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

Madina Kukenova[†] José-Antonio Monteiro[‡]

Draft: March 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 distributed as $(\mathbf{0}, \Omega)$. 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].

[1] The analysis of the spatial error panel model and the spatial lag with spatial error model is beyond the scope of this paper. For further details, see Elhorst (2005), Kapoor et al. (2007), Kelejian & Prucha (2007), Mutl & Pfaffermeier (2008) as well as Lee and Ju (2009).

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 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. The process is covariance stationary if

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

[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.

$\left|(I_N - \rho W_t)^{-1}\alpha\right| < 1$, or, equivalently, if

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

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:

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

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 υ_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)).

Table 1: Estimators Survey

Model	**Estimation Methods**	**Endogenous**
$Y_t = \alpha Y_{t-1} + \beta EX_t + \epsilon_t$	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))	$Y_{t-1}; EN_t$
$Y_t = \alpha Y_{t-1} + \rho WY_{t-1} + \beta EX_t + \gamma EN_t + \epsilon_t$	LSDV-IV (Korniotis (2008))	$Y_{t-1}; EN_t$
$Y_t = \rho WY_t + \beta EX_t + \epsilon_t$	Spatial-MLE (Anselin (1988) (2001), Elhorst (2003)) Spatial 2SLS (Anselin (1988) (2001)))	WY_t
$Y_t = \rho WY_t + \beta EX_t + \gamma EN_t + \epsilon_t$	Spatial 2SLS (Dall'erba & Le Gallo (2007))	$WY_t; EN_t$
$Y_t = \alpha Y_{t-1} + \rho WY_t + \beta EX_t + \epsilon_t$	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))	$WY_t; Y_{t-1}$
$Y_t = \alpha Y_{t-1} + \rho WY_t + \beta EX_t + \gamma EN_t + \epsilon_t$	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[4]. 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[5].

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[6]. The main difference

[4] 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 ($W_t \cdot \left(I_N - \widetilde{\widehat{\rho}} W_t\right)^{-1} \cdot EX_t \widetilde{\widehat{\beta}}$, where $\widetilde{\widehat{\rho}}$ and $\widetilde{\widehat{\beta}}$ are estimates obtained using the S2SLS proposed by Kelejian and Prucha (1998)).

[5] 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.

[6] 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[7]. 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.[8]. 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. As GMM framework is based on linear and quadratic moments, it is not necessary to impose parametric constraints On a practical ground, its implementation is computationally tractable as it avoids the inversion of high dimension spatial weights matrix W and the computation of its eigenvalues[9], 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 ($t = 1, .., T$):

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

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

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

[9]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

$$
\begin{aligned}
E\left(EN_{i,t-\tau}\triangle v_{it}\right) &= 0;\ \ \text{for } t=3...T \ \text{ and } \ 2\leq\tau\leq t-1 &(8)\\
E\left(\left[W_{t-\tau}Y_{t-\tau}\right]_i\triangle v_{it}\right) &= 0;\ \ \text{for } t=3...T \ \text{ and } \ 2\leq\tau\leq t-1 &(9)
\end{aligned}
$$

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:

$$
\begin{aligned}
E\left(\triangle Y_{i,t-1}v_{it}\right) &= 0;\ \ \text{for } t=3,...,T &(10)\\
E\left(\triangle EX_{it}v_{it}\right) &= 0;\ \ \text{for } t=2,...,T &(11)\\
E\left(\triangle EN_{it-1}v_{it}\right) &= 0;\ \ \text{for } t=3,...,T &(12)\\
E\left(\triangle\left[W_{t-1}Y_{t-1}\right]_i v_{it}\right) &= 0;\ \ \text{for } t=3,...,T &(13)
\end{aligned}
$$

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. Second, the Arellano-Bond test examines the serial correlation property of the level residuals.

Another issue lies in the fact that the instrument count grows as the sample size T rises. A large number of instruments can overfit endogenous variables (i.e. fail to correct for endogeneity) and leads to inaccurate estimation of the optimal weight matrix, downward biased two-step standard errors and wrong inference in the Hansen test of instruments validity. Okui (2008) demonstrates that the bias of extended GMM does not result from the total number of instruments, but from the number of instruments for each equation. As pointed out by Roodman (2009), it is advised to restrict the number of instruments by defining a maximum number of lags and/or by collapsing the instruments[10]. The *collapse* option consists of combining instruments through addition into subsets. For instance, if the instruments are collapsed, the moment condition (5) becomes:

$$E\left(Y_{i,t-\tau}\triangle v_{it}\right)=0;\quad \text{for } 2\leq\tau\leq t-1 \tag{14}$$

This modified moment condition still imposes the orthogonality of $Y_{i,t-\tau}$ and $\triangle v_{it}$, but rather to hold for each t and τ, it is only valid for each τ. Roodman (2009) shows that collapsed instruments lead to less biased estimates, although the associated standard error increases. GMM results for the remaining of the paper are thus based on collapsed instruments[11]

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[12]. Simulation studies already showed that bias associated with the spatial lag is rather small (Franzese and Hays (2007), Elhorst (2008)), but none analyzes the consequences of an additional endogenous explanatory variable in a spatial dynamic context. The data generating process (DGP) is defined as follows:

$$\begin{aligned}
Y_{it} &= \alpha Y_{i,t-1}+\rho\left[WY_t\right]_i+\beta EX_{it}+\gamma EN_{it}+\eta_i+v_{it} && (15)\\
EX_{it} &= \delta EX_{i,t-1}+u_{it} && (16)\\
EN_{it} &= \lambda EN_{i,t-1}+\psi\eta_i+\theta v_{it}+e_{it} && (17)
\end{aligned}$$

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

[10] This approach has been adopted in several empirical papers, including Beck & Levine (2004) and Crackovic & Levine (2005).

[11] See appendix for further details on extended GMM.

[12] Simulations are performed using Matlab R2008b.

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 dependent variable is generated according to the reduced form of equation (14):

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

Following Kapoor et al. (2007) and Kelejian and Prucha (1999), we consider three different types of spatial weight matrix. In each case, the matrice is row-standardized so that all non zero elements in each row sum to one. The matrices considered rely on a perfect "idealized" circular world. More precisely, 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.

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

$$T \in \{10, 20, 30, 40\};$$
$$N \in \{20;\ 30;\ 50;\ 70\};$$
$$\alpha \in; \{0.2;\ 0.4;\ 0.5;\ 0.7\}$$
$$\rho \in \{0.1;\ 0.3;\ 0.5;\ 0.7\};$$
$$\beta = 1;\ \delta = 0.65;\ \gamma = 0.5;\ \lambda = 0.45;\ \psi = 0.25; \theta = 0.6;$$
$$\sigma_u^2 = 0.05\ \sigma_v^2 = 0.05;\ \sigma_e^2 = 0.05;$$

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. The total number of designs is restricted to 160. For each of these designs, we performed 1000 trials. Note that for each design, the initial conditions and spatial weight matrices are generated once.

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.

The Monte Carlo investigation highlights several important facts[13]. First, the results are qualitatively similar with respect to different spatial weight schemes, that is why for sake of brevity we only present the results for "1 ahead and 1 behind" W. Second, the results in terms of bias and efficiency depend on the values assigned to the spatial and time lag parameters. The results based on RMSE and approximated RMSE are qualitatively similar, that is why we only present the results associated with standard RMSE. In order to assess the global properties of the estimators, we first discuss the results obtained by averaging over the parameters values, and ultimately describe the results in terms of response functions

[13]The full results are given in table 5.B. in appendix.

3.1 Extended GMM vs. Difference GMM

We first investigate the consistency and efficiency of the GMM estimators and least square dummy variables estimator in a spatial dynamic panel framework. As mentioned previously, we apply two-step difference and system GMM based on a collapsed instruments structure. To avoid imprecise estimate of the optimal weight matrix (when $T > N$), we restrict the lags to two and three. In other words, each endogenous variables (Y_{t-1}, WY_t, EN_t) is instrumented with their 2th and 3rd lags values (using the *collapse* option) and the exogenous variables X_t.

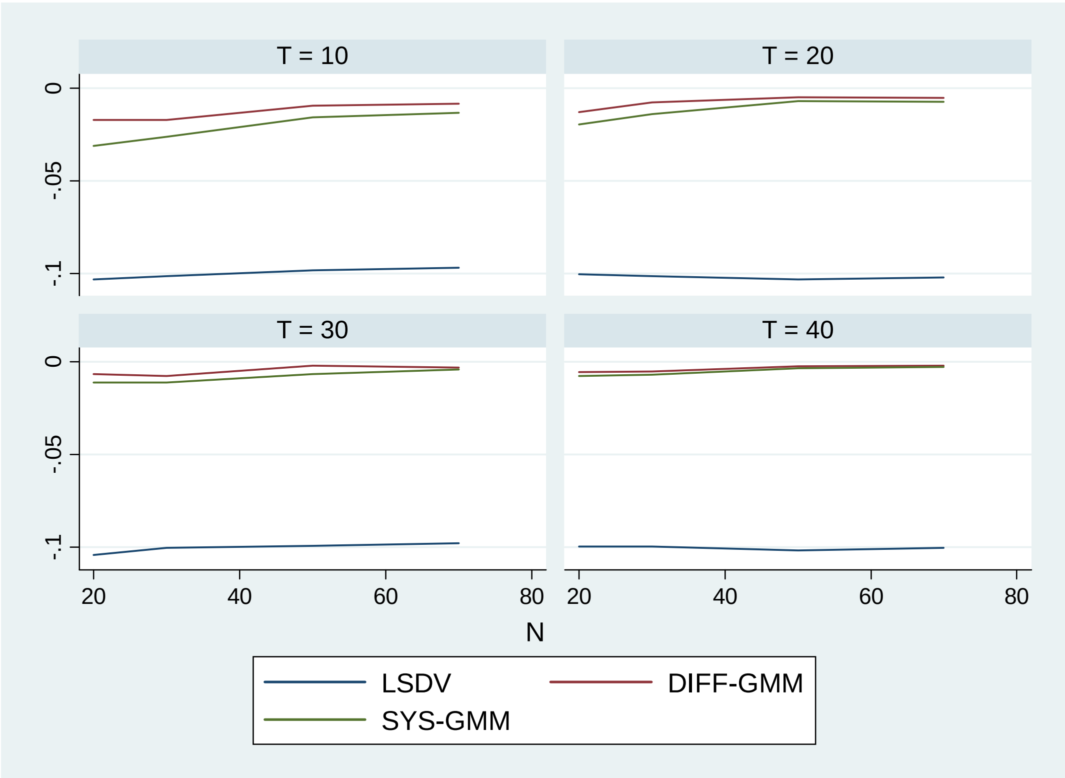

Figure 1: Figure 1: Bias GMM

Since the main goal of this paper is not to compare GMM estimators between themselves, we consider global consistency and efficiency (the bias and RMSE results are averaged over all parameters). For illustrative purpose, we also compare the global performance of LSDV estimator. As figure 1 and 2 show it, system and difference GMM outperform the fixed effect estimator in terms of bias and efficiency. In fact, LSDV estimator presents a negative bias which does not decrease with the size of the panel, making it clearly an inconsistent and inefficient estimator. Consequently, spatial dynamic panel model should definitively not be estimated by applying LSDV. The differences in terms of performance between difference and system GMM are relatively marginal. Hayakawa (2007) shows that difference and level GMM estimators can be more biased than system GMM, because the bias of extended-GMM is the sum of two elements. The first one is the weighted sum of the bias associated with difference and level GMM, while the second one originates from using the level and difference estimators jointly. Since the bias of difference and level GMM evolve in opposite direction, the first component of the bias of system GMM will tend to be small due to a partially cancelling out effect. In addition, the weight of the first element plays also a role in the difference of the magnitudes of the biases. This could explain why both GMM estimators performs almost equivalently. However as depicted in figure 2, system GMM seems to be more efficient in small sample. That is why, we decide to focus on extended GMM.

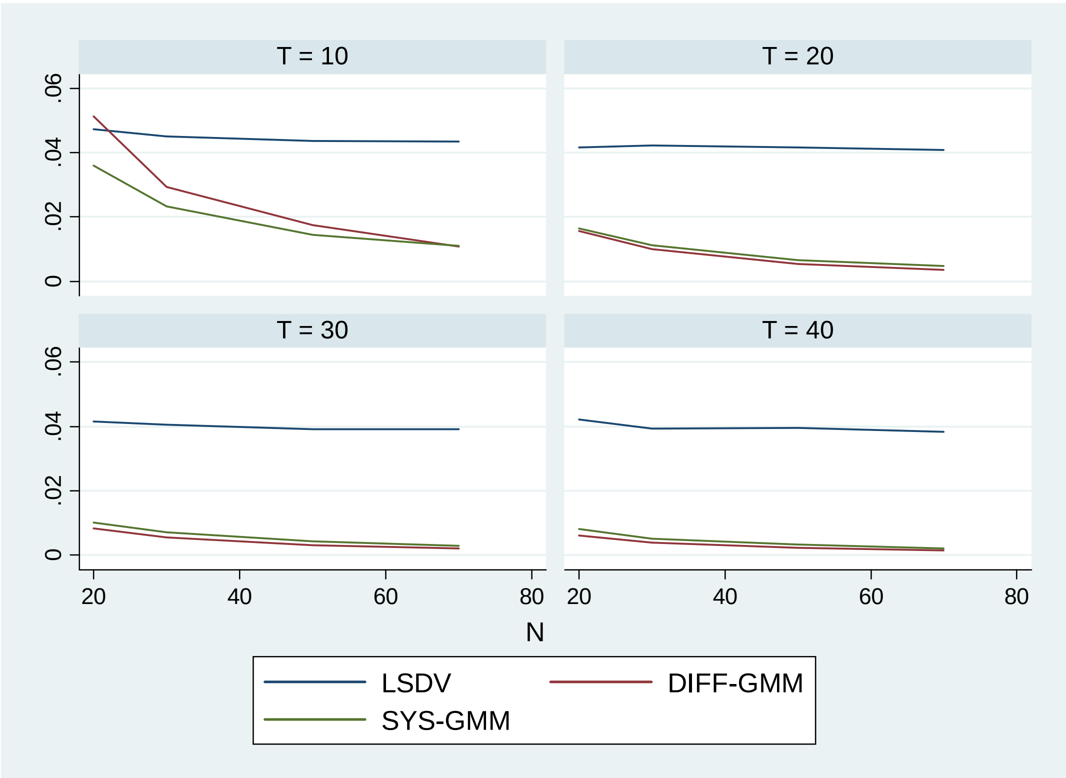

Figure 2: Figure 2: RMSE GMM

3.2 Extended GMM vs. Spatial Estimators

In this section we compare the performance of extended GMM with respect to spatial ML estimators. In terms of unbiasness, there are differences according to the parameter considered, as shown in figure 3. But overall, system-GMM is characterized by greater unbiasness than the other estimators. The only exception concerns the spatial lag parameter. This is not surprising since spatial estimators explicitly account for the spatial structure of the data by estimating the reduced form of the model. However, the estimation of the spatial lag becomes less bias using extended GMM as the sample size increases. Moreover, the rate of consistency is faster than the soatial estimators. Interestingly, all estimators tend to overestimate the spatial lag parameter and underestimate the time lag, the exogenous variable as well as the endogenous variable.

Most importantly, the plot on the bottom left of figure 3 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. Moreover, the magnitude of the bias for the endogenous covariate does not seem to depend on the value of the spatial and time lag and the sample dimension (N and T). This finding suggests that estimating a spatial dynamic panel model with several endogenous variables using traditional spatial ML estimator would not be advisable. In any case, beside extended-GMM, spatial dynamic QMLE is the estimator which displays lower bias for all coefficients.

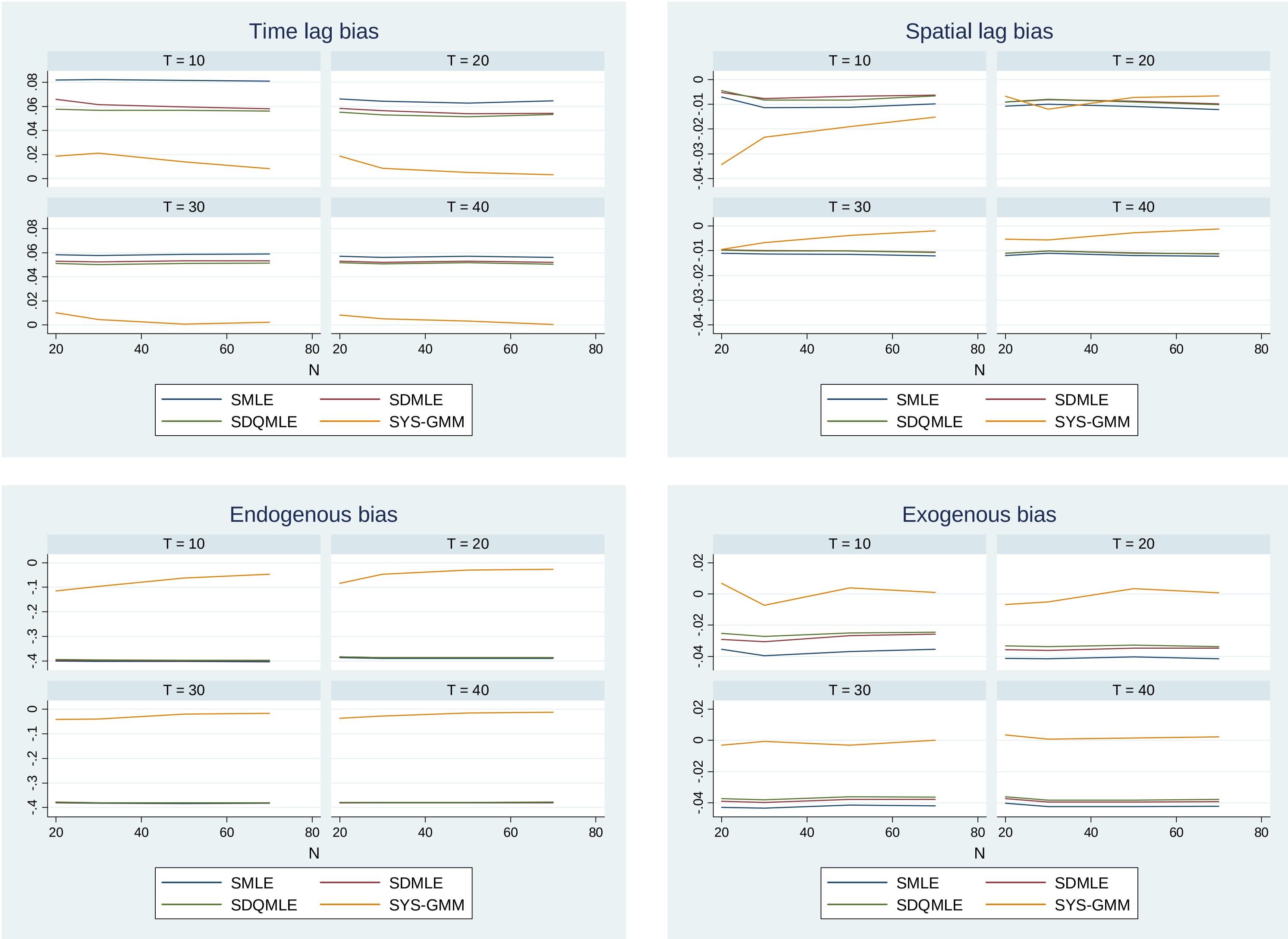

Figure 3: Bias Extended GMM vs Spatial Estimators

In terms of efficiency, summary of the results of RMSE (plotted in figure 4) are slighty different then the ones made in the analysis of the bias. Despite the fact that spatial ML estimators yield more bias, they tend to be more efficient than extended GMM in small sample for most parameters except the endogenous parameter. This is confirmed by figure 5, which plots the histogram distribution for each estimated parameter based on 1000 trials for $N = 70$ and $T = 10$. However, as the panel size increases, the performance of GMM converge to the spatial ML estimators. In some cases, it becomes even more efficient, this is the case for the time lag and exogenous variable. In fact, the comparison between system GMM and spatial maximum likelihood estimators suggests that in moderate samples, extended GMM might be more consistent and efficient than spatial ML estimators. This result might seems to be a paradox with respect to the accepted notion that ML is more efficient than GMM, but it is not. This finding is an extension to the panel framework of Das et al.'s (2003) conclusion that FG2SLS can outperform ML estimators in moderate cross-section samples The main reason for this contradictory result is that the spatial ML approach requires the estimation of more parameters than does extended GMM. ML involves the estimation of the additional parameter σ^2. Moreover, Elhorst's approach implies estimating additional parameters, while Yu et al's method is explicitly designed for large N and T. Therefore the classical arguments relating to relative efficiency do not apply here. In addition, as figures 3 and 4 show it, spatial estimators seem to reach efficiency for relatively small samples (making the slope of RMSE and bias almost flat), while extended GMM tends to gain efficiency in moderate sample (the slope of RMSE and bias being more steeper).

As noted in the bias analysis, the estimate of the endogenous covariate is clearly more efficient with extended GMM than any other maximum likelihood estimators. The rate of decrease of RMSE is almost null for the QMLE and MLE, which suggests that increasing the dimension sample cannot improve efficiency of the estimate of the endogenous variable. In other words, the use of spatial ML estimators is not recommended in the presence of endogenous or predetermined variable. Among the spatial estimators, SDQMLE is again the best one in terms of efficiency. Note that the estimation of the spatial lag by simple spatial MLE is as efficient as the other spatial estimators. This is in line with Elhorst's suggestion to estimate a dynamic panel model with exogenous variables by first estimate the spatial lag with SMLE and then estimate the remaining parameters using SDMLE or GMM (Elhorst (2008)).

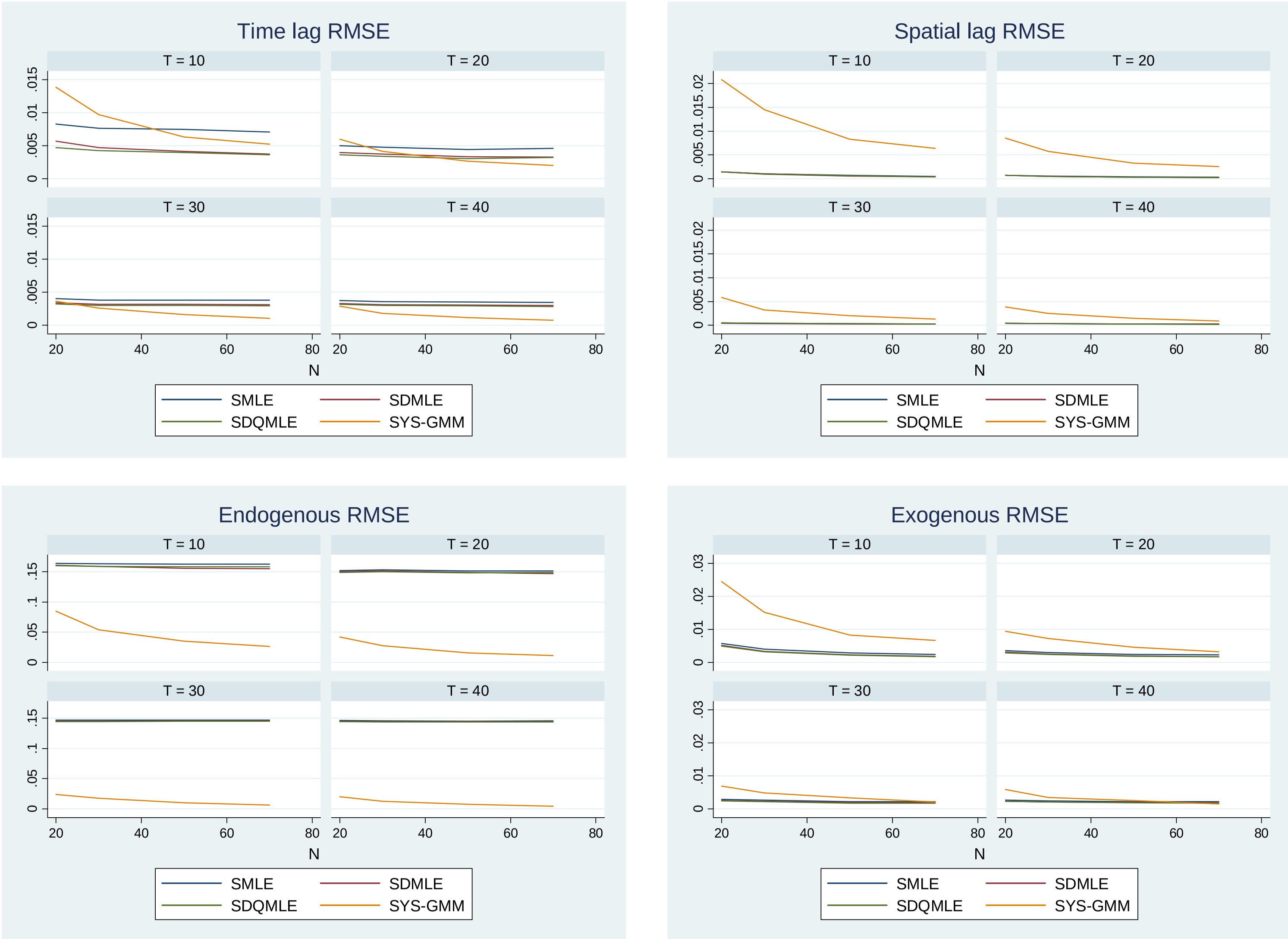

Figure 4: RMSE Extended GMM vs Spatial Estimators

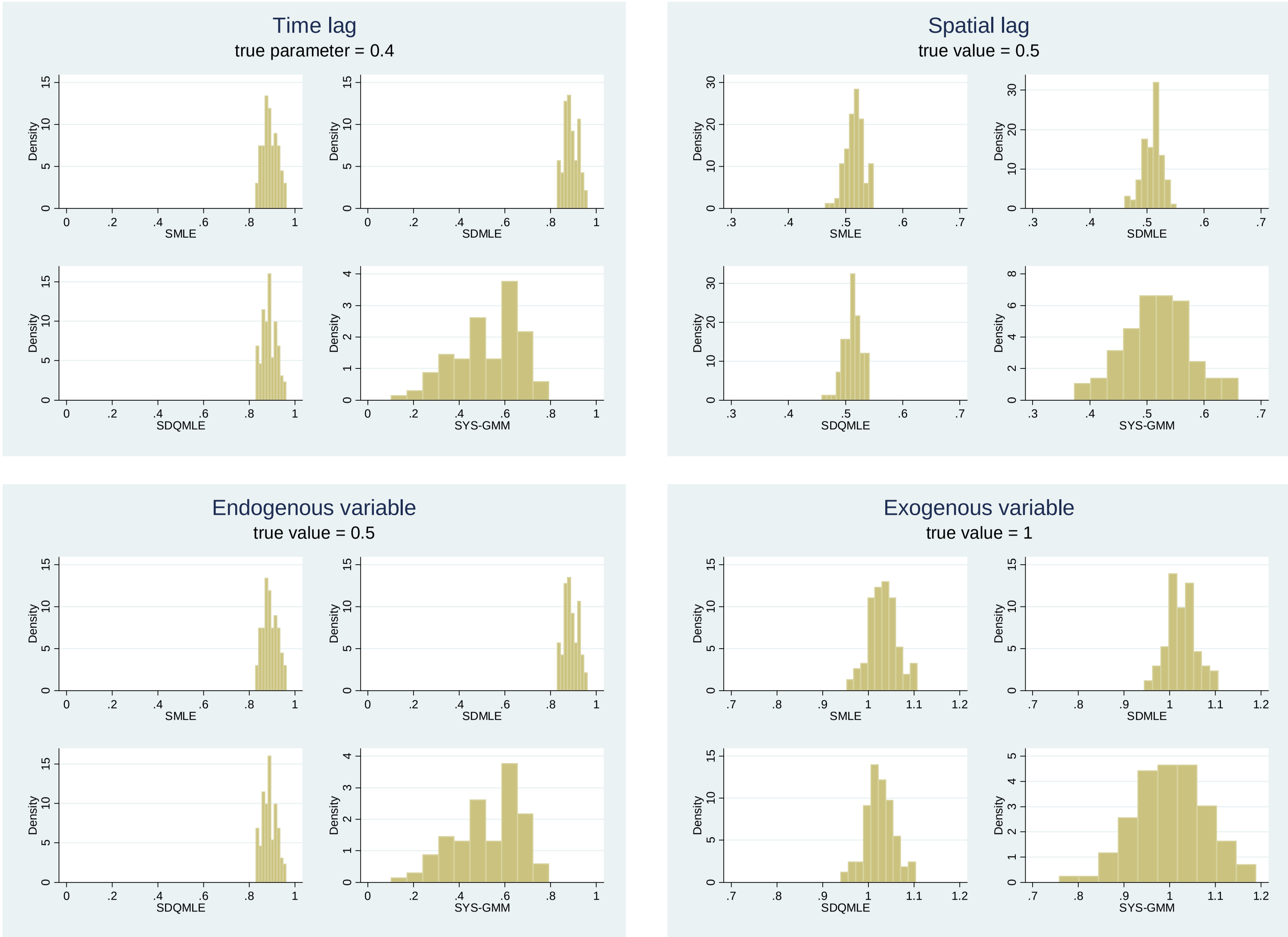

Figure 5: Parameters Histogram ($N = 70$ and $T = 10$)

3.3 RMSE response function

As previously commented, the relationship between the performance of the estimators and the model parameters is not easily determined. That is why, we report the results in terms of response functions. Using the RMSE for each estimator and parameters of the entire set of designs based on the normal distribution, the following equation is estimated by OLS:

$$\begin{aligned}\log\left(\sqrt{N \cdot T} \cdot RMSE_i\right) &= a_1 + a_2 \frac{1}{W_i} + a_3 \alpha_i + a_4 \rho_i + a_5 \left(\alpha_i \rho_i\right) + \\ & \quad a_6 \frac{W_i}{N_i} + a_7 \frac{W_i}{T_i} + a_8 \frac{1}{N_i} + a_9 \frac{1}{T_i} + \epsilon_i\end{aligned}$$

where $RMSE_i$ is the RMSE of a given parameter obtained using a given estimator in the i-th design. Note that W_i corresponds to the value attributed to the construction of the spatial weight matrix (e.g. $W_i \in (1, 3, 5)$).

The RMSE response function estimation results are displayed in table 2. Although some estimations are affected by multicollinearity, the fit of the response functions to the data is relatively good as confirmed by R^2. The comments made in the previous section are confirmed by the estimation results. As expected, the main factor that contributes to the efficiency of the GMM estimator is the panel size (N and T). This is not the case for the spatial estimators, because the rate of efficiency is marginally affected by the panel size (i.e. since the dependent variable is expressed in log, a negative coefficient implies a small effect). This was already confirmed by the figures 2 and 3, where the slope of the spatial estimators' RMSE were relatively flat. Concerning the value taken by the time and spatial lag, they only play a minor role individually, while their cross-product has a bigger impact on RMSE. This might be related to the time and spatial stationarity conditions. Another important finding relates to the specifiction of the spatial weight matrix W. The performances of the spatial estimators are definitively more sensitive to the spatial weight matrix than the extended GMM estimator. The estimation of the spatial lag by GMM is the only parameter which is directly affected by the spatial matrix This finding could suggest that spatial estimators are more sensitive to the specification of the spatial weight matrix than GMM.

Table 2: RMSE Response Functions

	Time lag: α			Spatial lag: ρ			Endogenous variable: γ			Exogenous variable: β		
	SYS-GMM	SDMLE	SDQMLE	SYS-GMM	SDMLE	SDQMLE	SYS-GMM	SDMLE	SDQMLE	SYS-GMM	SDMLE	SDQMLE
$1/W_i$	0.0251	-0.308***	-0.295***	-0.909***	-0.819***	-0.770***	-0.0352	-0.0138	-0.0146	-0.0396	-0.284***	-0.295***
	(-0.057)	(-0.064)	(-0.061)	(-0.081)	(-0.093)	(-0.092)	(-0.049)	(-0.018)	(-0.018)	(-0.063)	(-0.072)	(-0.072)
α_i	-0.229**	-3.520***	-3.821***	-2.887***	-1.858***	-1.800***	-1.002***	-0.195***	-0.187***	0.0625	-1.822***	-1.907***
	(-0.105)	(-0.122)	(-0.115)	(-0.143)	(-0.15)	(-0.151)	(-0.087)	(-0.032)	(-0.032)	(-0.109)	(-0.12)	(-0.121)
ρ_i	-0.736***	-3.072***	-3.131***	-2.615***	-2.156***	-2.012***	-0.500***	-0.157***	-0.159***	-0.441***	-1.911***	-1.899***
	(-0.116)	(-0.159)	(-0.153)	(-0.187)	(-0.205)	(-0.207)	(-0.1)	(-0.038)	(-0.039)	(-0.14)	(-0.147)	(-0.148)
$\alpha_i\rho_i$	1.437***	4.837***	5.041***	0.0916	5.945***	5.876***	0.348	0.146	0.152	0.786*	2.471***	2.380***
	(-0.384)	(-0.436)	(-0.406)	(-0.556)	(-0.646)	(-0.66)	(-0.317)	(-0.121)	(-0.123)	(-0.433)	(-0.461)	(-0.461)
W_i/N_i	0.706**	0.0427	-0.0181	1.454***	1.991***	2.055***	0.393	0.0724	0.0782	0.582	-0.155	-0.182
	(-0.357)	(-0.384)	(-0.374)	(-0.483)	(-0.484)	(-0.495)	(-0.299)	(-0.107)	(-0.11)	(-0.414)	(-0.369)	(-0.363)
W_i/T_i	-0.297*	0.0969	0.147	0.517**	1.366***	1.310***	-0.0852	0.0107	0.0049	-0.322*	-0.211	-0.206
	(-0.166)	(-0.188)	(-0.177)	(-0.24)	(-0.294)	(-0.297)	(-0.137)	(-0.048)	(-0.049)	(-0.186)	(-0.201)	(-0.195)
$1/N_i$	13.40***	-11.79***	-13.13***	13.28***	4.830***	3.885**	16.85***	-16.87***	-17.16***	15.56***	-1.461	-1.483
	(-1.317)	(-1.482)	(-1.41)	(-1.731)	(-1.657)	(-1.691)	(-1.079)	(-0.38)	(-0.391)	(-1.402)	(-1.46)	(-1.444)
$1/T_i$	13.22***	-3.730***	-4.408***	12.70***	1.976**	3.291***	11.04***	-7.673***	-7.513***	9.222***	-3.755***	-3.329***
	(-0.613)	(-0.723)	(-0.669)	(-0.865)	(-0.92)	(-0.939)	(-0.52)	(-0.166)	(-0.173)	(-0.647)	(-0.723)	(-0.711)
Constant	-3.409***	0.202***	0.302***	-0.637***	-3.258***	-3.345***	-1.191***	2.476***	2.474***	-2.762***	-1.206***	-1.228***
	(-0.057)	(-0.072)	(-0.068)	(-0.084)	(-0.09)	(-0.091)	(-0.049)	(-0.02)	(-0.019)	(-0.065)	(-0.072)	(-0.072)
R^2	0.824	0.824	0.856	0.893	0.797	0.791	0.85	0.962	0.961	0.698	0.612	0.618

The standard deviation is reported in parenthesis. ** $p<0.01$, ** $p<0.05$, * $p<0.1$

4 Robustness Check

In this section, departures from the model are investigated through 3 modifications. First, we check the robustness of the estimators with respect to a miss-specification of the spatial weight matrix. Second, we drop the assumption of gaussian process for the error terms and individual. Third, we investigate the consequence of dealing with spatially dependent endogenous and exogenous variables[14].

In order to study the consequences of the miss-specification of the spatial weight matrix, we estimate the spatial dynamic panel model using the data generated by the "3 ahead and 3 behind" W matrix, but assuming the spatial weight is "1 ahead and 1 behind" and "5 ahead and 5 behind". In the first case, we assume the spatial dependence to be more local than it really is. The opposite happens in the second scenario.

Overall and as expected, extended GMM seems to be relatively more robust in terms of bias to the miss-specification of the spatial weight matrix than the (overestimated) spatial estimators, with the exception of the spatial lag which converges at faster pace compared to the spatial estimators. The same cannot be said to the spatial ML estimators. In fact, it seems that most of the parameters estimated through spatial approaches are severely biased by the spatial miss-specification. The (underestimated) spatial coefficient remains the only parameter which is less affected. The main explanation for this finding lies in the fact that the spatial methods estimate the reduced form of the model, which means that the entire set of parameters, beside the spatial lag, are affected by the miss-specification. In GMM, only the instruments for the spatial lag are contaminated by the miss-specification of the spatial weight matrix. Once again, the simple SMLE is shown to estimate the spatial lag as accurate as SDMLE and SDQMLE (Elhorst (2008)).

Spatial MLE approaches rely on the assumption of normality of the individual effect and error term. The first non-gaussian error process considered is a Student distribution, which is characterized heavier tails than the normal distribution. The comments made earlier apply here also. More precisely, the performance of extended-GMM tends to improve, while there is some deteriorations for the spatial ML estimators. Among the latters, SDQMLE continues to displays the more robustness. Following Binder et al. (2005), we also investigate the way the individual effects are generated, since the

[14] To conserve space, we dont display the results tables, but they remain available upon request.

performance of extended-GMM depends on the ratio of the individual effect variance with respect to the variance of the error term (Hayakawa (2006)). Hence, the individual specific effects are non longer normally distributed but generated as follows:

$$\eta_i = \sqrt{\tau}\left(\frac{q_i - 1}{\sqrt{2}}\right)m_i, \qquad q_i \overset{iid}{\sim} \chi^2(1), \qquad m_i \overset{iid}{\sim} N(0, 0.05)$$

The parameter τ measures the degree of cross-section to the time-series variations. Two values are considered $\tau = 1$ and $\tau = 5$. As expected, spatial estimators are not affected by the way the individual specific effects are generated. The performance of system GMM deteriorates slightly, but continues to display better results for the endogenous variable.

Finally, we modify the data generating process to account for spatial dependence in the exogenous and endogenous variables. Each one follows a spatial moving average process:

$$\begin{aligned} EX_{it} &= \delta EX_{i,t-1} + \rho_{EX}\left[Wu_t\right]_i + u_{it} \\ EN_{it} &= \lambda EN_{i,t-1} + \psi\eta_i + \theta v_{it} + \rho_{EN}\left[We_t\right]_i + e_{it} \end{aligned}$$

The performance of the spatial estimators and GMM remain relatively robust to the presence of spatial dependence. Unlike in the baseline case, spatial ML estimators tend to underestimate the spatial lag parameter, while SYS-GMM continues to overestimate spatial lag parameter. Although not reported here, in the presence of positive spatial dependence in the exogenous and endogenous variables, extended GMM performs even better in terms of bias and efficiency, when the spatially weighted sum of the exogenous variable $W \cdot EX$ is included as additional instrument. In practice, this finding suggests to include spatial weighted sum of the exogenous variables, once the presence of spatial dependence is verified (through a Moran test for instance).

5 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 than any maximum likelihood estimators. In addition, the efficiency of the extended GMM could be improve through iterated GMM and continuously updated GMM (Hansen et al. (1996)).

Recently, a lot of attention has been drawn to the impact of heterogenous and cross-section error on the bias in dynamic panel estimation with fixed effects. One possible extension of this work would be to extended the spatial HAC framework proposed by Keleijan and Prucha (2007) or Driscoll and Kray's (1998) approach to the system GMM in a spatial dynamic panel model.

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] Beck, T., and R. Levine, 2004, "Stock markets, banks, and growth: Panel evidence", Journal of Banking and Finance, 28, 423–442.

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

[12] Binder, M., C. Hsiao and M H. Pesaran, 2005, "Estimation and Inference in Short Panel Vector Autoregressions with Unit Roots and Cointegration, Econometric Theory, 21, 795-837.

[13] 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.

[14] 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.

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

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

[17] Carkovic, M., and R. Levine, 2005, "Does Foreign Direct Investment Accelerate Economic Growth?" in T.H. Moran, E.M. Graham, and M. Blomström. Does foreign direct investment promote development? Washington, DC: Institute for International Economics and Center for Global Development.

[18] 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.

[19] Das, D., H. H. Kelejian and I. R. Prucha, 2003, "Finite sample properties of estimators of spatial autoregressive models with autoregressive disturbances", Journal of Economics, 82, 1-26.

[20] Driscoll, J. and A. Kraay, 1998, "Consistent Covariance Matrix Estimation With Spatially Dependent Panel Data", The Review of Economics and Statistics, 80, 549-560

[21] 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.

[22] 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

[23] 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

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

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

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

[27] 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.

[28] 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.

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

[30] 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.

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

[32] Hansen, L.P., J. Heaton and A. Yaron, 1996, "Finite-Sample Properties of Some Alternative GMM Estimators", Journal of Business and Economic Statistics, 14, 262-280.

[33] Hayakawa, K., 2007, "Small sample bias properties of the system GMM estimator in dynamic panel data models," Economics Letters, 95, 32-38.

[34] 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

[35] 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.

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

[37] 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.

[38] 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.

[39] Kelejian, H., and I. R. Prucha, 2007, "HAC estimation in a spatial framework", Journal of Econometrics, 140, 131-154.

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

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

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

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

[44] 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

[45] Lee L. and J. Ju, 2009, "Some Recent Developments in Spatial Panel Data Models", University of Kentucky, http://gatton.uky.edu/Faculty/yu/Research/Survey-SAR-Panel-0211-09.pdf

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

[47] Mutl J. and M. Pfaffermayr, 2008, "The Spatial Random Effects and the Spatial Fixed Effects Model: The Hausman Test in a Cliff and Ord Panel Model, Institute for Advanced Studies, http://www.ihs.ac.at/publications/eco/es-229.pdf

[48] Okui, R., 2008, "The Optimal Choice of Moments in Dynamic Panel Data Models", University of Pennsylvania, http://www.econ.upenn.edu/~okui/doc/okui_upenn_da.pdf

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

[50] Roodman, D., 2009, "A Note on the Theme of Too Many Instruments", Oxford Bulletin of Economics and Statistics, 71, 135-158, February 2009

[51] 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.

[52] 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

6 Appendices

6.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.

6.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 $\widetilde{\ }$.
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:
   $\begin{bmatrix} \widehat{\alpha} \\ \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} = b_0 - \widehat{\rho} b_L \qquad \text{and} \qquad \widehat{\sigma}^2 = \frac{1}{NT}\left(\widehat{U}_0 - \widehat{\rho}\widehat{U}_L\right)'\left(\widehat{U}_0 - \widehat{\rho}\widehat{U}_L\right).$

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

6.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 $\theta = \sigma^2_\xi/\sigma^2$ and σ^2_ξ 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
$$\underline{\Delta X} = \begin{bmatrix} \imath_N & \Delta X_1 & \cdots & \Delta X_T & 0 \\ 0 & 0 & \cdots & 0 & \Delta X_2 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \Delta X_T \end{bmatrix};$$

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

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

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

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

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

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 improves marginally the results.

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

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

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.

6.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 $\widehat{\sigma}^2$

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

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

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

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:

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

6.B GMM Estimators

This appendix section presents the procedure associated with the different GMM estimators.

. 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.

As mentioned previously, the time lag ($Y_{i,t-1}$), spatial lag ($[WY_t]_i$) and endogenous (EN_{it}) are treated as endogenous covariates, while the exogenous (EN_{it}) is considered as strictly exogenous. Each endogenous is instrumented by the strictly exogenous variables and the second and third lags of each endogenous variable. In order to restrict the number of instruments, the instruments matrix is constructed applying the "collapse" option[15].

6.B.1 Difference-GMM

The difference-GMM estimator, proposed by Arellano and Bond (1991), consists of estimating the model expressed in first-difference. More specifically, the estimation steps are:

1. Construct the "collapsed" instruments matrix for each cross-section i, :

$$Z_i^D = \begin{bmatrix} Y_{i,0} & 0 & [WY_0]_i & 0 & \Delta EX_{i,1} & EN_{i,0} & 0 \\ Y_{i,1} & Y_{i,0} & [WY_1]_i & [WY_0]_i & \Delta EX_{i,2} & EN_{i,1} & EN_{i,0} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_{i,T-2} & Y_{i,T-3} & [WY_{T-2}]_i & [WY_{T-3}]_i & \Delta EX_{i,T} & EN_{i,T-2} & EN_{i,T-3} \end{bmatrix}$$

2. Construct the weighting matrix:

$$A^{D1} = \left(\sum_i Z_i^{D\prime} \cdot H_i^{D1} \cdot Z_i^D \right)^{-1}$$

$$\text{where } H_i^{D1} = \begin{bmatrix} 1 & -0.5 & 0 & \cdots & 0 \\ -0.5 & 1 & -0.5 & \cdots & 0 \\ 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 & -0.5 \\ 0 & 0 & 0 & -0.5 & 1 \end{bmatrix}$$

3. Carry out the one-step estimation given by

$$\begin{bmatrix} \widehat{\alpha}_1 \\ \widehat{\rho}_1 \\ \widehat{\beta}_1 \\ \widehat{\gamma}_1 \end{bmatrix} = \left[Q_{XZ} \cdot A^{D1} \cdot Q'_{XZ} \right]^{-1} \cdot Q_{XZ} \cdot A^{D1} \cdot Q_{ZY}$$

where $Q_{XZ} = \sum_i X_i^{*\prime} \cdot Z_i$ and $Q_{ZX} = \sum_i Z_i' \cdot Y_i^*$

$$X_i^* = \begin{bmatrix} \Delta Y_{i,1} & [W\Delta Y_2]_i & \Delta EX_{i,2} & \Delta EN_{i,2} \\ \vdots & \vdots & \vdots & \vdots \\ \Delta Y_{i,T-1} & [W\Delta Y_T]_i & \Delta EX_{i,T} & \Delta EN_{i,T} \end{bmatrix} \text{ and } Y_i^* = \begin{bmatrix} \Delta Y_{i,2} \\ \vdots \\ \Delta Y_{i,T} \end{bmatrix}$$

[15] Numerous instruments lead to two types of small-sample issues. The first problem leads to overfitting endogenous variables, i.e. failure to remove endogeneity. The second problem concerns imprecise estimation of the optimal weighting matrix in the two-step procedure. This affects the computation of two-step standard errors and the validity of the Hansen's weak instruments (Roodman (2007)).

4. The associated variance are computed as follows:
$\widehat{V}_1 = \widehat{\sigma}_1^2 \cdot \left[Q_{XZ} \cdot A_i^{D1} \cdot Q'_{XZ}\right]^{-1}$
where $\widehat{\sigma}_1^2 = \frac{1}{N-4} \sum_i \left(Y_i^* - X_i^* \cdot \left[\widehat{\alpha}_1, \widehat{\rho}_1, \widehat{\beta}_1, \widehat{\gamma}_1\right]'\right)' \left(Y_i^* - X_i^* \cdot \left[\widehat{\alpha}_1, \widehat{\rho}_1, \widehat{\beta}_1, \widehat{\gamma}_1\right]'\right)$

5. The robust one-step variance is given by:
$\widehat{V}_{1,Robust} = \left[Q_{XZ} \cdot A^{D1} \cdot Q'_{XZ}\right]^{-1} \cdot A^{D1} \cdot \left(A^{D2}\right)^{-1} \cdot A^{D1} \cdot Q'_{XZ} \cdot \left[Q_{XZ} \cdot A^{D1} \cdot Q'_{XZ}\right]^{-1}$
where $A^{D2} = \left(\sum_i Z_i^{D\prime} \cdot H_i^{D2} \cdot Z_i^D\right)^{-1}$
$H_i^{D2} = \left(Y_i^* - X_i^* \cdot \left[\widehat{\alpha}, \widehat{\rho}, \widehat{\beta}, \widehat{\gamma}\right]'\right) \left(Y_i^* - X_i^* \cdot \left[\widehat{\alpha}, \widehat{\rho}, \widehat{\beta}, \widehat{\gamma}\right]'\right)'$

6. The two-step estimates are given by
$$\begin{bmatrix} \widehat{\alpha}_2 \\ \widehat{\rho}_2 \\ \widehat{\beta}_2 \\ \widehat{\gamma}_2 \end{bmatrix} = \left[Q_{XZ} \cdot A^{D2} \cdot Q'_{XZ}\right]^{-1} \cdot Q_{XZ} \cdot A^{D2} \cdot Q_{ZY}$$

7. The associated two-step variance is computed as
$\widehat{V}_2 = \left[Q_{XZ} \cdot A^{D2} \cdot Q'_{XZ}\right]^{-1}$

6.B.2 Extended-GMM

The system-GMM estimator, proposed by Arellano and Bover (1995) and Blundell and Bond (1998), consists of combining the moment conditions from the model in first-difference with the moment conditions from the model in levels. These are the estimation steps:

1. Construct the "collapsed" instruments matrix for each cross-section i, :
$$Z_i = \begin{bmatrix} Y_{i,0} & 0 & 0 & [WY_0]_i & 0 & 0 & \Delta EX_{i,1} & EN_{i,0} & 0 & 0 \\ Y_{i,1} & Y_{i,0} & 0 & [WY_1]_i & [WY_0]_i & 0 & \Delta EX_{i,2} & EN_{i,1} & EN_{i,0} & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ Y_{i,T-2} & Y_{i,T-3} & 0 & [WY_{T-2}]_i & [WY_{T-3}]_i & 0 & \Delta EX_{i,T} & EN_{i,T-2} & EN_{i,T-3} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & EX_{i1} & 0 & 0 & 0 \\ 0 & 0 & \Delta Y_{i,1} & 0 & 0 & [W\Delta Y_2]_i & EX_{i2} & 0 & 0 & \Delta EN_{i2} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \Delta Y_{i,T-1} & 0 & 0 & [W\Delta Y_{T-1}]_i & EX_{iT} & 0 & 0 & \Delta EN_{i,T-1} \end{bmatrix}$$

2. Construct the weighting matrix:

$$A^1 = \left(\sum_i Z_i' \cdot H_i \cdot Z_i \right)^{-1}$$

where $H_i = \begin{bmatrix} H_i^D & 0 \\ 0 & I_i \end{bmatrix}$

3. Carry out the one-step estimation given by

$$\begin{bmatrix} \widehat{\alpha}_1 \\ \widehat{\rho}_1 \\ \widehat{\beta}_1 \\ \widehat{\gamma}_1 \end{bmatrix} = [Q_{XZ} \cdot A^1 \cdot Q_{XZ}']^{-1} \cdot Q_{XZ} \cdot A^1 \cdot Q_{ZY}$$

where $Q_{XZ} = \sum_i X_i^{*\prime} \cdot Z_i$ and $Q_{ZX} = \sum_i Z_i' \cdot Y_i^*$

$$X_i^* = \begin{bmatrix} \Delta Y_{i,1} & [W\Delta Y_2]_i & \Delta EX_{i,2} & \Delta EN_{i,2} \\ \vdots & \vdots & \vdots & \vdots \\ \Delta Y_{i,T-1} & [W\Delta Y_T]_i & \Delta EX_{i,T} & \Delta EN_{i,T} \\ Y_{i,0} & [WY_1]_i & EX_{i,1} & EN_{i,1} \\ \vdots & \vdots & \vdots & \vdots \\ Y_{i,T-1} & [WY_T]_i & EX_{i,T} & EN_{i,T} \end{bmatrix} \text{ and } Y_i^* = \begin{bmatrix} \Delta Y_{i,2} \\ \vdots \\ \Delta Y_{i,T} \\ Y_{i,1} \\ \vdots \\ Y_{i,T} \end{bmatrix}$$

4. The associated variance are computed as follows:

$$\widehat{V}_1 = \widehat{\sigma}_1^2 \cdot [Q_{XZ} \cdot A_i^1 \cdot Q_{XZ}']^{-1}$$

where $\widehat{\sigma}_1^2 = \frac{1}{N-4} \sum_i \left(Y_i^* - X_i^* \cdot \left[\widehat{\alpha}_1, \widehat{\rho}_1, \widehat{\beta}_1, \widehat{\gamma}_1 \right]' \right)' \left(Y_i^* - X_i^* \cdot \left[\widehat{\alpha}_1, \widehat{\rho}_1, \widehat{\beta}_1, \widehat{\gamma}_1 \right]' \right)$

5. The robust one-step variance is given by:

$$\widehat{V}_{1,Robust} = [Q_{XZ} \cdot A^1 \cdot Q_{XZ}']^{-1} \cdot A^1 \cdot (A^2)^{-1} \cdot A^1 \cdot Q_{XZ}' \cdot [Q_{XZ} \cdot A^1 \cdot Q_{XZ}']^{-1}$$

where $A^2 = \left(\sum_i Z_i' \cdot H_i^2 \cdot Z_i \right)^{-1}$

$$H_i^2 = \left(Y_i^* - X_i^* \cdot \left[\widehat{\alpha}, \widehat{\rho}, \widehat{\beta}, \widehat{\gamma} \right]' \right) \left(Y_i^* - X_i^* \cdot \left[\widehat{\alpha}, \widehat{\rho}, \widehat{\beta}, \widehat{\gamma} \right]' \right)'$$

6. The two-step estimates are given by

$$\begin{bmatrix} \widehat{\alpha} \\ \widehat{\rho} \\ \widehat{\beta} \\ \widehat{\gamma} \end{bmatrix} = [Q_{XZ} \cdot A^2 \cdot Q_{XZ}']^{-1} \cdot Q_{XZ} \cdot A^2 \cdot Q_{ZY}$$

7. The associated two-step variance is computed as

$$\widehat{V}_2 = [Q_{XZ} \cdot A^2 \cdot Q_{XZ}']^{-1}$$

6.C Monte Carlo Results: Bias

Time lag variable α: Bias

T	N	α	ρ	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.2	0.1	0.1011	0.0797	0.0752	0.0909	0.0467	0.0056
10	20	0.5	0.3	0.0798	0.0644	0.0526	0.0623	0.0403	0.0156
10	20	0.7	0.1	0.0683	0.0527	0.0372	0.0427	0.0544	-0.0541
20	20	0.2	0.1	0.0906	0.0793	0.0776	0.0414	0.0412	0.0218
20	20	0.5	0.3	0.0588	0.0525	0.0479	0.0619	0.0285	0.0265
20	20	0.7	0.1	0.0514	0.0445	0.0394	0.0238	0.0148	0.0039
30	20	0.2	0.1	0.0847	0.0775	0.0761	-0.0576	0.0142	0.0141
30	20	0.5	0.3	0.0529	0.0477	0.0448	0.0603	0.0144	0.0033
30	20	0.7	0.1	0.0409	0.0360	0.0326	-0.0120	0.0007	-0.0023
40	20	0.2	0.1	0.0824	0.0767	0.0761	0.0802	0.0143	0.0067
40	20	0.5	0.3	0.0519	0.0486	0.0464	-0.0635	0.0125	0.0058
40	20	0.7	0.1	0.0379	0.0347	0.0325	-0.0077	0.0129	0.0058
10	30	0.2	0.1	0.1123	0.0867	0.0846	0.1076	0.0365	0.0388
10	30	0.5	0.3	0.0803	0.0572	0.0516	-0.0050	0.0363	-0.0055
10	30	0.7	0.1	0.0681	0.0419	0.0342	0.0657	0.0359	-0.0145
20	30	0.2	0.1	0.0895	0.0786	0.0767	0.0897	0.0156	0.0140
20	30	0.5	0.3	0.0597	0.0524	0.0487	0.0685	0.0127	-0.0033
20	30	0.7	0.1	0.0475	0.0416	0.0353	0.0086	0.0146	-0.0146
30	30	0.2	0.1	0.0796	0.0724	0.0712	-0.0533	0.0125	0.0090
30	30	0.5	0.3	0.0520	0.0473	0.0442	0.0379	0.0157	0.0022
30	30	0.7	0.1	0.0387	0.0344	0.0315	-0.0141	0.0039	-0.0117
40	30	0.2	0.1	0.0805	0.0746	0.0738	0.0635	0.0049	0.0038
40	30	0.5	0.3	0.0492	0.0458	0.0437	0.0259	0.0064	0.0070
40	30	0.7	0.1	0.0371	0.0336	0.0316	0.0106	-0.0013	-0.0107
10	50	0.2	0.1	0.1104	0.0867	0.0851	0.1110	0.0299	0.0147
10	50	0.5	0.3	0.0765	0.0531	0.0492	0.0630	0.0175	-0.0048
10	50	0.7	0.1	0.0696	0.0426	0.0372	0.0598	0.0228	-0.0127
20	50	0.2	0.1	0.0874	0.0760	0.0743	0.0820	0.0061	0.0034
20	50	0.5	0.3	0.0553	0.0470	0.0438	0.0669	0.0110	0.0174
20	50	0.7	0.1	0.0472	0.0361	0.0330	0.0460	0.0212	-0.0107
30	50	0.2	0.1	0.0860	0.0786	0.0774	0.0500	-0.0016	0.0012
30	50	0.5	0.3	0.0533	0.0482	0.0454	0.0367	0.0057	-0.0033
30	50	0.7	0.1	0.0382	0.0340	0.0305	0.0176	-0.0043	-0.0131
40	50	0.2	0.1	0.0831	0.0776	0.0767	0.0597	0.0025	0.0026
40	50	0.5	0.3	0.0517	0.0484	0.0464	0.0611	0.0045	-0.0005
40	50	0.7	0.1	0.0356	0.0323	0.0310	-0.0452	0.0092	-0.0043
10	70	0.2	0.1	0.1047	0.0799	0.0793	0.0964	0.0211	0.0191
10	70	0.5	0.3	0.0791	0.0546	0.0523	0.0907	0.0186	0.0005
10	70	0.7	0.1	0.0716	0.0448	0.0392	0.0396	0.0201	-0.0107
20	70	0.2	0.1	0.0913	0.0790	0.0783	0.0918	0.0069	0.0057
20	70	0.5	0.3	0.0565	0.0455	0.0445	0.0631	0.0083	-0.0015
20	70	0.7	0.1	0.0478	0.0359	0.0348	0.0357	0.0101	-0.0070
30	70	0.2	0.1	0.0855	0.0780	0.0769	0.0720	0.0042	0.0069
30	70	0.5	0.3	0.0511	0.0456	0.0435	0.0248	0.0031	0.0001
30	70	0.7	0.1	0.0388	0.0333	0.0309	0.0195	0.0074	0.0015
40	70	0.2	0.1	0.0822	0.0764	0.0753	0.0613	-0.0006	0.0014
40	70	0.5	0.3	0.0498	0.0462	0.0442	0.0608	0.0024	-0.0003
40	70	0.7	0.1	0.0359	0.0330	0.0303	0.0319	-0.0009	-0.0102

Spatial lag variable ρ: Bias

T	N	α	ρ	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.2	0.1	-0.0020	-0.0010	-0.0021	-0.0781	-0.0111	-0.0558
10	20	0.5	0.3	-0.0124	-0.0092	-0.0097	-0.0793	0.0017	0.0067
10	20	0.7	0.1	-0.0082	-0.0094	-0.0088	-0.0800	-0.0697	-0.0153
20	20	0.2	0.1	-0.0079	-0.0075	-0.0074	-0.0418	-0.0256	0.0268
20	20	0.5	0.3	-0.0150	-0.0129	-0.0119	-0.0476	-0.0164	-0.0172
20	20	0.7	0.1	-0.0085	-0.0083	-0.0072	-0.0674	-0.0135	-0.0012
30	20	0.2	0.1	0.0000	0.0005	0.0000	-0.1466	-0.0161	-0.0237
30	20	0.5	0.3	-0.0148	-0.0140	-0.0127	-0.0199	-0.0062	-0.0023
30	20	0.7	0.1	-0.0069	-0.0060	-0.0068	0.0747	0.0122	0.0114
40	20	0.2	0.1	-0.0067	-0.0065	-0.0063	-0.0250	-0.0021	0.0052
40	20	0.5	0.3	-0.0150	-0.0137	-0.0142	-0.0148	0.0023	-0.0073
40	20	0.7	0.1	-0.0049	-0.0040	-0.0041	-0.0113	0.0037	0.0008
10	30	0.2	0.1	-0.0062	-0.0047	-0.0041	-0.0229	-0.0530	-0.0369
10	30	0.5	0.3	-0.0172	-0.0123	-0.0136	-0.0484	-0.0132	-0.0206
10	30	0.7	0.1	-0.0094	-0.0080	-0.0105	-0.0111	0.0141	0.0120
20	30	0.2	0.1	-0.0050	-0.0038	-0.0041	-0.0312	0.0045	-0.0070
20	30	0.5	0.3	-0.0178	-0.0152	-0.0142	-0.0515	-0.0022	-0.0130
20	30	0.7	0.1	-0.0022	-0.0011	-0.0017	0.0075	0.0161	-0.0025
30	30	0.2	0.1	-0.0049	-0.0046	-0.0047	0.0958	0.0050	0.0138
30	30	0.5	0.3	-0.0167	-0.0159	-0.0149	0.0040	0.0016	-0.0001
30	30	0.7	0.1	-0.0071	-0.0055	-0.0062	0.0065	0.0012	-0.0013
40	30	0.2	0.1	0.0006	0.0009	0.0008	0.0120	-0.0165	-0.0156
40	30	0.5	0.3	-0.0162	-0.0149	-0.0145	-0.0198	-0.0076	-0.0061
40	30	0.7	0.1	-0.0075	-0.0072	-0.0070	0.0049	0.0023	0.0037
10	50	0.2	0.1	-0.0068	-0.0009	-0.0064	-0.0297	-0.0189	-0.0032
10	50	0.5	0.3	-0.0134	-0.0074	-0.0082	-0.1175	0.0050	-0.0258
10	50	0.7	0.1	-0.0061	-0.0047	-0.0065	-0.0289	-0.0068	-0.0052
20	50	0.2	0.1	0.0018	0.0021	0.0026	-0.0134	0.0061	0.0048
20	50	0.5	0.3	-0.0147	-0.0124	-0.0121	-0.0481	-0.0081	-0.0065
20	50	0.7	0.1	-0.0080	-0.0069	-0.0072	-0.0226	0.0050	0.0003
30	50	0.2	0.1	-0.0039	-0.0035	-0.0033	-0.0529	0.0038	-0.0051
30	50	0.5	0.3	-0.0175	-0.0153	-0.0148	0.0022	-0.0061	-0.0025
30	50	0.7	0.1	-0.0058	-0.0050	-0.0051	0.0068	-0.0007	0.0018
40	50	0.2	0.1	-0.0050	-0.0048	-0.0048	-0.0463	-0.0022	0.0037
40	50	0.5	0.3	-0.0188	-0.0178	-0.0174	-0.0541	0.0063	0.0000
40	50	0.7	0.1	-0.0080	-0.0075	-0.0087	-0.0201	-0.0020	-0.0030
10	70	0.2	0.1	-0.0029	-0.0009	-0.0019	-0.0290	0.0167	-0.0031
10	70	0.5	0.3	-0.0142	-0.0091	-0.0091	-0.0537	0.0029	-0.0087
10	70	0.7	0.1	-0.0059	-0.0045	-0.0041	0.0113	0.0162	-0.0003
20	70	0.2	0.1	-0.0049	-0.0045	-0.0038	-0.0234	-0.0111	-0.0105
20	70	0.5	0.3	-0.0172	-0.0136	-0.0143	-0.0441	-0.0050	-0.0030
20	70	0.7	0.1	-0.0092	-0.0072	-0.0082	-0.0055	-0.0109	-0.0060
30	70	0.2	0.1	-0.0035	-0.0029	-0.0030	-0.0124	0.0100	0.0135
30	70	0.5	0.3	-0.0166	-0.0141	-0.0144	-0.0154	-0.0032	-0.0002
30	70	0.7	0.1	-0.0058	-0.0047	-0.0052	0.0151	0.0022	-0.0012
40	70	0.2	0.1	-0.0039	-0.0036	-0.0035	-0.0293	-0.0025	0.0039
40	70	0.5	0.3	-0.0173	-0.0168	-0.0164	-0.0568	0.0058	0.0012
40	70	0.7	0.1	-0.0085	-0.0080	-0.0083	-0.0097	0.0006	0.0031

Endogenous variable γ: Bia

T	N	α	ρ	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.2	0.1	-0.4157	-0.4092	-0.4064	-0.4134	-0.0990	-0.1114
10	20	0.5	0.3	-0.3865	-0.3837	-0.3824	-0.3956	-0.1219	-0.1097
10	20	0.7	0.1	-0.3961	-0.3909	-0.3906	-0.4014	-0.1193	-0.0940
20	20	0.2	0.1	-0.3875	-0.3834	-0.3829	-0.4024	-0.0792	-0.1058
20	20	0.5	0.3	-0.3725	-0.3708	-0.3689	-0.3758	-0.0472	-0.0596
20	20	0.7	0.1	-0.3745	-0.3746	-0.3744	-0.3785	-0.0393	-0.0546
30	20	0.2	0.1	-0.3863	-0.3839	-0.3833	-0.4237	-0.0218	-0.0242
30	20	0.5	0.3	-0.3753	-0.3726	-0.3725	-0.3808	-0.0343	-0.0395
30	20	0.7	0.1	-0.3740	-0.3721	-0.3708	-0.4020	-0.0128	-0.0270
40	20	0.2	0.1	-0.3910	-0.3893	-0.3887	-0.3932	-0.0417	-0.0413
40	20	0.5	0.3	-0.3797	-0.3782	-0.3769	-0.4265	-0.0305	-0.0401
40	20	0.7	0.1	-0.3708	-0.3693	-0.3684	-0.3847	-0.0373	-0.0381
10	30	0.2	0.1	-0.4209	-0.4157	-0.4143	-0.4201	-0.1160	-0.1249
10	30	0.5	0.3	-0.3985	-0.3955	-0.3937	-0.4453	-0.0566	-0.0617
10	30	0.7	0.1	-0.4048	-0.4053	-0.4023	-0.4070	-0.0631	-0.0596
20	30	0.2	0.1	-0.3951	-0.3910	-0.3905	-0.3937	-0.0528	-0.0483
20	30	0.5	0.3	-0.3880	-0.3857	-0.3844	-0.3866	-0.0466	-0.0430
20	30	0.7	0.1	-0.3776	-0.3748	-0.3738	-0.3836	-0.0272	-0.0283
30	30	0.2	0.1	-0.3916	-0.3888	-0.3882	-0.4438	-0.0318	-0.0404
30	30	0.5	0.3	-0.3830	-0.3818	-0.3807	-0.4072	-0.0437	-0.0466
30	30	0.7	0.1	-0.3748	-0.3736	-0.3709	-0.3924	-0.0255	-0.0239
40	30	0.2	0.1	-0.3954	-0.3935	-0.3931	-0.4020	-0.0052	-0.0207
40	30	0.5	0.3	-0.3779	-0.3765	-0.3754	-0.3910	-0.0359	-0.0454
40	30	0.7	0.1	-0.3762	-0.3749	-0.3740	-0.3870	-0.0091	-0.0211
10	50	0.2	0.1	-0.4093	-0.4024	-0.4017	-0.4082	-0.0693	-0.0768
10	50	0.5	0.3	-0.3958	-0.3939	-0.3907	-0.3911	-0.0508	-0.0738
10	50	0.7	0.1	-0.4002	-0.3970	-0.3948	-0.4019	-0.0559	-0.0344
20	50	0.2	0.1	-0.4026	-0.3993	-0.3991	-0.4027	0.0005	-0.0158
20	50	0.5	0.3	-0.3782	-0.3754	-0.3742	-0.3768	-0.0434	-0.0479
20	50	0.7	0.1	-0.3798	-0.3755	-0.3759	-0.3797	-0.0163	-0.0053
30	50	0.2	0.1	-0.3933	-0.3907	-0.3903	-0.4018	-0.0257	-0.0352
30	50	0.5	0.3	-0.3801	-0.3784	-0.3775	-0.3970	-0.0061	-0.0092
30	50	0.7	0.1	-0.3758	-0.3738	-0.3725	-0.3859	0.0044	-0.0070
40	50	0.2	0.1	-0.3937	-0.3916	-0.3912	-0.3989	-0.0038	-0.0136
40	50	0.5	0.3	-0.3782	-0.3767	-0.3759	-0.3776	-0.0164	-0.0086
40	50	0.7	0.1	-0.3710	-0.3698	-0.3680	-0.3963	-0.0076	-0.0018
10	70	0.2	0.1	-0.4112	-0.4053	-0.4045	-0.4133	-0.0305	-0.0313
10	70	0.5	0.3	-0.4014	-0.3952	-0.3951	-0.3974	-0.0437	-0.0309
10	70	0.7	0.1	-0.3939	-0.3899	-0.3890	-0.4030	-0.0686	-0.0533
20	70	0.2	0.1	-0.3976	-0.3934	-0.3931	-0.3970	-0.0233	-0.0379
20	70	0.5	0.3	-0.3844	-0.3805	-0.3803	-0.3853	-0.0175	-0.0114
20	70	0.7	0.1	-0.3828	-0.3799	-0.3793	-0.3879	-0.0221	-0.0279
30	70	0.2	0.1	-0.3913	-0.3887	-0.3883	-0.3956	-0.0207	-0.0259
30	70	0.5	0.3	-0.3805	-0.3782	-0.3773	-0.4009	-0.0174	-0.0194
30	70	0.7	0.1	-0.3787	-0.3767	-0.3756	-0.3853	-0.0127	-0.0144
40	70	0.2	0.1	-0.3898	-0.3877	-0.3874	-0.3965	0.0123	0.0079
40	70	0.5	0.3	-0.3801	-0.3787	-0.3779	-0.3783	0.0005	0.0019
40	70	0.7	0.1	-0.3761	-0.3747	-0.3735	-0.3773	-0.0015	0.0061

Exogenous variable β: Bias

T	N	α	ρ	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.2	0.1	-0.0454	-0.0334	-0.0298	-0.0308	0.0256	-0.0259
10	20	0.5	0.3	-0.0246	-0.0194	-0.0129	-0.0003	-0.0247	0.0001
10	20	0.7	0.1	-0.0347	-0.0325	-0.0255	-0.0261	0.0247	0.0523
20	20	0.2	0.1	-0.0517	-0.0451	-0.0440	-0.0172	-0.0091	-0.0151
20	20	0.5	0.3	-0.0397	-0.0351	-0.0314	-0.0377	0.0021	-0.0269
20	20	0.7	0.1	-0.0372	-0.0312	-0.0287	-0.0203	-0.0106	0.0200
30	20	0.2	0.1	-0.0614	-0.0567	-0.0557	0.0481	-0.0059	-0.0140
30	20	0.5	0.3	-0.0408	-0.0364	-0.0346	-0.0438	0.0042	-0.0094
30	20	0.7	0.1	-0.0449	-0.0386	-0.0381	-0.0032	-0.0017	-0.0032
40	20	0.2	0.1	-0.0565	-0.0524	-0.0518	-0.0550	-0.0049	0.0110
40	20	0.5	0.3	-0.0391	-0.0366	-0.0346	0.0622	0.0036	-0.0058
40	20	0.7	0.1	-0.0376	-0.0353	-0.0331	0.0065	0.0015	0.0003
10	30	0.2	0.1	-0.0528	-0.0376	-0.0367	-0.0492	0.0190	-0.0308
10	30	0.5	0.3	-0.0388	-0.0305	-0.0284	0.0196	0.0149	0.0026
10	30	0.7	0.1	-0.0298	-0.0225	-0.0149	-0.0288	0.0096	0.0120
20	30	0.2	0.1	-0.0560	-0.0493	-0.0486	-0.0560	0.0030	-0.0078
20	30	0.5	0.3	-0.0361	-0.0292	-0.0267	-0.0350	0.0057	0.0007
20	30	0.7	0.1	-0.0327	-0.0273	-0.0226	-0.0084	0.0016	0.0154
30	30	0.2	0.1	-0.0596	-0.0548	-0.0539	0.0241	-0.0074	-0.0044
30	30	0.5	0.3	-0.0385	-0.0348	-0.0326	-0.0271	0.0011	0.0107
30	30	0.7	0.1	-0.0329	-0.0286	-0.0260	0.0070	0.0038	0.0092
40	30	0.2	0.1	-0.0595	-0.0560	-0.0555	-0.0456	-0.0030	-0.0033
40	30	0.5	0.3	-0.0398	-0.0367	-0.0349	-0.0151	0.0015	-0.0013
40	30	0.7	0.1	-0.0369	-0.0335	-0.0314	-0.0117	-0.0025	0.0018
10	50	0.2	0.1	-0.0447	-0.0321	-0.0309	-0.0453	-0.0002	-0.0102
10	50	0.5	0.3	-0.0410	-0.0288	-0.0277	-0.0023	0.0033	0.0188
10	50	0.7	0.1	-0.0277	-0.0185	-0.0159	-0.0269	-0.0064	0.0211
20	50	0.2	0.1	-0.0484	-0.0412	-0.0403	-0.0474	0.0036	-0.0012
20	50	0.5	0.3	-0.0427	-0.0373	-0.0349	-0.0425	-0.0121	-0.0224
20	50	0.7	0.1	-0.0280	-0.0210	-0.0185	-0.0260	0.0024	0.0229
30	50	0.2	0.1	-0.0545	-0.0488	-0.0479	-0.0233	0.0007	-0.0125
30	50	0.5	0.3	-0.0390	-0.0356	-0.0335	-0.0277	0.0015	-0.0032
30	50	0.7	0.1	-0.0318	-0.0282	-0.0254	-0.0163	0.0011	0.0031
40	50	0.2	0.1	-0.0591	-0.0553	-0.0547	-0.0369	0.0005	0.0007
40	50	0.5	0.3	-0.0393	-0.0365	-0.0348	-0.0378	-0.0007	0.0054
40	50	0.7	0.1	-0.0337	-0.0305	-0.0284	0.0463	0.0012	0.0101
10	70	0.2	0.1	-0.0536	-0.0410	-0.0404	-0.0483	0.0109	0.0038
10	70	0.5	0.3	-0.0319	-0.0227	-0.0203	-0.0266	-0.0053	0.0158
10	70	0.7	0.1	-0.0273	-0.0168	-0.0132	-0.0114	0.0132	0.0166
20	70	0.2	0.1	-0.0533	-0.0456	-0.0450	-0.0540	-0.0038	-0.0010
20	70	0.5	0.3	-0.0402	-0.0329	-0.0326	-0.0344	-0.0048	0.0078
20	70	0.7	0.1	-0.0281	-0.0197	-0.0192	-0.0187	0.0081	0.0073
30	70	0.2	0.1	-0.0582	-0.0531	-0.0523	-0.0484	-0.0015	-0.0037
30	70	0.5	0.3	-0.0380	-0.0342	-0.0325	-0.0133	-0.0026	0.0032
30	70	0.7	0.1	-0.0324	-0.0272	-0.0255	-0.0184	0.0005	-0.0013
40	70	0.2	0.1	-0.0593	-0.0558	-0.0552	-0.0412	-0.0066	-0.0012
40	70	0.5	0.3	-0.0373	-0.0346	-0.0328	-0.0361	-0.0032	0.0017
40	70	0.7	0.1	-0.0309	-0.0281	-0.0253	-0.0279	0.0057	0.0144

6.D Monte Carlo Results: RMSE

Time lag variable α: RMSE

T	N	α	ρ	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.2	0.1	0.0120	0.0080	0.0073	0.0098	0.0244	0.0178
10	20	0.5	0.3	0.0073	0.0051	0.0038	0.0049	0.0143	0.0116
10	20	0.7	0.1	0.0060	0.0039	0.0026	0.0033	0.0818	0.0172
20	20	0.2	0.1	0.0085	0.0066	0.0063	0.0023	0.0064	0.0075
20	20	0.5	0.3	0.0041	0.0032	0.0028	0.0045	0.0061	0.0071
20	20	0.7	0.1	0.0030	0.0023	0.0018	0.0011	0.0098	0.0038
30	20	0.2	0.1	0.0079	0.0067	0.0065	0.0039	0.0023	0.0036
30	20	0.5	0.3	0.0030	0.0025	0.0023	0.0039	0.0021	0.0025
30	20	0.7	0.1	0.0020	0.0016	0.0014	0.0004	0.0032	0.0042
40	20	0.2	0.1	0.0074	0.0065	0.0063	0.0071	0.0018	0.0026
40	20	0.5	0.3	0.0028	0.0024	0.0023	0.0046	0.0018	0.0030
40	20	0.7	0.1	0.0016	0.0013	0.0012	0.0002	0.0016	0.0020
10	30	0.2	0.1	0.0133	0.0088	0.0083	0.0130	0.0108	0.0097
10	30	0.5	0.3	0.0069	0.0039	0.0034	0.0009	0.0199	0.0132
10	30	0.7	0.1	0.0052	0.0023	0.0018	0.0047	0.0209	0.0090
20	30	0.2	0.1	0.0087	0.0068	0.0065	0.0087	0.0032	0.0041
20	30	0.5	0.3	0.0036	0.0028	0.0024	0.0049	0.0039	0.0043
20	30	0.7	0.1	0.0026	0.0020	0.0015	0.0003	0.0051	0.0037
30	30	0.2	0.1	0.0070	0.0059	0.0057	0.0031	0.0016	0.0031
30	30	0.5	0.3	0.0030	0.0025	0.0022	0.0015	0.0019	0.0027
30	30	0.7	0.1	0.0017	0.0014	0.0012	0.0003	0.0019	0.0021
40	30	0.2	0.1	0.0068	0.0059	0.0058	0.0044	0.0011	0.0015
40	30	0.5	0.3	0.0025	0.0021	0.0020	0.0008	0.0010	0.0017
40	30	0.7	0.1	0.0014	0.0012	0.0011	0.0002	0.0014	0.0018
10	50	0.2	0.1	0.0126	0.0077	0.0075	0.0125	0.0056	0.0068
10	50	0.5	0.3	0.0059	0.0028	0.0026	0.0040	0.0157	0.0076
10	50	0.7	0.1	0.0054	0.0023	0.0019	0.0039	0.0080	0.0046
20	50	0.2	0.1	0.0080	0.0061	0.0059	0.0071	0.0020	0.0027
20	50	0.5	0.3	0.0034	0.0025	0.0022	0.0048	0.0017	0.0029
20	50	0.7	0.1	0.0023	0.0015	0.0013	0.0022	0.0031	0.0032
30	50	0.2	0.1	0.0075	0.0063	0.0061	0.0026	0.0008	0.0018
30	50	0.5	0.3	0.0029	0.0024	0.0021	0.0015	0.0009	0.0015
30	50	0.7	0.1	0.0016	0.0013	0.0011	0.0004	0.0012	0.0016
40	50	0.2	0.1	0.0070	0.0061	0.0060	0.0036	0.0007	0.0010
40	50	0.5	0.3	0.0027	0.0023	0.0021	0.0037	0.0006	0.0010
40	50	0.7	0.1	0.0014	0.0012	0.0011	0.0020	0.0006	0.0010
10	70	0.2	0.1	0.0116	0.0068	0.0067	0.0100	0.0035	0.0041
10	70	0.5	0.3	0.0065	0.0031	0.0029	0.0085	0.0063	0.0067
10	70	0.7	0.1	0.0053	0.0023	0.0018	0.0018	0.0068	0.0068
20	70	0.2	0.1	0.0083	0.0063	0.0062	0.0084	0.0011	0.0019
20	70	0.5	0.3	0.0034	0.0021	0.0022	0.0040	0.0014	0.0018
20	70	0.7	0.1	0.0024	0.0014	0.0013	0.0013	0.0023	0.0028
30	70	0.2	0.1	0.0072	0.0060	0.0058	0.0052	0.0005	0.0008
30	70	0.5	0.3	0.0027	0.0022	0.0020	0.0008	0.0007	0.0009
30	70	0.7	0.1	0.0015	0.0011	0.0010	0.0004	0.0007	0.0008
40	70	0.2	0.1	0.0068	0.0059	0.0058	0.0037	0.0005	0.0008
40	70	0.5	0.3	0.0025	0.0022	0.0020	0.0037	0.0004	0.0008
40	70	0.7	0.1	0.0013	0.0011	0.0010	0.0011	0.0005	0.0009

Spatial lag variable ρ: RMSEs

T	N	α	ρ	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.2	0.1	0.0018	0.0018	0.0018	0.0078	0.0979	0.0324
10	20	0.5	0.3	0.0015	0.0014	0.0015	0.0093	0.0333	0.0153
10	20	0.7	0.1	0.0015	0.0015	0.0016	0.0078	0.2125	0.0129
20	20	0.2	0.1	0.0011	0.0011	0.0011	0.0041	0.0320	0.0219
20	20	0.5	0.3	0.0008	0.0007	0.0007	0.0031	0.0100	0.0058
20	20	0.7	0.1	0.0006	0.0006	0.0006	0.0052	0.0331	0.0043
30	20	0.2	0.1	0.0005	0.0005	0.0005	0.0231	0.0111	0.0105
30	20	0.5	0.3	0.0005	0.0005	0.0005	0.0008	0.0046	0.0039
30	20	0.7	0.1	0.0003	0.0003	0.0003	0.0061	0.0054	0.0038
40	20	0.2	0.1	0.0005	0.0005	0.0005	0.0012	0.0075	0.0080
40	20	0.5	0.3	0.0005	0.0004	0.0005	0.0010	0.0020	0.0020
40	20	0.7	0.1	0.0002	0.0002	0.0002	0.0004	0.0022	0.0014
10	30	0.2	0.1	0.0015	0.0015	0.0015	0.0026	0.0393	0.0304
10	30	0.5	0.3	0.0011	0.0010	0.0010	0.0045	0.0252	0.0099
10	30	0.7	0.1	0.0010	0.0010	0.0011	0.0012	0.0307	0.0090
20	30	0.2	0.1	0.0006	0.0006	0.0006	0.0016	0.0122	0.0101
20	30	0.5	0.3	0.0006	0.0005	0.0005	0.0030	0.0082	0.0046
20	30	0.7	0.1	0.0003	0.0003	0.0003	0.0007	0.0108	0.0027
30	30	0.2	0.1	0.0004	0.0004	0.0004	0.0103	0.0061	0.0050
30	30	0.5	0.3	0.0006	0.0005	0.0005	0.0004	0.0029	0.0018
30	30	0.7	0.1	0.0002	0.0002	0.0002	0.0004	0.0027	0.0016
40	30	0.2	0.1	0.0003	0.0003	0.0003	0.0005	0.0059	0.0049
40	30	0.5	0.3	0.0004	0.0004	0.0004	0.0006	0.0016	0.0012
40	30	0.7	0.1	0.0002	0.0002	0.0002	0.0002	0.0033	0.0016
10	50	0.2	0.1	0.0009	0.0007	0.0009	0.0021	0.0140	0.0101
10	50	0.5	0.3	0.0007	0.0005	0.0006	0.0143	0.0771	0.0083
10	50	0.7	0.1	0.0006	0.0005	0.0006	0.0014	0.0174	0.0046
20	50	0.2	0.1	0.0004	0.0003	0.0003	0.0008	0.0094	0.0057
20	50	0.5	0.3	0.0005	0.0004	0.0004	0.0025	0.0032	0.0025
20	50	0.7	0.1	0.0002	0.0002	0.0002	0.0007	0.0058	0.0016
30	50	0.2	0.1	0.0003	0.0003	0.0003	0.0035	0.0042	0.0044
30	50	0.5	0.3	0.0004	0.0003	0.0003	0.0002	0.0019	0.0016
30	50	0.7	0.1	0.0001	0.0001	0.0001	0.0002	0.0016	0.0011
40	50	0.2	0.1	0.0002	0.0002	0.0002	0.0024	0.0030	0.0029
40	50	0.5	0.3	0.0004	0.0004	0.0004	0.0029	0.0012	0.0010
40	50	0.7	0.1	0.0002	0.0001	0.0002	0.0006	0.0012	0.0006
10	70	0.2	0.1	0.0005	0.0004	0.0004	0.0016	0.0143	0.0142
10	70	0.5	0.3	0.0006	0.0004	0.0005	0.0032	0.0199	0.0048
10	70	0.7	0.1	0.0004	0.0003	0.0004	0.0008	0.0121	0.0025
20	70	0.2	0.1	0.0002	0.0002	0.0002	0.0008	0.0050	0.0050
20	70	0.5	0.3	0.0005	0.0003	0.0004	0.0024	0.0024	0.0015
20	70	0.7	0.1	0.0002	0.0002	0.0002	0.0002	0.0032	0.0018
30	70	0.2	0.1	0.0002	0.0002	0.0002	0.0004	0.0026	0.0026
30	70	0.5	0.3	0.0004	0.0003	0.0003	0.0004	0.0012	0.0010
30	70	0.7	0.1	0.0001	0.0001	0.0001	0.0003	0.0009	0.0006
40	70	0.2	0.1	0.0001	0.0001	0.0001	0.0010	0.0020	0.0020
40	70	0.5	0.3	0.0004	0.0003	0.0003	0.0033	0.0007	0.0005
40	70	0.7	0.1	0.0001	0.0001	0.0001	0.0002	0.0007	0.0005

Endogenous variable γ: RMSE

T	N	α	ρ	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.2	0.1	0.1737	0.1688	0.1678	0.1737	0.0700	0.0909
10	20	0.5	0.3	0.1575	0.1554	0.1540	0.1606	0.0692	0.0803
10	20	0.7	0.1	0.1615	0.1564	0.1584	0.1664	0.0852	0.0470
20	20	0.2	0.1	0.1538	0.1508	0.1503	0.1677	0.0558	0.0572
20	20	0.5	0.3	0.1433	0.1416	0.1406	0.1438	0.0260	0.0281
20	20	0.7	0.1	0.1459	0.1442	0.1431	0.1491	0.0243	0.0221
30	20	0.2	0.1	0.1513	0.1490	0.1488	0.1813	0.0178	0.0236
30	20	0.5	0.3	0.1429	0.1413	0.1407	0.1469	0.0206	0.0209
30	20	0.7	0.1	0.1425	0.1383	0.1403	0.1607	0.0157	0.0193
40	20	0.2	0.1	0.1519	0.1502	0.1500	0.1519	0.0183	0.0225
40	20	0.5	0.3	0.1451	0.1440	0.1433	0.1824	0.0107	0.0147
40	20	0.7	0.1	0.1404	0.1391	0.1384	0.1501	0.0121	0.0131
10	30	0.2	0.1	0.1746	0.1695	0.1686	0.1733	0.0748	0.0738
10	30	0.5	0.3	0.1609	0.1570	0.1570	0.1934	0.0587	0.0384
10	30	0.7	0.1	0.1629	0.1555	0.1602	0.1650	0.0401	0.0345
20	30	0.2	0.1	0.1587	0.1558	0.1552	0.1579	0.0223	0.0269
20	30	0.5	0.3	0.1510	0.1493	0.1480	0.1496	0.0165	0.0214
20	30	0.7	0.1	0.1420	0.1403	0.1391	0.1519	0.0188	0.0173
30	30	0.2	0.1	0.1535	0.1513	0.1510	0.1941	0.0168	0.0254
30	30	0.5	0.3	0.1463	0.1449	0.1442	0.1632	0.0106	0.0120
30	30	0.7	0.1	0.1420	0.1407	0.1396	0.1547	0.0106	0.0119
40	30	0.2	0.1	0.1561	0.1544	0.1542	0.1600	0.0131	0.0131
40	30	0.5	0.3	0.1416	0.1406	0.1400	0.1540	0.0094	0.0117
40	30	0.7	0.1	0.1412	0.1401	0.1390	0.1483	0.0076	0.0079
10	50	0.2	0.1	0.1684	0.1613	0.1623	0.1669	0.0390	0.0426
10	50	0.5	0.3	0.1616	0.1534	0.1572	0.1583	0.0521	0.0446
10	50	0.7	0.1	0.1576	0.1559	0.1553	0.1613	0.0226	0.0201
20	50	0.2	0.1	0.1606	0.1576	0.1571	0.1615	0.0131	0.0177
20	50	0.5	0.3	0.1454	0.1435	0.1427	0.1441	0.0164	0.0198
20	50	0.7	0.1	0.1439	0.1399	0.1407	0.1441	0.0104	0.0106
30	50	0.2	0.1	0.1544	0.1522	0.1519	0.1608	0.0083	0.0131
30	50	0.5	0.3	0.1424	0.1412	0.1403	0.1568	0.0066	0.0084
30	50	0.7	0.1	0.1417	0.1405	0.1394	0.1475	0.0072	0.0066
40	50	0.2	0.1	0.1545	0.1528	0.1525	0.1584	0.0077	0.0089
40	50	0.5	0.3	0.1443	0.1432	0.1426	0.1441	0.0057	0.0068
40	50	0.7	0.1	0.1383	0.1374	0.1361	0.1586	0.0046	0.0048
10	70	0.2	0.1	0.1683	0.1611	0.1625	0.1702	0.0217	0.0234
10	70	0.5	0.3	0.1607	0.1536	0.1562	0.1581	0.0183	0.0237
10	70	0.7	0.1	0.1549	0.1518	0.1512	0.1647	0.0184	0.0178
20	70	0.2	0.1	0.1578	0.1546	0.1543	0.1574	0.0114	0.0150
20	70	0.5	0.3	0.1481	0.1400	0.1451	0.1495	0.0078	0.0084
20	70	0.7	0.1	0.1489	0.1430	0.1458	0.1525	0.0066	0.0082
30	70	0.2	0.1	0.1537	0.1515	0.1512	0.1570	0.0063	0.0077
30	70	0.5	0.3	0.1454	0.1439	0.1432	0.1599	0.0049	0.0066
30	70	0.7	0.1	0.1426	0.1410	0.1404	0.1489	0.0045	0.0045
40	70	0.2	0.1	0.1518	0.1501	0.1499	0.1565	0.0044	0.0047
40	70	0.5	0.3	0.1450	0.1440	0.1434	0.1435	0.0029	0.0041
40	70	0.7	0.1	0.1418	0.1408	0.1399	0.1429	0.0031	0.0037

Exogenous variable β: RMSE

T	N	α	ρ	SMLE	SDMLE	SDQMLE	LSDV	DIF-GMM	SYS-GMM
10	20	0.2	0.1	0.0064	0.0055	0.0054	0.0057	0.0129	0.0366
10	20	0.5	0.3	0.0057	0.0054	0.0052	0.0065	0.0115	0.0147
10	20	0.7	0.1	0.0055	0.0052	0.0049	0.0053	0.0245	0.0311
20	20	0.2	0.1	0.0042	0.0035	0.0034	0.0023	0.0037	0.0115
20	20	0.5	0.3	0.0035	0.0031	0.0029	0.0035	0.0052	0.0097
20	20	0.7	0.1	0.0029	0.0025	0.0024	0.0023	0.0062	0.0089
30	20	0.2	0.1	0.0046	0.0040	0.0039	0.0039	0.0036	0.0063
30	20	0.5	0.3	0.0027	0.0024	0.0023	0.0030	0.0030	0.0063
30	20	0.7	0.1	0.0024	0.0020	0.0018	0.0010	0.0019	0.0079
40	20	0.2	0.1	0.0041	0.0037	0.0036	0.0039	0.0021	0.0049
40	20	0.5	0.3	0.0022	0.0020	0.0019	0.0052	0.0026	0.0064
40	20	0.7	0.1	0.0018	0.0016	0.0015	0.0009	0.0019	0.0049
10	30	0.2	0.1	0.0046	0.0037	0.0035	0.0046	0.0057	0.0146
10	30	0.5	0.3	0.0034	0.0028	0.0027	0.0039	0.0063	0.0164
10	30	0.7	0.1	0.0034	0.0029	0.0028	0.0036	0.0076	0.0095
20	30	0.2	0.1	0.0043	0.0035	0.0034	0.0041	0.0036	0.0092
20	30	0.5	0.3	0.0022	0.0019	0.0017	0.0022	0.0034	0.0063
20	30	0.7	0.1	0.0022	0.0019	0.0017	0.0015	0.0027	0.0069
30	30	0.2	0.1	0.0040	0.0035	0.0034	0.0022	0.0022	0.0051
30	30	0.5	0.3	0.0023	0.0020	0.0019	0.0016	0.0018	0.0050
30	30	0.7	0.1	0.0016	0.0014	0.0012	0.0011	0.0020	0.0044
40	30	0.2	0.1	0.0039	0.0035	0.0034	0.0028	0.0014	0.0023
40	30	0.5	0.3	0.0019	0.0017	0.0016	0.0008	0.0013	0.0032
40	30	0.7	0.1	0.0019	0.0017	0.0016	0.0009	0.0016	0.0037
10	50	0.2	0.1	0.0043	0.0033	0.0032	0.0042	0.0044	0.0070
10	50	0.5	0.3	0.0026	0.0020	0.0020	0.0018	0.0041	0.0077
10	50	0.7	0.1	0.0019	0.0015	0.0014	0.0019	0.0045	0.0081
20	50	0.2	0.1	0.0032	0.0025	0.0024	0.0028	0.0023	0.0043
20	50	0.5	0.3	0.0024	0.0020	0.0018	0.0024	0.0021	0.0056
20	50	0.7	0.1	0.0015	0.0012	0.0011	0.0014	0.0016	0.0051
30	50	0.2	0.1	0.0036	0.0031	0.0030	0.0015	0.0012	0.0030
30	50	0.5	0.3	0.0017	0.0015	0.0013	0.0012	0.0009	0.0036
30	50	0.7	0.1	0.0014	0.0012	0.0011	0.0007	0.0009	0.0028
40	50	0.2	0.1	0.0036	0.0032	0.0031	0.0019	0.0011	0.0024
40	50	0.5	0.3	0.0019	0.0017	0.0016	0.0017	0.0008	0.0020
40	50	0.7	0.1	0.0013	0.0011	0.0010	0.0027	0.0007	0.0023
10	70	0.2	0.1	0.0038	0.0027	0.0027	0.0034	0.0031	0.0042
10	70	0.5	0.3	0.0019	0.0014	0.0013	0.0018	0.0031	0.0104
10	70	0.7	0.1	0.0017	0.0013	0.0013	0.0014	0.0034	0.0079
20	70	0.2	0.1	0.0037	0.0029	0.0029	0.0036	0.0017	0.0033
20	70	0.5	0.3	0.0019	0.0012	0.0013	0.0016	0.0011	0.0030
20	70	0.7	0.1	0.0012	0.0008	0.0008	0.0010	0.0012	0.0038
30	70	0.2	0.1	0.0035	0.0030	0.0029	0.0027	0.0010	0.0021
30	70	0.5	0.3	0.0018	0.0015	0.0014	0.0006	0.0010	0.0020
30	70	0.7	0.1	0.0014	0.0011	0.0010	0.0007	0.0007	0.0017
40	70	0.2	0.1	0.0035	0.0031	0.0030	0.0020	0.0006	0.0016
40	70	0.5	0.3	0.0019	0.0016	0.0015	0.0016	0.0007	0.0014
40	70	0.7	0.1	0.0012	0.0011	0.0009	0.0010	0.0007	0.0024