Spatial Dynamic Panel Model and System GMM: A Monte Carlo Investigation

Kukenova, Madina and Monteiro, Jose-Antonio

University of Lausanne, University of Neuchatel

July 2008

Spatial Dynamic Panel Model and System GMM: A Monte Carlo Investigation*

Madina Kukenova† José-Antonio Monteiro‡

Draft: February 2009

Abstract

This paper investigates the finite sample properties of estimators for spatial dynamic panel models in the presence of several endogenous variables. So far, none of the available estimators in spatial econometrics allows considering spatial dynamic models with one or more endogenous variables. We propose to apply system-GMM, since it can correct for the endogeneity of the dependent variable, the spatial lag as well as other potentially endogenous variables using internal and/or external instruments. The Monte-Carlo investigation compares the performance of spatial MLE, spatial dynamic MLE (Elhorst (2005)), spatial dynamic QMLE (Yu et al. (2008)), LSDV, difference-GMM (Arellano & Bond (1991)), as well as extended-GMM (Arellano & Bover (1995), Blundell & Bover (1998)) in terms of bias, root mean squared error and standard-error accuracy. The results suggest that, in order to account for the endogeneity of several covariates, spatial dynamic panel models should be estimated using extended GMM. On a practical ground, this is also important, because system-GMM avoids the inversion of high dimension spatial weights matrices, which can be computationally unfeasible for large N and/or T.

Keywords: Spatial Econometrics, Dynamic Panel Model, System GMM, Monte Carlo Simulations

JEL classification: C15, C31, C33

*We thank Florian Pelgrin for suggesting the main idea of this paper.

†University of Lausanne, madina.kukenova@unil.ch

‡University of Neuchatel, jose-antonio.monteiro@unine.ch

1 Introduction

Although the econometric analysis of dynamic panel models (Arellano and Bond (1998), Blundell and Bover (1998), Baltagi and Kao (2000)) has drawn a lot of attention in the last decade, econometric analysis of spatial and dynamic panel models is almost inexistent (Elhorst (2005), Kapoor, Kelejian and Prucha (2007), Lee and Yu (2007), Yu et al. (2007) and Beenstock and Felsenstein (2007)). So far, none of the available estimators allows to consider a dynamic spatial lag panel model with one or more endogenous variables (besides the time and spatial lag) as explanatory variables. From an applied econometric point of view, this is an important issue because several reasons can explain the presence of endogeneity (measurement errors, variables omission, simultaneous relationship between the dependent and the explanatory variable). Empirically, there are numerous examples where the presence of a dynamic process, spatial dependence and endogeneity might occur.

This is the case with the analysis of the determinants of Foreign Direct Investment (FDI). In particular, complex FDI is characterized by a multinational firm from home country i which owns not only a production plant in host country j but also one in third country k, in order to exploit the comparative advantages of various locations (Baltagi, Egger and Pfaffermayr (2007)). This type of FDI can thus feature complementary/substitutive spatial dependence with respect to FDI to other host countries. The presence of complex FDI can be tested empirically by estimating a spatial lag model (as proposed by Blonigen, Davies, Waddell and Naughton (2007)), which can also include a lagged dependent variable to account for the fact that FDI decisions are part of a dynamic process, i.e. more FDI in a host country seems to attract more FDI in this same host country (Kukenova and Monteiro (2008)). This persistence effect is partly due to the fact that FDI is often accompanied by physical investments that are irreversible in the short run. Since the inclusion of the time lagged depend variable in the equation might lead to inconsistent estimates, dynamic spatial lag panel models are usually estimated using the system generalized method of moments (GMM) estimator, developed by Arellano and Bover (1995) and Blundell and Bond (1998). The main argument of applying the extended GMM in a spatial context is that it corrects for the endogeneity of the spatial lagged dependent variable and other potentially endogenous explanatory variables. Going beyond this intuitive motivation, this paper wants to determine if it is suitable to instrument the spatial lag variable using the instruments proposed by system GMM, i.e. lagged spatial lag values. This is done by comparing the results obtained by extended GMM with spatial dynamic estimators (Spatial MLE (SMLE), Spatial Dynamic MLE (SDMLE) and Spatial Dynamic QMLE (SDQMLE)) which assume only exogenous covariates.

The outline of the paper is as follows. The dynamic spatial lag model is defined and interpreted in section 2. The Monte Carlo investigation is described and performed in section 3. Finally, section 4 concludes.

2 Spatial Dynamic Panel Model

The development of empirical spatial models is intimately linked to the recent progress in spatial econometrics. The basic spatial model was suggested by Cliff and Ord (1981), but it did not receive important theoretical extensions until the middle of the 1990s. Anselin (2001) and Elhorst (2003b) provide thorough surveys of the different spatial models and suggest econometric strategies to estimate them. More generally, spatial data is characterized by the spatial arrangement of the observations. Following Tobler's First Law of Geography, *everything is related to everything else, but near things are more related than distant things*, the spatial linkages of the observations $i = 1, ..., N$ are measured by defining a spatial weight matrix, denoted by W_t for any year $t = 1, ..., T$:

$$W_t = \begin{pmatrix} 0 & w_t(d_{k,j}) & \cdots & w_t(d_{k,l}) \\ w_t(d_{j,k}) & 0 & \cdots & w_t(d_{j,l}) \\ \vdots & \vdots & \ddots & \vdots \\ w_t(d_{l,k}) & w_t(d_{l,j}) & \cdots & 0 \end{pmatrix}$$

where $w_t(d_{j,k})$ defines the functional form of the weights between any two pair of location j and k. In the construction of the weights themselves, the theoretical foundation for $w_t(d_{j,k})$ is quite general and the particular functional form of any single element in W_t is, therefore, not prescribed. In fact, the determination of the proper specification of W_t is one of the most difficult and controversial methodological issues in spatial data analysis. As is standard in spatial econometrics, for ease of interpretation, the weighting matrix W_t is row standardized so that each row in W_t sums to one.

As distances are time-invariant, it will generally be the case that $W_t = W_{t+1}$. However, when dealing with unbalanced panel data, this is no longer true (Egger et al (2005)). Stacking the data first by time and then by cross-section, the full weighting matrix, W, is given by:

$$W = \begin{pmatrix} W_1 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \ddots & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & W_T \end{pmatrix}$$

2.1 Dynamic Spatial Lag Model

A general spatial dynamic panel model can be described as follows:

$$\begin{aligned} Y_t &= \alpha Y_{t-1} + \rho W_{1t} Y_t + EX_t \beta + EN_t \gamma + \varepsilon_t \\ \varepsilon_t &= \eta + \phi W_{2t} \varepsilon_t + v_t, \quad t = 1, ..., T \end{aligned} \tag{1}$$

where Y_t is a $N \times 1$ vector, W_{1t} and W_{2t} are $N \times N$ spatial weight matrices which are non-stochastic and exogenous to the model, η is the vector of country effect, EX_t is a $N \times p$ matrix of p exogenous explanatory variables ($p \geq 0$) and EN_t is a $N \times q$ matrix of q endogenous explanatory variables with respect to Y_t ($q \geq 0$). Finally, v_t is assumed to be normally distributed ($N(\mathbf{0}, \Omega)$). By adding some restrictions to the parameters, two popular spatial model specifications can be derived from this general spatial model, namely the dynamic spatial lag model ($\phi = 0$) and the dynamic spatial error model ($\rho = 0$)[1].

The spatial lag model accounts directly for relationships between dependent variables that are believed to be related in some spatial way. Somewhat analogous to a lagged dependent variable in time series analysis, the estimated "spatial lag" coefficient[2] characterizes the contemporaneous correlation between one cross-section and other geographically-proximate cross-sections. The following equation gives the basic

[1] The analysis of the spatial error panel model is beyond the scope of this paper. For further details, see Elhorst (2005) and Kapoor et al. (2007).

[2] The spatial autoregressive term is also referred as endogenous interation effects in social economics or as interdependence process in political science.

spatial dynamic panel specification, also known as the "time-space simultaneous" model (Anselin (1988, 2001))[3]:

$$Y_t = \alpha Y_{t-1} + \rho W_t Y_t + EX_t \beta + EN_t \gamma + \eta + v_t \tag{2}$$

The spatial autoregressive coefficient (ρ) associated with $W_t Y_t$ represents the effect of the weighted average ($w_t(d_{ij})$ being the weights) of the neighborhood, i.e. $[W_t Y_t]_i = \sum_{j=1..N_t} w_t(d_{ij}) \cdot Y_{jt}$. The spatial lag term allows to determine if the dependent variable Y_t is (positively/negatively) affected by the Y_t from other close locations weighted by a given criterion (usually distance or contiguity). In other words, the spatial lag coefficient captures the impact of Y_t from neighborhood locations. Let $\omega_{\min}$ and $\omega_{\max}$ be the smallest and highest characteristic root of the spatial matrix W, then this spatial effect is assumed to lie between $\frac{1}{\omega_{\min}}$ and $\frac{1}{\omega_{\max}}$. Most of the spatial econometrics literature constrains the spatial lag to lie between -1 and +1. However, this might be restrictive, because if the spatial matrix is row-normalized, then the highest characteristic root is equal to unity ($\omega_{\max} = 1$), but the smallest eigenvalue can be bigger than -1, which would lead the lower bound to be smaller than -1.

Given that expression (2) is a combination of a time and spatial autoregressive models, we need to ensure that the resulting process is stationary. The stationarity restrictions in this model are stronger than the individual restrictions imposed on the coefficients of a spatial or dynamic model[4]. The process is covariance stationary if $\left|(I_N - \rho W_t)^{-1} \alpha\right| < 1$, or, equivalently, if

$$\begin{aligned} |\alpha| &< 1 - \rho\omega_{\max} \quad \text{if } \rho \geq 0 \\ |\alpha| &< 1 - \rho\omega_{\min} \quad \text{if } \rho < 0 \end{aligned}$$

[3] Beside the "time-space simulatenous" model, Anselin distinguishes three other distinct spatial lag panel models: the "pure space recursive" model which only includes a lagged spatial lag coefficient; the "time-space recursive" specification which considers a lagged dependent variable as well as a lagged spatial lag (see Korniotis (2007)); and the "time-space dynamic" model, which includes a time lag, a spatial lag and a lagged spatial lag.

[4] In time series, the autoregressive coefficient must satisfy: $\alpha \in (-1; 1)$, while in the spatial econometrics the spatial lag coefficient must lie $\rho \in \left(\frac{1}{\omega_{\min}}; \frac{1}{\omega_{\max}}\right)$.

From an econometric viewpoint, equation (2) faces simultaneity and endogeneity problems, which in turn means that OLS estimation will be biased and inconsistent (Anselin (1988)). To see this point more formally, note that the reduced form of equation (2) takes the following form:

$$Y_t = (I_N - \rho W_t)^{-1} (\alpha Y_{t-1} + EX_t\beta + EN_t\gamma + \eta + v_t)$$

Each element of Y_t is a linear combination of all of the error terms. Moreover, as pointed out by Anselin (2003), assuming $|\rho| < 1$ and each element of W_t is smaller than one imply that $(I_N - \rho W_t)^{-1}$ can be reformulated as a Leontief expansion $(I_N - \rho W_t)^{-1} = I + \rho W_t + \rho^2 W_t^2 + ...$ Accordingly, the spatial lag model features two types of global spillovers effects: a multiplier effect for the predictor variables as well as a diffusion effect for the error process. Since the spatial lag term $W_t Y_t$ is correlated with the disturbances, even if v_t are independently and identically distributed, it must be treated as an endogenous variable and proper estimation method must account for this endogeneity.

Despite the fact that dynamic panel models have been the object of recent important developments (see survey by Baltagi and Kao (2000) or Phillips and Moon (2000)), econometric analysis of spatial dynamic panel models is almost inexistent. In fact, there is only a limited number of available estimators that deal with spatial and time dependence in a panel setting. Table 1 sums up the different estimators proposed in the literature:

In the absence of spatial dependence, there are three types of estimators available to estimate a dynamic panel model. The first type of estimators consists of estimating an unconditional likelihood function (Hsiao et al. (2002)). The second type of procedure corrects the bias associated with the least square dummy variables (LSDV) estimator (Bun and Carree (2005)). The last type, which is the most popular, relies on GMM estimators, like difference GMM (Arellano and Bond (1992)) or system GMM (Arellano and Bover (1995), Blundell and Bond (1998)).

Model	Estimation Methods	Endogenous Variable(s)
$Y_t = \alpha Y_{t-1} + \beta EX_t + \epsilon_t$	Difference GMM (Arellano & Bond (1991)) System-GMM (Arellano & Bover (1995), Blundell & Bond (1998)) MLE/MDE (Hsiao, Pesaran & Tahmiscioglu (2002)) CLSDV (Kiviet (1995), Hahn & Kuersteiner (2002) Bun & Carree (2005))	Y_{t-1};
$Y_t = \alpha Y_{t-1} + \beta EX_t + \gamma EN_t + \epsilon_t$	System-GMM (Arellano & Bover (1995), Blundell & Bond (1998))	$Y_{t-1}; EN_t$
$Y_t = \alpha Y_{t-1} + \rho W Y_{t-1} + \beta EX_t + \gamma EN_t + \epsilon_t$	LSDV-IV (Korniotis (2008))	$Y_{t-1}; EN_t$
$Y_t = \rho W Y_t + \beta EX_t + \epsilon_t$	Spatial-MLE (Anselin (1988) (2001), Elhorst (2003)) Spatial 2SLS (Anselin (1988) (2001)))	WY_t
$Y_t = \rho W Y_t + \beta EX_t + \gamma EN_t + \epsilon_t$	Spatial 2SLS (Dall'erba & Le Gallo (2007))	$WY_t; EN_t$
$Y_t = \alpha Y_{t-1} + \rho W Y_t + \beta EX_t + \epsilon_t$	Spatial Dynamic MLE (Elhorst (2003b, 2005, 2008)) Spatial Dynamic QMLE (Yu, de Jong & Lee (2007) (2008), Lee & Yu (2007)) C2SLSDV (Beenstock & Felsenstein (2007)) Spatial MLE-GMM / Spatial MLE-Spatial Dynamic MLE (Elhorst (2008))	$WY_t; Y_{t-1}$
$Y_t = \alpha Y_{t-1} + \rho W Y_t + \beta EX_t + \gamma EN_t + \epsilon_t$	System-GMM (Arellano & Bover (1995), Blundell & Bond (1998))	$WY_t; Y_{t-1}; EN_t$

Assuming all explanatory variables are exogenous beside the spatial autoregressive term, the spatial lag panel model without any time dynamic is usually estimated using spatial maximum likelihood (spatial ML) (Elhorst (2003b)) or spatial two-stage least squares methods (S2SLS) (Anselin (1988) (2001)). The ML approach consists of estimating the spatial coefficient by maximizing the non-linear reduced form of the spatial lag model. The spatial 2SLS uses the exogenous variables and their spatially weighted averages (EX_t, $W_t \cdot EX_t$) as instruments[5]. When the number of cross-sections is larger than the period sample, Anselin (1988) suggests to estimate the model using MLE, 2SLS or 3SLS in a spatial seemingly unrelated regression (SUR) framework. More recently, Dall'erba and Le Gallo (2007) suggest to estimate a spatial lag panel model, which includes several endogenous variables but no time dynamic, by applying spatial 2SLS with lower orders of the spatially weighted sum of the exogenous variables as instrument for the spatial autoregressive term[6].

In a dynamic context, the estimation of spatial lag panel models is usually based on a ML function. Elhorst (2003a, 2005) proposes to estimate the unconditional loglikelihood function of the reduced form of the model in first-difference. While the absence of explanatory variables besides the time and spatial lags leads to an exact likelihood function, this is no longer the case when additional regressors are included. Moreover, when the sample size T is relatively small the initial observations contribute greatly to the overall likelihood. That is why the pre-sample values of the explanatory variables and likelihood function are approximated using the Bhargava and Sargan approximation or the Nerlove and Balestra approximation. More recently, Yu et al. (2008) provide a theoretical analysis on the asymptotic properties of the quasi-maximum likelihood (Spatial Dynamic QML), which relies on the maximization of the concentrated likelihood function of the demeaned model. They show that the limit distribution is not centered around zero and propose a bias-corrected estimator[7]. The main difference

[5] In a cross-section setting, Kelejian and Prucha (1998) propose also additional instruments ($W_t^2 EX_t$, $W_t^3 EX_t$, ...). Lee (2003) shows that the estimator proposed by Kelejian and Prucha is not an asymptotically optimal estimator and suggests a three-steps procedure with an alternative instrument for the spatial autoregressive coefficient in the last step ($W_t \cdot \left(I_N - \widetilde{\widetilde{\rho}} W_t\right)^{-1} \cdot EX_t \widetilde{\widetilde{\beta}}$, where $\widetilde{\widetilde{\rho}}$ and $\widetilde{\widetilde{\beta}}$ are estimates obtained using the S2SLS proposed by Kelejian and Prucha (1998)).

[6] Recently, Fingleton and Le Gallo (2008) propose an extended feasible generalized spatial two-stage least squares estimator for spatial lag models with several endogenous variables and spatial error term in a cross-section framework.

[7] In two other related working papers, Lee and Yu (2007) and Yu et al.(2007) investigate the presence of non-stationarity and time fixed effects, respectively.

between Elhorst's and Yu et al.'s ML estimators lies in the asymptotic structure. Elhorst considers fixed T and large N $(N \to \infty)$, while Yu et al. assume large N and T $(N \to \infty;\ T \to \infty)$. Consequently, the way the individual effects are taken out differs: Elhorst considers first-difference variables, while Yu et al. demean the variables. Assuming large T avoids the problem associated with initial values and the use of approximation procedures. Finally Yu et al's approach allows to recover the estimated individual effects, which is not the case with the estimator proposed by Elhorst. In his most recent work, Elhorst (2008) analyzes the finite sample performance of several estimators for a spatial dynamic panel model with only exogenous variables. The estimators considered are the Spatial MLE, Spatial Dynamic MLE and GMM. His Monte Carlo study shows that Spatial Dynamic MLE has the better overall performance in terms of bias reduction and root mean squared errors (RMSE), although the Spatial MLE presents the smallest bias for the spatial autoregressive coefficient. Based on these results, Elhorst proposes two mixed estimators, where the spatial lag dependent variable is based on the spatial ML estimator and the remaining parameters are estimated using either GMM or Spatial Dynamic ML conditional on the spatial ML's estimate of the spatial autoregressive coefficient. These two mixed estimators outperform the original estimators. The mixed Spatial MLE/Spatial Dynamic MLE estimator shows superior performance in terms of bias reduction and RMSE in comparison with mixed Spatial MLE/GMM. However, the latter can be justified on a practical ground if the number of cross-sections in the panel is large, since the time needed to compute Spatial MLE/Spatial Dynamic MLE is substantial. In a spatial vector autoregression (VAR) setting, Beenstock and Felsenstein (2007) suggest a two-step procedure. The first step consists of applying LSDV to the model without the spatial lag and computing the fitted values $(\widehat{Y}_t)$. Then, in the second step, the full model is also estimated using LSDV, but with $W_t\widehat{Y}_t$ as instrument for W_tY_t. Finally, the authors suggest to correct the bias of the lagged dependent variable by using the asymptotic bias defined by Hsiao (1986).

If one is willing to consider some explanatory variables as potentially endogenous in a dynamic spatial panel setting, then no estimator is currently available. From an applied econometric point of view, this is an important issue because several grounds can lead to the presence of endogeneity including measurement errors, variables omission or the presence of simultaneous relationship(s) between the dependent and the explanatory variable(s). The main drawback of applying SMLE, SDMLE or SDQMLE is that, while the spatial autoregressive coefficient is considered endogenous, no instrumental treatment is applied to other potential endogenous variables. This can lead to biased estimates, which would invalidate empirical results.

2.2 System GMM

Empirical papers dealing with a dynamic spatial panel model with several endogenous variables usually apply system-GMM[8]. Haining (1978) already proposed to instrument a first order spatial autoregressive model using lagged dependent variables. While this method is not efficient in a cross-section setting, because it does not use efficiently all the available information (Anselin (1988)), this is no longer necessarily the case in a panel framework. The bias-corrected LSDV-IV estimator proposed by Korniotis (2007) is in line with this approach and considers lagged spatial lag and dependent variable as instruments. Accordingly, the use of system GMM might be justified in this trade-off situation, since the spatial lag would be instrumented by lagged values of the dependent variable and the spatial autoregressive variable.[9]. In particular, extended GMM can correct for the endogeneity of the spatial lag and lagged dependent variable as well as other potentially endogenous explanatory variables. It also allows to take into consideration some econometrics problems such as measurement error and weak instruments. Moreover it also controls for time-invariant individual-specific effects such as distance, culture and political structure. On a practical ground, it also avoids the inversion of high dimension spatial weights matrix W and the computation of its eigenvalues[10], which can be sometimes computationally unfeasible to estimate model with large N and/or T.

For simplicity, equation (2) is reformulated for a given cross-section i ($i = 1, .., N$) at time t ($t = 1, .., T$):

$$Y_{it} = \alpha Y_{it-1} + \rho \left[W_t Y_t\right]_i + EX_{it}\beta + EN_{it}\gamma + \eta_i + v_{it} \tag{3}$$

[8] See for example, Madriaga and Poncet (2007), Foucault, Madies and Paty (2008), or Hong, Sun and Li (2008).

[9] Badinger et al. (2004) recommend to apply system GMM, once the data has been spatially filtered. This approach can only be consider when spatial depence is viewed as a nuisance parameter.

[10] Kelejian and Prucha (1999) notice that the calculation of roots for moderate 400×400 nonsymmetric matrix involves accuracy problems.

According to the GMM procedure, one has to get rid of the individual effects (η_i) correlated with the covariates and the lagged dependent variable, by rewriting equation (3) in first order difference for individual i at time t:

$$\triangle Y_{it} = \alpha \triangle Y_{it-1} + \rho \triangle \left[W_t Y_t\right]_i + \triangle EX_{it}\beta + \triangle EN_{it}\gamma + \triangle v_{it} \tag{4}$$

Even if the fixed effects (within) estimator cancels the country individual fixed (η_i), the lagged endogenous variable ($\triangle Y_{it-1}$) is still correlated with the idiosyncratic error terms (v_{it}). Nickell (1981) as well as Anderson and Hsiao (1981) showed that the within estimator has a bias measured by $\mathrm{O}(\frac{1}{T})$ and is only consistent for large T. Given that this condition is usually not satisfied, the GMM estimator is also biased and inconsistent. Arellano and Bond (1991) propose the following moment conditions associated with equation (4):

$$E\left(Y_{i,t-\tau} \triangle v_{it}\right) = 0; \quad \text{for } t = 3, ..., T \ \text{ and } \ 2 \leq \tau \leq t-1 \tag{5}$$

But the estimation based only on these moment conditions (5) is insufficient, if the strict exogeneity assumption of the covariates (EX_{it}) has not been verified. The explicative variables constitute valid instruments to improve the estimator's efficiency, only when the strict exogeneity assumption is satisfied:

$$E\left(EX_{i\tau} \triangle v_{it}\right) = 0; \quad \text{for } t = 3, ..., T \ \text{ and } \ 1 \leq \tau \leq T \tag{6}$$

However, the GMM estimator based on the moment conditions (5) and (6) can still be inconsistent when $\tau < 2$ and in presence of inverse causality, i.e. $\mathrm{E}(EX_{it}v_{it}) \neq 0$. In order to overcome this problem, one can assume that the covariates are weakly exogenous for $\tau < t$, which means that the moment condition (6) can be rewritten as:

$$E\left(EX_{i,t-\tau} \triangle v_{it}\right) = 0; \quad \text{for } t = 3, ..., T \ \text{ and } \ 1 \leq \tau \leq t-1 \tag{7}$$

For the different endogenous variables, the valid moment conditions are

$$E\left(EN_{i,t-\tau}\triangle v_{it}\right) = 0; \quad \text{for } t = 3...T \text{ and } 2 \leq \tau \leq t-1 \tag{8}$$

$$E\left(\left[W_{t-\tau}Y_{t-\tau}\right]_i \triangle v_{it}\right) = 0; \quad \text{for } t = 3...T \text{ and } 2 \leq \tau \leq t-1 \tag{9}$$

For small samples, this estimator can still yield biased coefficients. Blundell and Bond (1998) showed that the imprecision of this estimator is bigger as the individual effects are important and as the variables are persistent over time. To overcome this limits, the authors propose the system GMM, which estimate simultaneously equation (3) and equation (4). The extra moment conditions for the extended GMM are thus:

$$E\left(\triangle Y_{i,t-1}v_{it}\right) = 0; \quad \text{for } t = 3,...,T \tag{10}$$

$$E\left(\triangle EX_{it}v_{it}\right) = 0; \quad \text{for } t = 2,...,T \tag{11}$$

$$E\left(\triangle EN_{it-1}v_{it}\right) = 0; \quad \text{for } t = 3,...,T \tag{12}$$

$$E\left(\triangle \left[W_{t-1}Y_{t-1}\right]_i v_{it}\right) = 0; \quad \text{for } t = 3,...,T \tag{13}$$

The consistency of the SYS-GMM estimator relies on the validity of these moment conditions, which depends on the assumption of absence of serially correlation of the level residuals and the exogeneity of the explanatory variables. Therefore, it is necessary to apply specification tests to ensure that these assumptions are justified. More generally, one should keep in mind that the estimation of the spatial autoregressive coefficient although "potentially" consistent is usually not the most efficient one. Efficiency relies on the "proper" choice of instruments, which is not an easy task to determine. Arellano and Bond suggest two specification tests in order to verify the consistency of the GMM estimator. First, the overall validity of the moment conditions is checked by the Sargan/Hansen test. The null hypothesis is that instruments are not correlated with the residuals. Aware that too many instrument variables tend to validated invalid results through the Hansen J test for joint validity of those instruments, as well as the difference-in-Sargan/Hansen tests for subsets of instruments, it is advised to restrict the number of instruments by defining a maximum number of lags or by collapsing the instruments (see Roodman (2006)). Second, the Arellano-Bond test examines the serial correlation property of the level residuals. In order to check first-correlation in levels, we rely on the Arellano bond test for second order autocorrelation (m_2).

3 A Monte-Carlo Study

In this section, we investigate the finite sample properties of several estimators including Spatial MLE, Spatial Dynamic MLE and Spatial Dynamic QMLE, LSDV, difference GMM and extended GMM to account for the endogeneity of the spatial lag as well as an additional regressor in a dynamic panel data context using Monte-Carlo simulations[11]. Simulation studies already showed that bias associated with the spatial lag is rather small (Franzese and Hays (2007), Elhorst (2008)), but none analyze the consequences of an additional endogenous explanatory variable in a spatial dynamic context. The data generating process (DGP) is defined as follows:

$$\begin{aligned} Y_{it} &= \alpha Y_{i,t-1} + \rho \left[WY_t\right]_i + \beta EX_{it} + \gamma EN_{it} + \eta_i + v_{it} && (14) \\ EX_{it} &= \delta EX_{i,t-1} + u_{it} && (15) \\ EN_{it} &= \lambda EN_{i,t-1} + \psi \eta_i + \theta v_{it} + e_{it} && (16) \end{aligned}$$

with $\eta_i \sim N\left(0, \sigma_\eta^2\right)$; $v_{it} \sim N\left(0, \sigma_v^2\right)$; $u_{it} \sim N\left(0, \sigma_u^2\right)$; $e_{it} \sim N\left(0, \sigma_e^2\right)$.

In order to avoid results being influenced by initial observations, the covariates Y_{i0}, EX_{i0} and $EN_{i0}$are set to 0 for all i and each variable is generated $(100 + T)$ times according to their respective DGP. The first 100 observations are then discarded. Note that the the dependent variable is generated accodring to the reduced form of equation (14):

$$Y_{it} = \left(1 - \rho\left[W\right]_i\right)^{-1}\left[\alpha Y_{i,t-1} + \beta EX_{it} + \gamma EN_{it} + \eta_i + v_{it}\right]$$

In order to check the consistency of the spatial autoregressive estimator, we consider the following different designs with different sample and cross-country sizes:

$$T \in \{10, 20, 30, 40\}\,;$$

$$N \in \{20;\ 30;\ 50;\ 70\}\,;$$

$$\beta = 1;\ \delta = 0.65;\ \gamma = 0.5;\ \lambda = 0.45;\ \psi = 0.25; \theta = 0.6;$$

$$\sigma_u^2 = 0.05;\ \sigma_v^2 = 0.05;\ \sigma_e^2 = 0.05;\ \sigma_\eta^2 = 0.05$$

[11]Simulations are performed using Matlab R2008b.

The parameter α and ρ are randomly generated within the interval $(0.1; 0.9)$ In order to ensure stationarity, only design which respect the restrictions $|\alpha| < 1 - \rho\omega_{\max}$ if $\rho \geq 0$ or $|\alpha| < 1 - \rho\omega_{\min}$ if $\rho < 0$ are considered. For each of these designs, we performed 1000 trials. Note that for each design, the initial conditions and spatial weight matrices are generated once.

Following Kapoor et al. (2007) and Kelejian and Prucha (1999), we consider different types of spatial weight matrix. In each case, the matrices are row-standardized so that all non zero elements in each row sum to one. The first three matrices rely on a perfect "idealized" circular world, while the last ones consider a real-word weighting scheme. The three "theoretical" spatial matrices, referred as "1 ahead and 1 behind", "3 ahead and 3 behind" and "5 ahead and 5 behind", respectively, are characterized by different degree of sparseness. Each are such that each location is related to the one/three/five locations immediately before and after it, so that each nonzero elements are equal to $0.5/0.\overline{3}/0.1$, respectively. The last two spatial weighting schemes are based on real distance data. We consider the distance between capitals among OECD countries and among non OECD countries[12], respectively. In order to avoid giving some positive weight to very remote countries (with weaker cultural, political and economic ties), we consider the negative exponential weighting scheme. This is done by dividing the distance between locations j and k by the minimum distance within the region r (where location j lies within region r): $w(d_{j,k}) = \exp(-d_{j,k}/MIN_{r,j})$ if $j \neq k$.

As a measure of consistency, we consider the root mean square error (RMSE). Theoretically, RMSE is defined as the square root of the weighted average of the mean and the variance. We not only consider this definition but also the approximation given in Kelejian and Prucha (1999) and Kapoor et al. (2007), which converges to the standard RMSE under a normal distribution:

$$RMSE = \sqrt{bias^2 + \left(\frac{IQ^2}{1.35}\right)^2}$$

where the *bias* is the difference between the true value of the coefficient and the median of the estimated coefficients; and IQ is the difference between the 75% and 25% quantile. This definition has the advantage of being more robust to outliers that may be generated by the Monte-Carlo simulations.

[12] The data is taken from CEPII database.

Since the results are qualitatively similar with respect to different spatial weight schemes, for sake of brievty we only present the results for "3 ahead and 3 behind" W. The full results are given in table 5.B. in appendix.

The Monte Carlo investigation highlights several important facts. First the use of different structure of instruments in system-GMM only affect marginally the unbiasness and efficiency of the estimates. Therefore, the extended-GMM results presented here are based on instrumenting each endogenous variables (Y_{t-1}, WY_t, EN_t) with their 2th and 3rd lags values (using the *collapse* option[13]) and the exogenous variables X_t and WX_t.

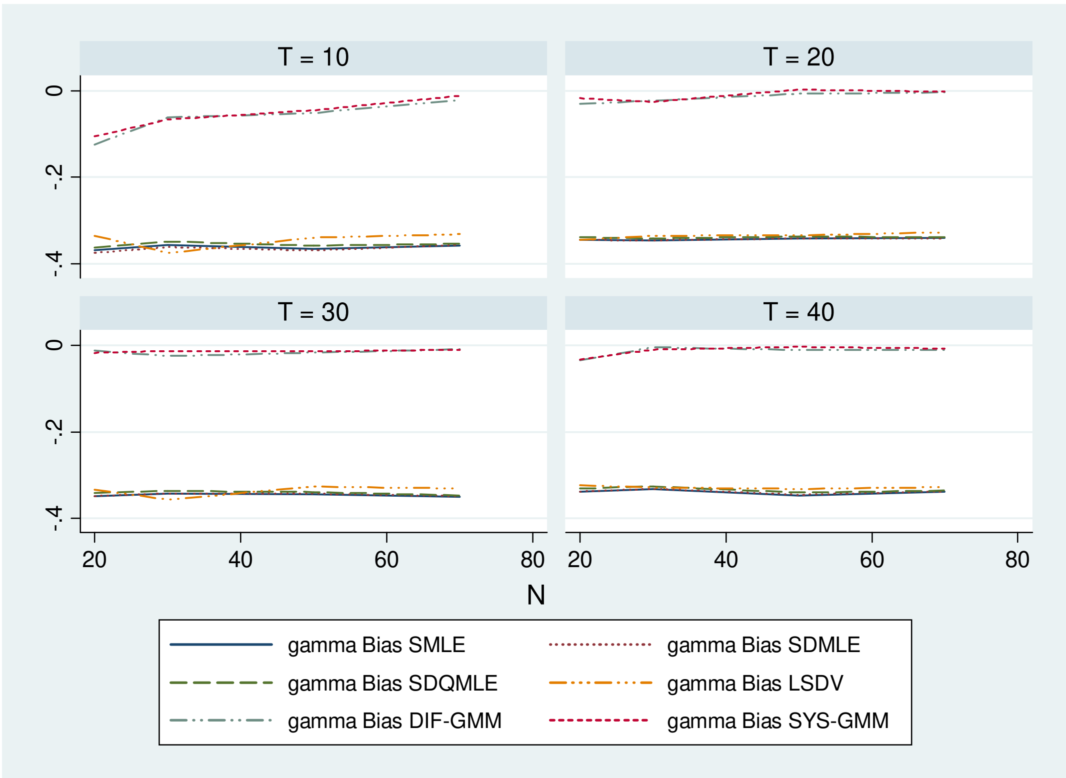

Figure 1: Endogenous variable γ Bias

[13]Instead of generating one column for each time period and lag available the instruments are collapsed. See Roodman (2006) for further details.

Monte Carlo results are reported in the appendix.In terms of unbiasness, there are differences according to the parameter considered. But overall, system-GMM is characterized by greater unbiasness than the other estimators. Moreover, extended-GMM is characterized by a faster rate of consistency than the other estimators. While system-GMM tends to overestimate the time lag coefficient to a smaller extent, the remaining estimators underestimate it. The same is true for the coefficient of the exogenous and endogenous. The latter shows how important it is to correct for the endogeneity. In fact, when endogeneity is not accounted for, the bias can represent more than 60% of the true value of the parameter, which is unacceptable (see Figure1). Moreover, the magnitude of the bias for the endogenous covariate does not seem to depend on the value of α and β and the sample dimension (N and T). On the contrary, independently of the estimator chosen except LSDV, the simultaneity bias associated with the spatial autoregressive is usually relatively small, especially when the spatial dependence is low. This result is in line with Franzese and Hays (2007), Yu et al. (2008) and Elhorst (2008). Beside extended-GMM, Spatial Dynamic QMLE is the estimator which displays lower bias for all coefficients,except for the endogenous variable, where LSDV performs better.

In terms of efficiency, summary of the results of RMSE and approximated RMSE, which are qualitatively similar, are less simpler. Despite the fact that spatial dynamic QMLE and MLE yield more bias but are more efficient than GMM for the spatial lag and exogenous variable, this is not the case for the time lag and endogenous variable. The estimate of the endogenous covariate is clearly more efficient with extended GMM than any other maximum likelihood estimators. Moreover, the rate of decrease of RMSE is almost null for the QMLE and MLE. In other words, increasing the dimension sample cannot improve efficiency of the estimate of the endogenous variable.

Finally, it is interesting to look at the accuracy of the estimators. This is measures as the ratio of the average of the reported standard-errors of each coefficient and the actual standard deviation of the estimated coefficient for each design. As it could be expected, GMM's results tend to be overconfident, which will result in the overrejection of the null hypothesis. However, the probability to obtain an estimation of the spatial autoregressive term which fails to respect the stationarity conditions is null in all designs for 1000 trials. Moreover, the results presented here for GMM are not based on robust standard errors. This could explain the discrepancy.

4 Conclusion

In the presence of endogenous covariates, our spatial dynamic panel simulations demonstrate that while the simultaneity bias of the spatial lag remains relatively low, the bias of the endogenous is large if it is not corrected. Proper correction leads to favour extended GMM. In fact, system-GMM emerges clearly dominant by an unbiasedness criterion for most variables, including the endogenous variable. Its RMSE decays at a faster rate as N or T increases and its standard error accuracy is acceptable. Moreover from a viewpoint purely practical, extended GMM avoids the inversion of a large spatial weight matrix, is easier to implement and its computation time is definitively lower (less than one minute) than any maximum likelihood estimators. Until a new estimator that allows to account for the endogeneity of the lagged dependent variable, spatial lag and other potentially endogenous variables is found, applied researchers can apply extended GMM to estimate "time-space simultaneous" models.

References

[1] Anselin, L., 1988, Spatial Econometrics: Methods and Models, Kluwer Academic Publishers, Dordrecht.

[2] Anselin, L., 1990, "What is Special about Spatial Data? Alternative Perspectives on Spatial Data Analysis", in D. Griffith, ed., Spatial Statistics: Past, Present and Future, Ann Arbor, Michingan.

[3] Anselin, L., 2001, "Spatial Econometrics", in A Companion to Theoretical Econometrics, ed., B.H. Baltagi, Blackwell Publishers Lte., Massachusetts.

[4] Anselin, L., 2003, "Spatial Externalities, Spatial Multipliers and Spatial Econometrics", International Regional Science Review 26, 153–166.

[5] Anselin, L., 2006, "Spatial Econometrics" In: Mills TC, Patterson K (eds) Palgrave handbook of econometrics: Volume 1, econometric theory. Palgrave MacMillan, Basingstoke.

[6] Arellano, M., and S. Bond, 1991, "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations", Review of Economic Studies 58, 277-97.

[7] Badinger, H., W. Müller and G. Tondl, 2004, "Regional Convergence in the European Union, 1985- 1999: A Spatial Dynamic Panel Analysis", Regional Studies 38, 241-253.

[8] Baltagi, B.H. and C. Kao, 2000, "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A Survey", Advances in Econometrics 68, 29-51.

[9] Baltagi, B. H., P. Egger, and M. Pfaffermayr, 2007, "Estimating Models of Complex FDI: Are There Third-Country Effects?", Journal of Econometrics 140, 260-281.

[10] Beenstock M. and D. Felsenstein, 2007, "Spatial Vector Autoregressions," Spatial Economic Analysis 2, 167-196.

[11] Blonigen, B. A., R. B. Davies, H. T. Naughton, and G. R. Waddell, 2008, "Spacey Parents: Spatial Autoregressive Patterns in Inbound FDI", in Inbound FDI, in S. Brakman and H. Garretsen (Eds.), Foreign Direct Investment and the Multinational Enterprise. Cambridge, MA: The MIT Press.

[12] Blonigen, B. A., R. B. Davies, G. R. Waddell, and H. T. Naughton, 2007, "FDI in Space: Spatial Autoregressive Relationships in Foreign Direct Investment", European Economic Review 51, 1303-1325.

[13] Blundell, R., and S. Bond, 1998, "Initial Conditions and Moment Restrictions in Dynamic Panel Data Models", Journal of Econometrics 87, 115-143.

[14] Cliff, A. D., and J. K. Ord, 1981, Spatial Processes: Models and Applications. Pion Ltd., London.

[15] Dall'erba S. and J. Le Gallo, 2008, "Regional Convergence and the Impact of European Structural Funds Over 1989–1999: A Spatial Econometric Analysis", Papers in Regional Science 87, 219–244.

[16] Dubin, R., 2004, "Spatial Lags and Spatial Errors Revisited: Monte Carlo Evidence", in Advances in Econometrics, Volume 18, Spatial and Spatiotemporal Econometrics, James P. LeSage and R. Kelley Pace (eds.), 75-98. Elsevier Ltd, Oxford.

[17] Elhorst, J. P., 2001, "Panel Data Models Extended to Spatial Error Autocorrelation or a Spatially Lagged Dependent Variable", Research Report 01C05, University of Groningen, Research Institute SOM (Systems, Organizations and Management). http://irs.ub.rug.nl/ppn/217984169

[18] Elhorst, J. P., 2003a, "Unconditional Maximum Likelihood Estimation of Dynamic Models for Spatial Panels", Research Report 03C27, University of Groningen, Research Institute SOM (Systems, Organizations and Management). http://irs.ub.rug.nl/ppn/25230568X

[19] Elhorst, J. P., 2003b, "Specification and Estimation of Spatial Panel Data Models", International Regional Science Review 26, 244-268.

[20] Elhorst, J. P., 2005, "Unconditional maximum likelihood estimation of linear and log-linear dynamic models for spatial panels," Geographical Analysis, 37, 85-106.

[21] Elhorst J.P., 2008, "Estimation of Dynamic Panels with Endogenous Interaction Effects", presented at the second World Conference of the Spatial Econometrics Association.

[22] Elhorst J.P., 2009, Spatial Panel Data Models, in Fischer MM, Getis A (Eds.), "Handbook of Applied Spatial Analysis", ch. C.2. Springer: Berlin Heidelberg New York.

[23] Fingleton, B. and J. Le Gallo, 2008, "Estimating Spatial Models with Endogenous Variables, a Spatial Lag and Spatially Dependent Disturbances: Finite Sample Properties", Papers in Regional Science, forthcoming.

[24] Foucault, M., T. Madies and S. Paty, 2008, "Public Spending Interactions and Local Politics. Empirical Evidence From French Municipalities," Public Choice 137, 57-80.

[25] Franzese, R. and J. C. Hays, 2007, "Spatial-Econometric Models of Cross-Sectional Interdependence in Political-Science Panel and Time-Series-Cross-Section Data," Political Analysis, 15, 140–164.

[26] Haining, R., "Estimating spatial-interaction models" Environment and Planning A10, 305–320.

[27] Hong, E., L. Sun, and T. Li, 2008, "Location of Foreign Direct Investment in China: A Spatial Dynamic Panel Data Analysis by Country of Origin", Discussion Paper 86, Department of Financial & Management Studies, University of London. http://www.cefims.ac.uk/documents/research-79.pdf

[28] Hsiao, C., M.H. Pesaran and A.K. Tahmiscioglu, 2002, "Maximum Likelihood Estimation of Fixed Effects Dynamic Panel Data Models Covering Short Time Periods", Journal of Econometrics, 109, 107-150.

[29] Kapoor M., H.H. Kelejian and I.R. Prucha, 2007, "Panel Data Models with Spatially Correlated Error Components". Journal of Econometrics 140, 97–130.

[30] Kelejian, H.H. and I.R. Prucha, 1998, "A Generalized Spatial Two-Stage Least Squares Procedure for Estimating a Spatial Autoregressive Model with Autoregressive Disturbances.", Journal of Real Estate Finance and Economics 17, 99–121.

[31] Kelejian, H.H. and I.R. Prucha, 1999, "A Generalized Moments Estimator for the Autoregressive Parameter in a Spatial Model". International Economic Review 40, 509–533.

[32] Keller, W. and C. Shiue, 2007, "The Origin of Spatial Interaction", Journal of Econometrics 140, 304-332.

[33] Korniotis, G. M., 2007, "Estimating Panel Models with Internal and External Habit Formation". http://ssrn.com/abstract=986726

[34] Kukenova, M. and J.-A. Monteiro, "Does Lax Environmental Regulation Attract FDI when accounting for "third-country" effects?". http://ssrn.com/abstract=1292705

[35] Lee, L., 2003, "Best Spatial Two-Stage Least Squares Estimators for a Spatial Autoregressive Model with Autoregressive Disturbances", Econometric Reviews 22, 307-335.

[36] Lee, L. and J. Yu, 2007, "A Spatial Dynamic Panel Data Model with Both Time and Individual Fixed Effects", University of Rochester. http://www.econ.ohio-state.edu/lee/wp/Time-Dummy-Spatial-Panel-0718.pdf

[37] Madariaga, N. and S. Poncet, 2007, "FDI in Chinese Cities: Spillovers and Impact on Growth", The World Economy 30, 837-862.

[38] Phillips, P. and H. Moon, 2000, "Nonstationary Panel Data Analysis: an Overview of some Recent Developments", Econometric Reviews 19: 263-286.

[39] Yu, J., R. de Jong and L. Lee, 2008, "Quasi-Maximum Likelihood Estimators For Spatial Dynamic Panel Data With Fixed Effects When Both n and T Are Large", Journal of Econometrics, 146, 118-134.

[40] Yu, J., R. de Jong and L. Lee, 2007, "Quasi-Maximum Likelihood Estimators For Spatial Dynamic Panel Data With Fixed Effects When Both n and T Are Large: A Nonstationary Case", University of Rochester. http://www.econ.ohio-state.edu/lee/wp/NonStationary-Spatial-Panel-0825.pdf

5 Appendices

5.A Spatial Estimators

This appendix section presents the procedure associated with the different spatial estimators. For further details, the reader is referred to Anselin (1988), Elhorst (2003a, 2005, 2008) and Yu et al. (2008). Let Y, Y_{-1}, WY, U be $N \cdot T$ column vectors, EX,is a $N \cdot T \times p$ matrix and EN is a $N \cdot T \times q$ matrix. Note that the data is first sorted by time T and then by cross-section N. Thus, $Y = (Y_1;\ Y_2;\ ...;\ Y_T)'$, where $Y_t = (Y_{1t};\ Y_{2t};\ ...;\ Y_{Nt})'$. The same structure is applied to the remaining vectors and matrices. These estimators can be implemented in Matlab.

5.A.1 Spatial MLE

The classical spatial maximum likelihood estimator relies on the concentrated likelihood in the spatial lag parameter, which is conditional upon the others' coefficient values. Operationally, "standard" spatial maximum estimation can be achieved in five steps:

1. Demean all variables, denoted by $\tilde{}$.
2. Carry out the following OLS regressions:
$\widetilde{Y} = \left[\widetilde{Y}_{-1}; \widetilde{EX}; \widetilde{EN}\right] b_0 + U_0$
$W\widetilde{Y} = \left[\widetilde{Y}_{-1}; \widetilde{EX}; \widetilde{EN}\right] b_L + U_L.$
3. Compute the associated residuals $\widehat{U}_0$ and $\widehat{U}_L$.
4. Given $\widehat{U}_0$ and $\widehat{U}_L$, find ρ that maximizes the following concentrated likelihood
$\ln L(\rho) = -\frac{NT}{2}\ln 2\pi - \frac{NT}{2}\ln\sigma^2 + T\ln|I_N - \rho W| - \frac{NT}{2}\ln\left[\left(\widehat{U}_0 - \rho\widehat{U}_L\right)'\left(\widehat{U}_0 - \rho\widehat{U}_L\right)\right].$
5. Given the estimate $\widehat{\rho}$, the remaining coefficient estimates are computed as follows:
$\begin{bmatrix} \widehat{\alpha} \\ \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} = b_0 - \widehat{\rho} b_L$ and $\widehat{\sigma}^2 = \frac{1}{NT}\left(\widehat{U}_0 - \widehat{\rho}\widehat{U}_L\right)'\left(\widehat{U}_0 - \widehat{\rho}\widehat{U}_L\right).$

As mentioned in Elhorst (2008), this spatial MLE is inconsistent, because of the presence of the lag dependent variable.

5.A.2 Spatial Dynamic MLE

The unconditional MLE, proposed by Elhorst (2005, 2008), involves a two-steps iterative procedure once the data has been first-differenced. Note that the initial observations are approximated using Bhargava and Sargan approach (1983). Estimation should proceed according to the following steps:

1. Take the first-difference of all variables;
2. Define some initial values for the parameters α, ρ and θ, where $\theta = \sigma_\xi^2/\sigma^2$ and σ_ξ^2 is the variance associated with the approximation of the initial observations.

3. The two-steps iterative procedure begins here with the computation of the coefficients π_i associated with the initial observations's approximation as well as the parameters of the exogenous and endogenous covariates, and the variance σ^2 :

$$\begin{bmatrix} \widehat{\pi}_1 \\ \widehat{\pi}_2 \\ \vdots \\ \widehat{\pi}_T \\ \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} = \left(\underline{\Delta X}' H_{V\theta}^{-1} \underline{\Delta X}\right)^{-1} \underline{\Delta X}' H_{V\theta}^{-1} \underline{\Delta Y} \qquad \text{and} \qquad \widehat{\sigma}^2 = \frac{\Delta \widehat{U}' H_{V\theta}^{-1} \Delta \widehat{U}}{NT}$$

where
$$\underline{\Delta X} = \begin{bmatrix} \imath_N & \Delta X_1 & \cdots & \Delta X_T & 0 \\ 0 & 0 & \cdots & 0 & \Delta X_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \Delta X_T \end{bmatrix};$$

$$\underline{\Delta Y} = \begin{bmatrix} (I_N - \rho W)\,\Delta Y_1 \\ (I_N - \rho W)\,\Delta Y_2 - \alpha \Delta Y_1 \\ \vdots \\ (I_N - \rho W)\,\Delta Y_T - \alpha \Delta Y_{T-1} \end{bmatrix};$$

$$H_{V\theta} \begin{bmatrix} V_\theta & -I_N & 0 & \cdots & 0 & 0 \\ -I_N & 2\cdot I_N & -I_N & \ddots & 0 & 0 \\ 0 & -I_N & 2\cdot I_N & \ddots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 2\cdot I_N & -I_N \\ 0 & 0 & 0 & \cdots & -I_N & 2\cdot I_N \end{bmatrix};$$

$$\begin{aligned} V_\theta = {} & \theta I_N + I_N + (\alpha S - I_N)\left(I_N - \alpha^2 S S'\right)^{-1} (\alpha S - I_N)' \\ & - (\alpha S - I_N)(\alpha S)^{m-1} \left(I_N - \alpha^2 S S'\right)^{-1} (\alpha S)^{m-1} (\alpha S - I_N)' \\ & - \left(\alpha^2 S S'\right)^{m-1}; \end{aligned}$$

$$S = (I_N - \rho W)^{-1};$$

$$\Delta \widehat{U} = \underline{\Delta Y} - \underline{\Delta X} \cdot \left(\widehat{\pi}_1;\ ...;\ \widehat{\pi}_T;\ \widehat{\beta}';\ \widehat{\gamma}'\right);$$

The parameter m, which represents the number of periods since the process started,should be defined in advance. It must be such that the eigenvalues of the matrix αS lie inside the unit circle, because otherwise the matrix $(\alpha S)^{m-1}$ would become infinite and yield a corner solution. Elhorst (2008) proposes to include a third step procedure to estimate m. Beside increasing the computation time, this additional step affects minorily the results.

4. Given the set of parameters obtained in step 3, maximize the unconditional likelihood function as follows:

$$\ln L(\alpha, \rho, \theta) = -\frac{NT}{2} \ln 2\pi - \frac{NT}{2} \ln \sigma^2 + T \ln |I_N - \rho W| - \frac{1}{2} \ln |H_{V\theta}| - \frac{1}{2\sigma^2} \Delta \widehat{U}' H_{V\theta}^{-1} \Delta \widehat{U}$$

w.r.t. $|\alpha| < 1 - \rho\omega_{\max}$ and $|\alpha| < 1 - \rho\omega_{\min}$

5. Repeat step 3, with the estimates obtained in step 4 and so on.., until convergence is met.

Note that to reduce the computation time the jacobian term, $\ln |I_N - \rho W|$, in the loglikelihood function is approximation by $\sum_{i=1}^{N} \ln(1 - \rho\omega_i)$, where ω_i is the eigenvalue of the matrix W. The inverse of matrix $H_{V\theta}$ is also estimated using summation operations instead of matrix calculus.

5.A.3 Spatial Dynamic QMLE

The QMLE, presented by Yu et al. (2008), requires first the maximization of the concentrated likelihood and then a bias correction. The estimation process involves the following steps:

1. Demean all variables, denoted by $\tilde{}$.
2. Maximize the following concentrated likelihood function in order to estimate $\widehat{\alpha}$, $\widehat{\rho}$, $\widehat{\beta}$, $\widehat{\gamma}$ and $\widehat{\sigma}^2$

$$\ln L\left(\alpha,\rho,\beta,\gamma,\sigma^2\right) = -\frac{NT}{2}\ln 2\pi - \frac{NT}{2}\ln\sigma^2 + T\ln\left|I_N-\rho W\right| - \frac{1}{2\sigma^2}\sum_{t=1}^{T}\tilde{U}_t'\tilde{U}_t$$

$$\begin{aligned} \text{w.r.t.}\quad & \sum_{t=1}^{T}\tilde{Y}_{-1}'\tilde{U}_t = 0 \\ & \sum_{t=1}^{T}\left(W\tilde{Y}_{-1}'\right)'\tilde{U}_t = tr\left(W\left(I_N-\rho W\right)^{-1}\right) \\ & \sum_{t=1}^{T}\widetilde{EX}'\tilde{U}_t = 0 \\ & \sum_{t=1}^{T}\widetilde{EN}'\tilde{U}_t = 0 \\ & \sum_{t=1}^{T}\tilde{U}_t'\tilde{U}_t = N\sigma^2 \end{aligned}$$

where $\tilde{U}_t = \left(I_N-\rho W\right)\tilde{Y}_t - \left[\tilde{Y}_{-1};\widetilde{EX};\widetilde{EN}\right]\left[\alpha;\ \beta';\ \gamma'\right]'$

3. The bias-corrected estimator is then given by:

$$\begin{bmatrix} \widehat{\alpha}^c \\ \widehat{\rho}^c \\ \widehat{\beta}^c \\ \widehat{\gamma}^c \\ \widehat{\sigma}^{2c} \end{bmatrix} = \begin{bmatrix} \widehat{\alpha} \\ \widehat{\rho} \\ \widehat{\beta} \\ \widehat{\gamma} \\ \widehat{\sigma}^2 \end{bmatrix} - \frac{1}{T}\left(-\widehat{\Sigma}^{-1}b\right)$$

where $\widehat{\Sigma}^{-1}$ can be approximated by the empirical Hessian matrix of the concentrated log likelihood function (an analytical expression for the matrix Σ can also be found in Yu et al.) and the column matrix b is given by:

$$b = \begin{bmatrix} \frac{1}{N}tr\left(\left(I_N-\widehat{\alpha}\left(I_N-\widehat{\rho}W\right)^{-1}\right)\left(I_N-\widehat{\rho}W\right)^{-1}\right) \\ \frac{\widehat{\alpha}}{N}tr\left(W\left(I_N-\widehat{\rho}W\right)^{-1}\left(I_N-\widehat{\alpha}\left(I_N-\widehat{\rho}W\right)^{-1}\right)\left(I_N-\widehat{\rho}W\right)^{-1}\right)+\frac{1}{N}tr\left(W\left(I_N-\widehat{\rho}W\right)^{-1}\right) \\ 0 \\ 0 \\ \frac{1}{2\widehat{\sigma}^2} \end{bmatrix}$$

4. Finally, the individual effects are recovered as follows:

$$\widehat{\eta} = \frac{1}{T}\sum_{t=1}^{T}\left(I_N-\widehat{\rho}^cW\right)Y_t - \left[Y_{-1};EX;EN\right]\left[\widehat{\alpha}^c;\ \widehat{\beta}^{c\prime};\ \widehat{\gamma}^{c\prime}\right]'$$

5.B Monte Carlo Results: Bias

Time lag variable α: Bias

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.042	0.035	0.031	0.053	0.036	0.002
20	20	0.030	0.024	0.027	0.041	0.008	0.000
30	20	0.027	0.024	0.032	0.036	0.009	0.005
40	20	0.026	0.023	0.032	0.040	0.011	0.002
10	30	0.044	0.031	0.030	0.020	0.022	0.004
20	30	0.029	0.024	0.029	0.039	0.012	−0.004
30	30	0.031	0.026	0.030	−0.002	0.010	0.004
40	30	0.024	0.022	0.027	0.024	0.000	0.001
10	50	0.044	0.029	0.034	0.056	0.018	0.000
20	50	0.033	0.027	0.033	0.036	0.010	0.003
30	50	0.027	0.023	0.033	0.041	0.004	−0.001
40	50	0.031	0.029	0.034	0.041	0.006	0.001
10	70	0.040	0.027	0.025	0.061	0.010	−0.003
20	70	0.029	0.021	0.024	0.046	0.002	−0.003
30	70	0.045	0.040	0.041	0.052	0.004	0.000
40	70	0.030	0.026	0.030	0.041	0.002	0.001

Spatial lag variable ρ: Bias

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.013	0.025	0.011	−0.063	−0.007	−0.012
20	20	0.015	0.018	0.017	−0.005	−0.012	−0.005
30	20	0.015	0.018	0.014	−0.029	−0.001	−0.001
40	20	0.011	0.013	0.007	−0.029	−0.009	−0.008
10	30	0.017	0.031	0.020	−0.001	−0.009	−0.029
20	30	0.018	0.022	0.021	−0.031	0.001	0.004
30	30	0.014	0.016	0.015	−0.030	−0.002	−0.003
40	30	0.012	0.014	0.012	−0.030	−0.002	−0.004
10	50	0.016	0.024	0.021	−0.040	−0.003	−0.003
20	50	0.012	0.017	0.014	−0.027	−0.004	−0.002
30	50	0.016	0.019	0.018	−0.032	−0.008	−0.007
40	50	0.014	0.015	0.015	−0.026	−0.004	−0.006
10	70	0.012	0.021	0.018	−0.045	−0.003	−0.002
20	70	0.012	0.016	0.015	−0.027	−0.004	−0.008
30	70	0.028	0.031	0.030	−0.025	−0.004	−0.001
40	70	0.016	0.017	0.017	−0.020	−0.004	−0.002

Exogenous lag variable β: Bias

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	−0.046	−0.050	−0.033	0.040	0.014	0.030
20	20	−0.048	−0.050	−0.035	−0.027	−0.001	0.036
30	20	−0.037	−0.038	−0.020	0.014	0.007	0.014
40	20	−0.043	−0.040	−0.033	−0.009	0.000	0.013
10	30	−0.039	−0.041	−0.030	0.006	0.013	0.046
20	30	−0.037	−0.035	−0.028	0.005	0.017	0.006
30	30	−0.045	−0.044	−0.037	0.040	0.008	−0.003
40	30	−0.041	−0.040	−0.034	0.020	−0.003	0.009
10	50	−0.048	−0.052	−0.035	0.006	−0.006	0.008
20	50	−0.041	−0.041	−0.032	−0.003	0.002	0.012
30	50	−0.048	−0.046	−0.043	0.002	0.001	0.006
40	50	−0.047	−0.046	−0.041	−0.012	0.007	0.008
10	70	−0.037	−0.035	−0.026	0.021	0.017	0.013
20	70	−0.051	−0.049	−0.042	−0.024	0.009	0.021
30	70	−0.064	−0.062	−0.060	−0.020	0.004	0.001
40	70	−0.046	−0.044	−0.042	−0.014	0.002	0.008

Endogenous lag variable γ: Bias

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	-0.369	-0.375	-0.363	-0.336	-0.124	-0.106
20	20	-0.345	-0.346	-0.339	-0.345	-0.031	-0.018
30	20	-0.349	-0.349	-0.342	-0.334	-0.012	-0.017
40	20	-0.338	-0.338	-0.332	-0.324	-0.035	-0.032
10	30	-0.357	-0.362	-0.349	-0.376	-0.062	-0.067
20	30	-0.346	-0.346	-0.342	-0.336	-0.023	-0.026
30	30	-0.343	-0.343	-0.337	-0.357	-0.025	-0.013
40	30	-0.332	-0.332	-0.326	-0.329	-0.004	-0.010
10	50	-0.367	-0.370	-0.359	-0.340	-0.052	-0.045
20	50	-0.343	-0.342	-0.338	-0.335	-0.007	0.003
30	50	-0.344	-0.343	-0.339	-0.327	-0.017	-0.014
40	50	-0.347	-0.346	-0.341	-0.332	-0.010	-0.003
10	70	-0.359	-0.357	-0.354	-0.332	-0.021	-0.012
20	70	-0.341	-0.342	-0.339	-0.328	-0.004	-0.002
30	70	-0.350	-0.349	-0.347	-0.331	-0.009	-0.010
40	70	-0.338	-0.338	-0.336	-0.328	-0.011	-0.007

5.C Monte Carlo Results: RMSE

Time lag variable α: RMSE

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.003	0.026	0.002	0.004	0.010	0.011
20	20	0.002	0.021	0.001	0.002	0.003	0.005
30	20	0.002	0.011	0.002	0.002	0.002	0.002
40	20	0.001	0.011	0.001	0.002	0.001	0.002
10	30	0.003	0.002	0.002	0.002	0.006	0.006
20	30	0.002	0.012	0.001	0.002	0.002	0.002
30	30	0.001	0.001	0.001	0.001	0.001	0.002
40	30	0.001	0.011	0.001	0.001	0.001	0.001
10	50	0.003	0.019	0.002	0.004	0.003	0.004
20	50	0.002	0.029	0.001	0.002	0.001	0.002
30	50	0.001	0.012	0.001	0.002	0.001	0.001
40	50	0.002	0.020	0.002	0.002	0.001	0.001
10	70	0.003	0.034	0.001	0.004	0.002	0.003
20	70	0.001	0.001	0.001	0.002	0.001	0.001
30	70	0.002	0.005	0.002	0.003	0.000	0.001
40	70	0.001	0.001	0.001	0.002	0.000	0.001

Spatial lag variable ρ: RMSE

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.002	0.008	0.002	0.006	0.015	0.013
20	20	0.001	0.008	0.001	0.002	0.003	0.008
30	20	0.001	0.003	0.001	0.003	0.003	0.002
40	20	0.001	0.001	0.001	0.002	0.002	0.002
10	30	0.001	0.002	0.002	0.004	0.011	0.009
20	30	0.001	0.004	0.001	0.002	0.003	0.002
30	30	0.001	0.001	0.001	0.002	0.002	0.002
40	30	0.001	0.003	0.001	0.002	0.001	0.001
10	50	0.001	0.007	0.001	0.003	0.005	0.005
20	50	0.001	0.013	0.001	0.001	0.002	0.002
30	50	0.001	0.003	0.001	0.002	0.001	0.002
40	50	0.001	0.001	0.001	0.001	0.001	0.001
10	70	0.001	0.013	0.001	0.003	0.004	0.004
20	70	0.000	0.001	0.001	0.001	0.001	0.001
30	70	0.001	0.002	0.001	0.003	0.001	0.001
40	70	0.001	0.001	0.001	0.001	0.001	0.001

Exogenous lag variable β: RMSE

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.006	0.006	0.006	0.008	0.016	0.032
20	20	0.004	0.004	0.004	0.004	0.006	0.020
30	20	0.003	0.003	0.003	0.003	0.004	0.006
40	20	0.003	0.008	0.003	0.002	0.003	0.006
10	30	0.005	0.005	0.005	0.005	0.011	0.019
20	30	0.003	0.003	0.003	0.002	0.004	0.007
30	30	0.003	0.003	0.003	0.004	0.003	0.007
40	30	0.003	0.003	0.003	0.002	0.002	0.003
10	50	0.004	0.006	0.004	0.003	0.007	0.010
20	50	0.003	0.003	0.003	0.002	0.002	0.005
30	50	0.003	0.003	0.003	0.001	0.002	0.004
40	50	0.003	0.020	0.003	0.002	0.001	0.002
10	70	0.003	0.003	0.003	0.002	0.004	0.007
20	70	0.003	0.003	0.003	0.001	0.002	0.004
30	70	0.004	0.004	0.004	0.003	0.001	0.002
40	70	0.003	0.003	0.002	0.001	0.001	0.002

Endogenous lag variable γ: RMSE

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.143	0.145	0.137	0.120	0.085	0.088
20	20	0.121	0.121	0.116	0.121	0.029	0.034
30	20	0.121	0.121	0.114	0.115	0.015	0.019
40	20	0.118	0.122	0.111	0.107	0.010	0.011
10	30	0.129	0.131	0.124	0.146	0.042	0.045
20	30	0.123	0.123	0.118	0.113	0.017	0.023
30	30	0.119	0.118	0.114	0.127	0.009	0.012
40	30	0.114	0.113	0.107	0.109	0.006	0.008
10	50	0.135	0.139	0.129	0.118	0.026	0.030
20	50	0.119	0.118	0.113	0.113	0.008	0.010
30	50	0.119	0.118	0.113	0.108	0.006	0.008
40	50	0.118	0.125	0.113	0.110	0.004	0.006
10	70	0.129	0.129	0.126	0.111	0.016	0.014
20	70	0.118	0.117	0.114	0.108	0.007	0.009
30	70	0.123	0.122	0.120	0.109	0.004	0.006
40	70	0.115	0.115	0.112	0.108	0.004	0.005

5.D Monte Carlo Results: SE accuracy

Time lag variable α: SE accuracy

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.125	0.095	0.138	0.156	0.479	0.445
20	20	0.086	0.067	0.093	0.102	0.305	0.281
30	20	0.059	0.057	0.060	0.072	0.202	0.192
40	20	0.050	0.048	0.052	0.061	0.169	0.161
10	30	0.106	0.113	0.121	0.117	0.399	0.397
20	30	0.062	0.058	0.069	0.076	0.227	0.216
30	30	0.048	0.054	0.053	0.059	0.168	0.158
40	30	0.044	0.042	0.047	0.055	0.145	0.142
10	50	0.081	0.065	0.093	0.097	0.287	0.288
20	50	0.046	0.035	0.051	0.055	0.170	0.171
30	50	0.041	0.038	0.043	0.048	0.139	0.139
40	50	0.037	0.031	0.039	0.044	0.128	0.118
10	70	0.067	0.047	0.083	0.085	0.276	0.270
20	70	0.038	0.044	0.046	0.049	0.152	0.151
30	70	0.059	0.054	0.063	0.066	0.173	0.170
40	70	0.032	0.035	0.036	0.037	0.106	0.105

Spatial lag variable ρ: SE accuracy

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.116	0.090	0.119	0.120	0.353	0.388
20	20	0.083	0.069	0.079	0.079	0.258	0.247
30	20	0.063	0.054	0.055	0.063	0.189	0.184
40	20	0.052	0.044	0.045	0.052	0.157	0.159
10	30	0.107	0.088	0.116	0.120	0.364	0.356
20	30	0.072	0.061	0.069	0.075	0.210	0.219
30	30	0.056	0.048	0.053	0.064	0.172	0.172
40	30	0.042	0.036	0.038	0.045	0.123	0.123
10	50	0.076	0.063	0.086	0.080	0.250	0.259
20	50	0.052	0.044	0.051	0.057	0.170	0.176
30	50	0.044	0.038	0.041	0.046	0.137	0.134
40	50	0.038	0.033	0.035	0.040	0.116	0.117
10	70	0.084	0.068	0.102	0.091	0.267	0.292
20	70	0.067	0.057	0.068	0.067	0.201	0.208
30	70	0.051	0.045	0.051	0.049	0.153	0.153
40	70	0.034	0.031	0.034	0.035	0.102	0.104

Exogenous lag variable β: SE accuracy

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.857	35.026	0.778	0.807	0.952	0.705
20	20	0.805	13.777	0.648	0.661	1.062	0.553
30	20	0.790	8.734	0.504	0.592	0.943	0.752
40	20	0.735	2.839	0.451	0.652	0.986	0.694
10	30	0.835	21.602	0.767	0.780	0.946	0.767
20	30	0.740	8.765	0.582	0.697	1.030	0.783
30	30	0.661	4.865	0.494	0.541	0.980	0.596
40	30	0.705	3.876	0.420	0.515	1.049	0.734
10	50	0.765	9.543	0.678	0.714	0.888	0.806
20	50	0.657	4.673	0.465	0.633	1.137	0.688
30	50	0.660	2.954	0.425	0.561	0.970	0.607
40	50	0.548	0.416	0.356	0.470	1.120	0.769
10	70	0.752	8.827	0.735	0.728	1.022	0.783
20	70	0.553	2.743	0.489	0.639	1.055	0.689
30	70	0.646	2.030	0.534	0.356	0.977	0.692
40	70	0.622	1.434	0.463	0.518	1.071	0.707

Endogenous lag variable γ: SE accuracy

T	N	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.801	21.269	0.737	0.758	0.881	0.887
20	20	0.792	9.052	0.684	0.792	0.879	0.787
30	20	0.876	6.540	0.613	0.680	0.985	0.862
40	20	0.827	2.529	0.540	0.713	1.092	0.962
10	30	0.894	14.899	0.894	0.807	0.880	0.894
20	30	0.794	6.101	0.669	0.794	0.932	0.816
30	30	0.787	3.800	0.615	0.703	1.045	0.921
40	30	0.751	2.694	0.518	0.674	1.098	0.935
10	50	0.840	5.838	0.816	0.796	0.904	0.883
20	50	0.807	3.657	0.599	0.779	1.052	0.948
30	50	0.820	2.394	0.539	0.636	1.007	0.885
40	50	0.730	0.525	0.478	0.658	0.993	0.855
10	70	0.812	5.820	0.831	0.889	0.939	1.018
20	70	0.885	2.806	0.728	0.917	0.943	0.855
30	70	0.718	1.501	0.629	0.546	1.017	0.883
40	70	0.738	1.150	0.603	0.635	0.946	0.825