)-m(m-1)/2+l-m} = P_N'W_{m,N}'W_{l,N}P_N$, m = 1, …, R-1, l = m +1, … R.

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 $R_{ij,N}$ (i = 1, .., 6) in light of Remark A1. The other elements of G_N^* and g_N^* are 0,1 or of the form $\frac{1}{N}tr(W_{r,N}'W_{r,N})$ 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 A3.

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 $N \to \infty$, provided $\tilde{\beta}_N \xrightarrow{P} \beta$ as $N \to \infty$.

Proof.

In (A.16), elements of $g_N^* = [g_{i,N}^*]$, i = 1, …, (4R + 2), and $G_N^* = [g_{ij,N}^*]$, i = 1, …, (4R + 2) and j = 1, …, [2R + R(R-1)/2 + 2] were shown to be of the form

$$\varphi_{ij,N} = \frac{1}{N} u_N' C_{ij,N} u_N , \tag{A.7}$$

where $C_{ij,N}$ are nonstochastic $NT \times NT$ 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 $C_{ij,N}$ in light of Remark A1. The elements of G_N and g_N defined in (13) are – again apart from a constant – given by

$$\tilde{\varphi}_{ij,N} = \frac{1}{N} \tilde{u}_N' C_{ij,N} \tilde{u}_N . \tag{A.8}$$

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

$$\tilde{u}_N = y_N - X_N \tilde{\beta}_N = u_N - X_N (\dot{\beta}_N - \beta) . \tag{A.9}$$

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

$$\tilde{\varphi}_{ij,N} - \varphi_{ij,N} = (\dot{\beta}_N - \beta)'(N^{-1} X_N' C_{ij,N} X_N)(\dot{\beta}_N - \beta) - 2(\dot{\beta}_N - \beta)'(N^{-1} X_N' C_{ij,N} u_N) . \tag{A.10}$$

Regarding the first term on the right hand side of (A.10), the row and column sums of $C_{ij,N}$ 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 $N^{-1} X_N' C_{ij,N} X_N$ are O(1). Thus, the first term on the right hand side converges in probability to zero since $(\dot{\beta}_N - \beta) \xrightarrow{P} 0$ as $N \to \infty$.

Regarding the second term on the right hand side of (A.10), consider the vector $\zeta_N = N^{-1} X_N' C_{ij,N} u_N$. The mean of ζ_N is zero and its variance covariance matrix is

$$N^{-1} (N^{-1} X_N' C_{ij,N} \Omega_{u,N} C_{ij,N}' X_N) , \tag{A.11}$$

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 $C_{ij,N}\Omega_{u,N}C'_{ij,N}$. Since the elements of X_N are uniformly bounded in absolute value by Assumption A3, it follows that all K^2 elements of $N^{-1}X'_N C_{ij,N}\Omega_{u,N}C'_{ij,N}X_N$ are O(1) in light of remark A1. As a consequence, the variance covariance matrix of ζ_N converges to zero and hence ζ_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

$$G_N - \Gamma_N \xrightarrow{P} 0 \text{ and } g_N - \gamma_N \xrightarrow{P} 0 \text{ as } N \to \infty. \tag{A.12}$$

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 θ 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

$$R_N^0(\underline{\theta}^0) = (g_N^0 - G_N^0\underline{\alpha}^0]'[g_N^0 - G_N^0\underline{\alpha}^0] \text{ and} \tag{A.13a}$$

$$\overline{R}_N^0(\underline{\theta}^0) = [\gamma_N^0 - \Gamma_N^0\underline{\alpha}^0]'[\gamma_N^0 - \Gamma_N^0\underline{\alpha}^0]. \tag{A.13b}$$

Since $\gamma_N^0 - \Gamma_N^0\alpha^0 = 0$, we have $\overline{R}_N^0(\theta^0) = 0$, i.e., $\overline{R}_N^0(\underline{\theta}^0) = 0$ at the true parameter vector $\theta^0 = (\rho_1, ..., \rho_R, \sigma_v^2)$.

Then,

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

$$\overline{R}_N^0(\underline{\theta}^0) - \overline{R}_N^0(\theta^0) = [\underline{\alpha}^0 - \alpha^0]' \Gamma_N^{0\prime} \Gamma_N^0 [\underline{\alpha}^0 - \alpha^0]. \tag{A.14}$$

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

$$\overline{R}_N^0(\underline{\theta}^0) - \overline{R}_N^0(\theta^0) \geq \lambda_{\min}(\Gamma_N^{0\prime} \Gamma_N^0)[\underline{\alpha}^0 - \alpha^0]'[\underline{\alpha}^0 - \alpha^0] \text{ and}$$

$$\overline{R}_N^0(\underline{\theta}^0) - \overline{R}_N^0(\theta^0) \geq \lambda_*[\underline{\alpha}^0 - \alpha^0]'[\underline{\alpha}^0 - \alpha^0] \text{ by Assumption 5.}$$

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

$$\lim_{N\to\infty} \inf_{\{\theta^0 : \|\underline{\theta}^0 - \theta^0\| \geq \varepsilon\}} [\overline{R}_N^0(\underline{\theta}^0) - \overline{R}_N^0(\theta^0)] \geq \inf_{\{\theta^0 : \|\underline{\theta}^0 - \theta^0\| \geq \varepsilon\}} \lambda_* \left\|\underline{\theta}^0 - \theta^0\right\|^2 = \lambda_* \varepsilon^2 > 0 \tag{A.15}$$

which proves that the true parameter θ^0 is identifiable unique.

Next, let $F_N^0 = [g_N^0, -G_N^0]$ and $\Phi_N^0 = [\gamma_N^0, -\Gamma_N^0]$, then the objective function and its nonstochastic counterpart can be written as

$$R_N^0(\underline{\theta}^0) = (1, \underline{\alpha}^{0\prime}) F_N^{0\prime} F_N^0 (1, \underline{\alpha}^{0\prime})' \text{ and}$$

$$\overline{R}_N^0(\underline{\theta}^0) = (1, \underline{\alpha}^{0\prime}) \Phi_N^{0\prime} \Phi_N^0 (1, \underline{\alpha}^{0\prime})'.$$

Hence for $\rho \in [-a, a]$ and $\sigma_v^2 \in [0, b]$ it holds that

$$\left\|R_N^0(\underline{\theta}^0) - \overline{R}_N^0(\underline{\theta}^0)\right\| = \left\|(1, \underline{\alpha}^{0\prime})[F_N^{0\prime} F_N^0 - \Phi_N^{0\prime} \underline{\Phi_N^0}](1, \underline{\alpha}^{0\prime})'\right\|.$$

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

$$\left|R_N^0(\underline{\theta}^0) - \overline{R}_N^0(\underline{\theta}^0)\right| \leq \left\|F_N^{0\prime} F_N^0 - \Phi_N^{0\prime} \Phi_N^0\right\| \left\|(1, \underline{\alpha}^{0\prime})\right\|^2$$

$$\leq \left\|F_N^{0\prime} F_N^0 - \Phi_N^{0\prime} \Phi_N^0\right\| [1 + Ra^2 + \frac{2R + R(R-1)}{2} a^4 + b^2].$$

In light of Theorem A1, we have $\left\| F_N^0 - \Phi_N^0 \right\| \xrightarrow{P} 0$ as $N \to \infty$. Observing that, by Lemma A2, the elements of Φ_N^0 are 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 $N \to \infty$. As a consequence, we have (for finite R)

$$\sup_{\rho \in [-a,a], \sigma_v^2 \in [0,b]} \left| R_N^0(\underline{\theta}^0) - \overline{R}_N^0(\underline{\theta}^0) \right| \leq \left\| [F_N^{0\prime} F_N^0 - \Phi_N^{0\prime} \Phi_N^0] \right\| [1 + Ra^2 + \frac{R(R-1)}{2} a^4 + b^2] \xrightarrow{P} 0 \text{ as } N \to \infty. \quad \text{(A.16)}$$

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

Having proved that the estimators $\tilde{\rho}_{1,N} ..., \tilde{\rho}_{R,N}, \tilde{\sigma}_{v,N}^2$ are consistent for $\rho_{1,} ..., \rho_R, \sigma_v^2$, we now show that σ_1^2 can be estimated consistently from the last line ($4R$ + 2) of equation system (12), using

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

Since $\gamma_N - \Gamma_N \alpha = 0$, we have

$$\begin{aligned} &\tilde{\sigma}_1^2 - \sigma_1^2 = (g_{4R+2} - \gamma_{4R+2}) - (g_{4R+2,1} - \gamma_{4R+2,1})\tilde{\rho}_{1,N} - ... - (g_{4R+2,R} - \gamma_{4R+2,R})\tilde{\rho}_{R,N} \\ &- (g_{4R+2,R+1} - \gamma_{4R+2,R+1})\tilde{\rho}_{1,N}^2 - ... - (g_{4R+2,2R} - \gamma_{4R+2,2R})\tilde{\rho}_{R,N}^2 \\ &- (g_{4R+2,2R+1} - \gamma_{4R+2,2R+1})\tilde{\rho}_{1,N}\tilde{\rho}_{2,N} - ... - (g_{4R+2,2R+R(R-1)/2} - \gamma_{4R+2,2R+R(R-1)/2})\tilde{\rho}_{R-1,N}\tilde{\rho}_{R,N} \\ &- \gamma_{4R+2,1}(\tilde{\rho}_{1,N} - \rho_1) - ... - \gamma_{4R+2,R}(\tilde{\rho}_{R,N} - \rho_R) \\ &- \gamma_{4R+2,R+1}(\tilde{\rho}_{1,N}^2 - \rho_1^2) ... - \gamma_{4R+2,2R}(\tilde{\rho}_{R,N}^2 - \rho_R^2) \\ &- \gamma_{4R+2,2R+1}(\tilde{\rho}_{1,N}\tilde{\rho}_{2,N}^2 - \rho_1\rho_2) ... - \gamma_{4R+2,2R+R(R-1)/2}(\tilde{\rho}_{R-1,N}\tilde{\rho}_{R,N} - \rho_{R-1}\rho_R). \end{aligned} \quad \text{(A.17b)}$$

Observing by Theorem A1 that $F_N - \Phi_N \xrightarrow{P} 0$ as $N \to \infty$ and that the elements of Φ_N are O(1) it follows from the consistency of $\tilde{\rho}_{1,N}, ..., \tilde{\rho}_{R,N}$ that $\tilde{\sigma}_{1,N}^2 - \sigma_1^2 \xrightarrow{P} 0$ as $N \to \infty$.

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

$$R_N(\underline{\theta}) = (g_N - G_N \underline{\alpha}]' \tilde{\Xi}_N^{-1} [g_N - G_N \underline{\alpha}] \text{ and} \quad \text{(A.18a)}$$

$$\overline{R}_N(\underline{\theta}) = [\gamma_N - \varGamma_N \underline{\alpha}]' \varXi_N^{-1} [\gamma_N - \varGamma_N \underline{\alpha}] \tag{A.18b}$$

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

$$\lambda_{\min}(\varGamma'_N \varXi_N^{-1} \varGamma_N) \geq \lambda_0 \text{ for some } \lambda_0 > 0. \tag{A.19}$$

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

Now define $\breve{\varGamma}_N$ as

$$\breve{\varGamma}_N = \begin{bmatrix} \breve{\varGamma}_N^0 \\ \breve{\varGamma}_N^1 \end{bmatrix}, \tag{A.20a}$$

which differs from $\varGamma_N$ only by the ordering of the rows.

$\breve{\varGamma}_N^0$ corresponds to $\varGamma_N^0$ with a zero column appended as last column, i.e., $\breve{\varGamma}_N^0 = [\varGamma_N^0, 0]$, such that

$$\breve{\varGamma}_N^{0\prime} \breve{\varGamma}_N^0 = \begin{bmatrix} \breve{\varGamma}_N^{0\prime} \breve{\varGamma}_N^0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} a_{1,1} & & a_{1,2R+R(R-1)/2+1} & 0 \\ . & & & 0 \\ a_{2R+R(R-1)/2+1,1} & & a_{2R+R(R-1)/2+1,2R+R(R-1)/2+1} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{A.21a}$$

($\breve{\varGamma}_N^{0\prime} \breve{\varGamma}_N^0$ is of the same dimension as $\varGamma'_N \varGamma_N$, i.e., $[2R + R(R\text{-}1)/2 + 2] \times [2R + R(R\text{-}1)/2 + 2]$.)

$\breve{\varGamma}_N^1$ is a modified version of $\varGamma_N^1$, with a zero column included as second last column, such that

$$\breve{\varGamma}_N^{1\prime} \breve{\varGamma}_N^1 = \begin{bmatrix} b_{1,1} & & 0 & b_{1,2R+R(R-1)/2+1} \\ . & & 0 & . \\ 0 & 0 & 0 & 0 \\ b_{2R+R(R-1)/2+1,1} & . & 0 & b_{2R+R(R-1)/2+1,2R+R(R-1)/2+1} \end{bmatrix}. \tag{A.21b}$$

($\breve{\varGamma}_N^{1\prime} \breve{\varGamma}_N^1$ is of the same dimension as $\varGamma'_N \varGamma_N$, i.e., $[2R + R(R\text{-}1)/2 + 2] \times [2R + R(R\text{-}1)/2 + 2]$.)

Since $\breve{\Gamma}_N = \begin{bmatrix} \breve{\Gamma}_N^0 \\ \breve{\Gamma}_N^1 \end{bmatrix}$ differs from Γ_N only by the ordering of the rows, it follows that

$$\Gamma_N' \, \Gamma_N = \breve{\Gamma}_N' \, \breve{\Gamma}_N = \begin{bmatrix} \breve{\Gamma}_N^{0\,\prime} & \breve{\Gamma}_N^{1\,\prime} \end{bmatrix} \begin{bmatrix} \breve{\Gamma}_N^0 \\ \breve{\Gamma}_N^1 \end{bmatrix} = \breve{\Gamma}_N^{0\,\prime} \breve{\Gamma}_N^0 + \breve{\Gamma}_N^{1\,\prime} \breve{\Gamma}_N^1 \text{, i.e.,} \tag{A.22}$$

$$\Gamma_N' \, \Gamma_N = \begin{bmatrix} a_{1,1} & & a_{1,2R+R(R-1)/2+1} & 0 \\ . & & & 0 \\ a_{2R+R(R-1)/2+1,1} & & a_{2R+R(R-1)/2+1,2R+R(R-1)/2+1} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

$$+ \begin{bmatrix} b_{1,1} & & 0 & b_{1,2R+R(R-1)/2+1} \\ . & & 0 & . \\ 0 & 0 & 0 & 0 \\ b_{2R+R(R-1)/2+1,1} & . & 0 & b_{2R+R(R-1)/2+1,2R+R(R-1)/2+1} \end{bmatrix}.$$

Utilizing Assumption A5 we have

$$x'\Gamma_N{}'\Gamma_N x = x'\breve{\Gamma}_N^{0\,\prime}\breve{\Gamma}_N^0 x + x'\breve{\Gamma}_N^{1\,\prime}\breve{\Gamma}_N^1 x = x_A' A x_A + x_B' B x_B \,. \tag{A.23}$$

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

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

$$x_A' A x_A + x_B' B x_B \geq \lambda_{\min}(A) x_A' x_A + \lambda_{\min}(B) x_B' x_B \geq \lambda^* (x_A' x_A + x_B' x_B) \geq \lambda^* x' x \tag{A.24}$$

for any $x = [x_1, x_2, ..., x_{2R+2}]$.

Hence, we have shown that

$$x'\Gamma_N'\Gamma_N x \geq \lambda^* x'x \,,$$

or, equivalently,

$$\frac{x'\Gamma_N'\Gamma_N x}{x'x} \geq \lambda^* \quad \text{for } x \neq 0 \,. \tag{A.25}$$

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

$$\lambda_{\min}(\Gamma_N'\Gamma_N) = \inf_x \frac{x'\Gamma_N'\Gamma_N x}{x'x} \geq \lambda_* > 0 . \qquad \text{(A.26)}$$

Using Mittelhammer (1996, p. 254) we have

$$\lambda_{\min}(\Gamma_N'\Xi_N^{-1}\Gamma_N) = \inf_x \frac{x'\Gamma_N'\Xi_N^{-1}\Gamma_N x}{x'x} \geq \lambda_{\min}(\Xi_N^{-1}) \inf_x \frac{x'\Gamma_N'\Gamma_N x}{x'x}$$

$$= \lambda_{\min}(\Xi_N^{-1})\lambda_{\min}(\Gamma_N'\Gamma_N) \geq \lambda_0 > 0 , \qquad \text{(A.27)}$$

with $\lambda_0 = \overline{\lambda}_*\lambda_*$ since $= \lambda_{\min}(\Xi_N^{-1}) \geq \overline{\lambda}_* > 0$ by assumption (see Theorem 2).

This ensures that the true parameter vector $\theta = (\rho_1,...,\rho_R,\sigma_v^2,\sigma_1^2)$ is identifiable unique.

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

Analogous to the prove of theorem 1, observe that $\overline{R}_N(\theta) = 0$, i.e., $\overline{R}_N(\underline{\theta}) = 0$ at the true parameter vector $\theta = (\rho_1,...,\rho_R,\sigma_v^2,\sigma_1^2)$. It follows that

$$\overline{R}_N(\underline{\theta}) - \overline{R}_N(\theta) = [\underline{\alpha} - \alpha]'\Gamma_N'\,\Xi_N^{-1}\Gamma_N[\underline{\alpha} - \alpha] . \qquad \text{(A.30)}$$

Moreover, let $F_N = [g_N, -G_N]$ and $\Phi_N = [\gamma_N, -\Gamma_N]$, then,

$$R_N(\underline{\theta}) = (1,\underline{\alpha}')F_N'\tilde{\Xi}_N^{-1}F_N(1,\underline{\alpha}')' \text{ and} \qquad \text{(A.31a)}$$

$$\overline{R}_N(\underline{\theta}) = (1,\underline{\alpha}')\Phi_N'\Xi_N^{-1}\Phi_N(1,\underline{\alpha}')' . \qquad \text{(A.31b)}$$

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

Proof of Theorem 3. Consistency of partially weighted GM estimator

Let $\overline{\lambda}_* = \min[(T-1)\sigma_v^{-4},\sigma_1^{-4}]$ and $\overline{\lambda}_{**} = \max[(T-1)\sigma_v^{-4},\sigma_1^{-4}]$.

Then $0 < \overline{\lambda}_* \leq \lambda_{\min}(\Xi_N^p)^{-1} \leq \overline{\lambda}_{**} < \infty$. The proof of Theorem 3 is now analogous to that of Theorem 2 with Ξ_N and $\tilde{\Xi}_N$ replaced by Ξ_N^p and $\tilde{\Xi}_N^p$.

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