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

Madina Kukenova and Jose-Antonio Monteiro

University of Lausanne, University of Neuchatel

July 2008

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

Madina Kukenova* José-Antonio Monteiro†

First Draft: June, 2008
This Draft: October 2008

Abstract

Since there is so far no estimator that allows to estimate a dynamic panel model that includes a spatial lag as well as other potential endogenous variables. This paper wants to determine if it is suitable to instrument the spatial lag variable (which is by definition endogenous/simultaneous) using the instruments proposed by system GMM, i.e. lagged spatial lag values. The Monte Carlo investigation highlights the possibility to estimate a dynamic spatial lag model using the extended GMM proposed by Arellano and Bover (1995) and Blundell and Bover (1998), especially when N and T are large.

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

JEL classification: C15, C31, C33

*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 (2003), 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 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 several 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 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.

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 (2003) 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:

$$\begin{aligned} Y_t &= \delta Y_{t-1} + \rho W_{1t} Y_t + EX_t \beta + EN_t \gamma + \mu_t + \varepsilon_t \\ \varepsilon_t &= \varphi + \lambda W_{2t} \varepsilon_t + u_t, \quad t = 1, ..., T \end{aligned} \tag{1}$$

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, φ is the vector of country effect, μ is the vector of time 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, u_t is assumed to be normally distributed ($N(\mathbf{0}, \Omega)$). By adding some restrictions to the parameters, two popular spatial model specifications can be derived from this general spatial model, namely the dynamic spatial lag model ($\lambda = 0$) and the dynamic spatial error model ($\rho = 0$)[1].

The spatial lag model accounts directly for relationships between dependent variables that are believed to be related in some spatial way. Somewhat analogous to a lagged dependent variable in time series analysis, the estimated "spatial lag" coefficient characterizes the contemporaneous correlation between one country's Y and other geographically-proximate country' $Y's$. The following equation gives the basic specification for the "time-space simultaneous" model (Anselin (2001))[2]:

$$Y_t = \delta Y_{t-1} + \rho W_t Y_t + EX_t \beta + EN_t \gamma + \varphi + \mu_t + u_t \tag{2}$$

[1] The analysis of the spatial error panel model is beyond the scope of this paper. For further detail, see Elhorst (2003) and Kapoor et al. (2007).

[2] Beside the "time-space simulatenous" model, Anselin (2001) distinguishes three other distinct spatial lag panel models: the "pure space recursive" model which only includes a lagged spatial lag coefficient ($\rho W_{1,t-1} Y_{t-1}$); 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.

The spatial autoregressive coefficient (ρ) associated with W_tY_t represents the effect of the weighted average ($w_t(d_{ij})$ being the weights) of the neighborhood, i.e. $[W_tY_t]_i = \sum_{j=1..N_t} w_t(d_{ij}) \cdot Y_{jt}$. The spatial lag term allows to determine if the variable Y 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. This effect is assumed to lie between -1 and +1. As such, this model allows the data to reveal patterns of substitution or complementarity through the estimated spatial lag coefficient.

Note that the spatial lag term W_tY_t is correlated with the disturbances, even if u_t are independently and identically distributed. To see this point more formally, note that the reduced form of equation (2) take the following form:

$$Y_t = (I_N - \rho W_t)^{-1} (\delta Y_{t-1} + EX_t\beta + EN_t\gamma + \varphi + \mu_t + u_t)$$

Each element of Y_t is a linear combination of all of the error terms. Moreover, as pointed out by Anselin (2003), since $|\rho| < 1$ and each element of W_t is smaller than one implies 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. 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)). Therefore, the spatial lag coefficient must be treated as an endogenous variable and proper estimation methods must account for this endogeneity.

Despite the fact that dynamic panel models have been the object of recent important developments (Baltagi and Kao (2000), 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:

Table 1: Spatial Dynamic Panel Estimators

Model	Estimation Methods	Endogenous Variable(s)
$Y_t = \alpha Y_{t-1} + \beta EX_t + \epsilon_t$	Difference GMM (Arellano and Bond (1991)) MLE/MDE (Hsiao, Pesaran and Tahmiscioglu (2001)) System-GMM (Arellano and Bover (1995), Blundell and Bond (1998))	Y_{t-1};
$Y_t = \alpha Y_{t-1} + \beta EX_t + \gamma EN_t + \epsilon_t$	System-GMM (Arellano and Bover (1995), Blundell and Bond (1998))	$Y_{t-1}; EN_t$
$Y_t = \rho WY_t + \beta EX_t + \epsilon_t$	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 and Le Gallo (2007))	$WY_t; EN_t$
$Y_t = \alpha Y_{t-1} + \rho WY_t + \beta EX_t + \epsilon_t$	MLE (Elrhost (2003), Lee and Yu (2007)) QMLE (Yu, de Jong and Lee (2006) (2007), Lee and Yu (2007)) 2SLSDV (Beenstock and Felsenstein (2007))	$WY_t; Y_{t-1}$
$Y_t = \alpha Y_{t-1} + \rho WY_t + \beta EX_t + \gamma EN_t + \epsilon_t$	System-GMM (Arellano and Bover (1995), Blundell and 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 maximum likelihood (Elhorst (2003b)) or spatial two-stage least squares methods (Anselin (1988) (2001)). The ML approach consists of estimating the spatial coefficient using a non-linear optimization routine that maximizes 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[3]. 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, applying spatial 2SLS with lower orders of the spatial lags of the exogenous variables as instrument for the spatial autoregressive term[4].

In a dynamic context, Elhorst (2003a) proposes to estimate a reduced form of the model in first-difference using maximum likelihood. Yu et al. (2006, 2007) provide a theoretical analysis on the asymptotic properties of the ML and QML estimators, assuming the process is stationary and partially nonstationary, respectively. In order to account for not only unobservable individual effects but also unobserved time effects, Lee and Yu (2007) propose to transform the data to eliminate the time effects (bias) and then estimate the model using QML. Recently, Beenstock and Felsenstein (2007) suggest a two-step procedure to estimate a spatial panel vector autoregression model. The first step consists of applying least square dummy variables (LSDV) to the model omitting 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

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

[4]Recently, Fingleton and Le Gallo (2008) proposes 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.

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 MLE, S2SLS or spatial GMM 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

In the absence of spatial dependence, there are different estimators available to estimate a dynamic panel model, like classical GMM (Arellano and Bond (1992)) and MLE (Hsiao, Pesaran and Tahmiscioglu (2002)). However, 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 GMM estimator[5], suggested by Arellano and Bover (1995) and Blundell and Bond (1998).

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. 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.[6]. In particular, it can correct for the endogeneity of the spatial lag and lagged dependent variable as well as other potentially endogenous explanatory variables. Extended GMM allows also to take into consideration some econometrics problems such as measurement error and weak instruments. It also controls for time-invariant individual-specific effects such as distance, culture and political structure. On a practical ground, it also avoids the inversion of high dimension spatial weights matrix W and the computation of its eigenvalues[7], which can be sometimes computationally

[5] See for example, Madriaga and Poncet (2007) or Foucault, Madies and Paty (2008).

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

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

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 ($t = 1, .., T$):

$$Y_{it} = \delta Y_{it-1} + \rho \left[W_t Y_t\right]_i + EX_{it}\beta + EN_{it}\gamma + \varphi_i + \mu_t + u_{it} \tag{3}$$

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

$$\triangle Y_{it} = \delta \triangle Y_{it-1} + \rho \triangle \left[W_t y_t\right]_i + \triangle EX_{it}\beta + \triangle EN_{it}\gamma + \mu_t + \triangle u_{it} \tag{4}$$

Even if the fixed effects (within) estimator cancels the country individual fixed (φ_i), the lagged endogenous variable ($\triangle Y_{it-1}$) is still correlated with the idiosyncratic error terms (u_{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):

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

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:

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

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. $\mathrm{E}(EX_{it}u_{it}) \neq 0$. 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:

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

For the different endogenous variables, the valid moment conditions are

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

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:

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

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. The null hypothesis is that instruments are not correlated with the residuals. This null hypothesis of no misspecification is rejected if the minimized

GMM criterion function is greater than the value of a chi-squared distribution with the degree of freedom equal to the difference between the number of moment conditions and number of parameters. The validity of the moment conditions can also be evaluated with the Sargan/Hansen-difference test, which checks the validity of extra moment conditions over that of weak exogeneity. If the Sargan-difference test rejects the validity of these extra moment conditions, then the strong assumption of strict exogeneity will be in doubt. Aware that too many instrument variables tend to validated invalid results through the Hansen J test for joint validity of those instruments, as well as the difference-in-Sargan/Hansen tests for subsets of instruments, it is advised to restrict the number of instruments by defining a maximum number of lags or by collapsing the instruments (see Roodman (2006)).

Second, the Arellano-Bond test examines the serial correlation property of the level residuals. If the level residuals were serially uncorrelated, then the first-differenced residuals in (6) would, by construction, follow a MA(1) process. This would imply that autocorrelations of the first-order are different from zero, while the second (m_2) or higher-ones are equal to zero. Applied to the residuals in difference, the $m_1$and m_2 Arellano-Bond statistics test the null hypothesis of zero first-order and second-order autocorrelation of the idiosyncratic disturbances. Knowing that $\triangle u_{it}$ is mathematically related to $\triangle u_{it-1}$ via the shared term u_{it-1}, one can expect a first-order serial correlation in differences. That is why, in order to check first-correlation in levels, we rely on the Arellano bond test for second order autocorrelation (m_2). An insignificant $m_1$and/or significant m_2 suggest the likely presence of invalid moment conditions due to serial correlation in the level residuals.

3 A Monte-Carlo Study

In this section, we investigate the properties of using extended GMM to account for the endogeneity of the spatial lag in a dynamic panel data context using Monte-Carlo simulations. Simulation studies already showed that SYS-GMM is the right estimator when the panel model includes a time autoregressive coefficient and several endogenous variables. That is why we only focus here on the estimation of the spatial lag and its consistency. The dynamic spatial lag panel data model is thus defined as follows:

$$
\begin{aligned}
Y_t &= \delta Y_{t-1} + \rho W Y_t + \beta X_t + \varepsilon_t \\
\varepsilon_t &= \varphi + u_t
\end{aligned} \tag{14}
$$

where Y_t is $N \times 1$ vector, W is a $N \times N$ spatial weight matrix[8], X_t is an $N \times 1$ exogenous variable, η is the individual effect while u_t is the error term which is normally distributed.

In order to check the consistency of the spatial autoregressive estimator, we consider the following different designs with different period and cross-country sizes[9]:

$$
\begin{gathered}
T \in \{5;\ 20;\ 30;\ 40\} \\
N \in \{5;\ 20;\ 50;\ 70\} \\
\delta \in \{0;\ 0.25;\ 0.5;\ 0.75\} \\
\rho \in \{0;\ 0.25;\ 0.5;\ 0.75\} \\
\beta \in \{1\}
\end{gathered}
$$

There is a total of 256 different designs $(4 \times 4 \times 4 \times 4)$. For each of these designs, we performed 1000 trials. Note that for each design, the exogenous variables and spatial weight matrices are generated once according to a standard normal distribution. In order to compute the initial observations Y_0, we create a 30×1 vector of initial disturbances:

$$
\epsilon = \sum_{t=0}^{\infty} \delta^t \varepsilon_{-t}
$$

We then construct the $N \times 1$ vector of initial observations according to the following equation:

$$
Y_0 = (I_N - \rho W)^{-1} \left[\epsilon + (1-\delta)^{-1} \eta \right]
$$

[8]Note that we only consider a balanced panel model to simplify the simulation process.

[9]Note that the designs with δ and $\rho \in (0.5, 0.75)$ are subject to high multicollinearity.

The subsequent observations for $t = 1, ..., T$ are then generated according to the following reduced form:

$$Y_t = (I_N - \rho W)^{-1} [\delta Y_{t-1} + \beta X_t + \varepsilon_t]$$

Following Kapoor et al. (2007) and Kelejian and Prucha (1999), we consider different types of spatial weight matrix. In each case, the matrices are row-standardized so that all non zero elements in each row sum to one. The first three matrices rely on a perfect "idealized" circular world, while the last ones consider a real-word weighting scheme. The three "theoretical" spatial matrices, referred as "1 ahead and 1 behind" (W^{I1}), "3 ahead and 3 behind" (W^{I3}) and "5 ahead and 5 behind" (W^{I5}), respectively, are characterized by different degree of sparseness. Each are such that each location is related to the one/three/five locations immediately before and after it, so that each nonzero elements are equal to $0.5/0.\overline{3}/0.1$, respectively. The last two spatial weighting schemes are based on real distance data. We consider the distance between capitals among OECD countries and among non OECD countries[10], respectively. In order to avoid giving some positive weight to very remote countries (with weaker cultural, political and economic ties), we consider the negative exponential weighting scheme. This is done by dividing the distance between locations j and k by the minimum distance within the region r (where location j lies within region r):$w(d_{j,k}) = \exp(-d_{j,k}/MIN_{r,j})$ if $j \neq k$.

As mentioned previously, extended GMM relies on the specification of instruments. In order to check if the estimated spatial lag coefficient is sensitive to the instruments structure, we consider different approaches. Each endogenous variables (Y_{t-1}, WY_t) will thus be instrumented by their

1. 2th and 3rd lags values, using the *collapse* option and the exogenous variable X_t;
2. 2th and 3rd lags values, without the *collapse* option and the exogenous variable X_t;
3. 2th and lower lags values, using the *collapse* option and the exogenous variables X_t and WX_t;

[10] The data is taken from CEPII database.

4. 2th and lower lags values, without the *collapse* option and the exogenous variables X_t and WX_t;

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.

Since the results are qualitatively similar with respect to different spatial weight schemes, for sake of brievty we only present the results for "1 ahead and 1 behind" W. The full results are given in table 2 in appendix.

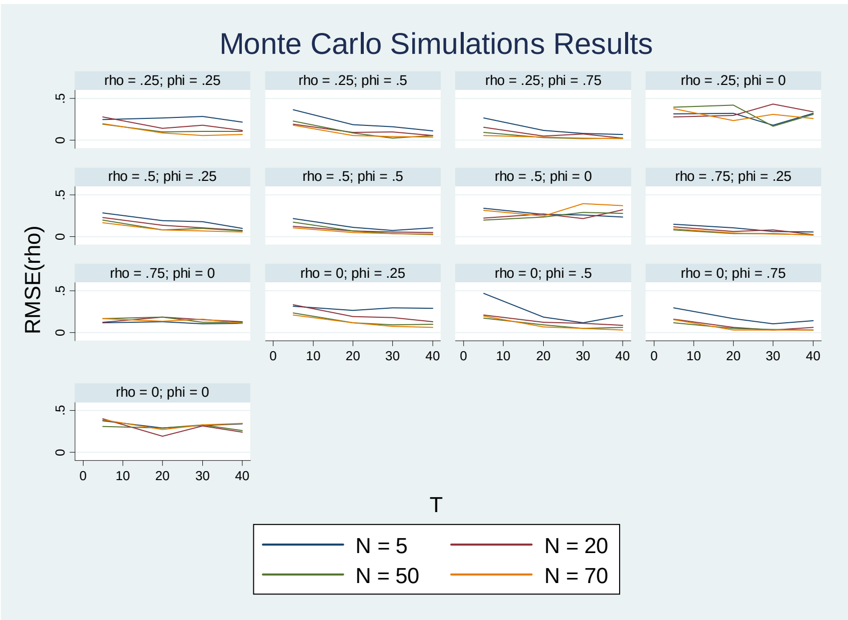

Figure 1: Monte-Carlo Simulation for T fixed

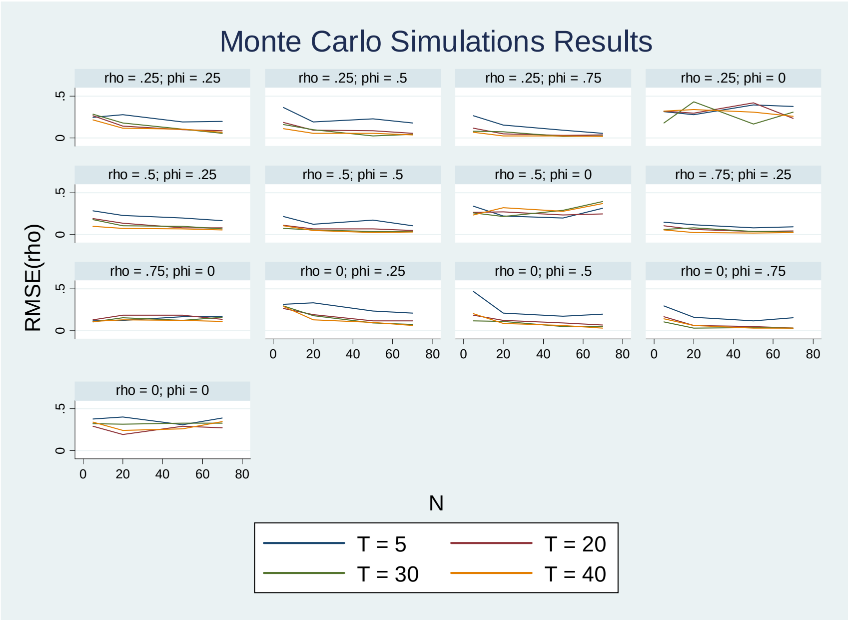

Figure 2: Monte Carlo Simulations for N fixed

The Monte Carlo investigation highlights several important facts:

- System GMM can estimate consistently the spatial lag ρ. However, the rate of consistency is faster when T is fixed and N increases than when T increases for a given cross-section N size.
- When the sample and period size are relatively small, one should use extended GMM carefully. The spatial autoregressive coefficient tend to be over-estimated. This comes from the fact that unlike MLE, the spatial lag term is not bounded. This becomes an important issue when the lagged dependent variable is significantly not different from zero. In reality, this is relatively intuitive: when there is no time dynamic, then extended GMM is no longer relevant. It is probably more suitable to estimate the model using MLE.
- When both spatial and time lagged coefficients are close to one, then the probability to face multicollinearity problem increases. This issue tends to disappear once the number of cross-section N increases with T fixed, but not the other way around.

- The presence of exogenous variables reduces the probability to overestimate the spatial autoregressive term. This was confirmed by comparing the results of a pure first order dnyamic spatial autoregressive model and a mixed spatial autoregressive specification.

4 Conclusion

This study shows that system GMM can estimate consistently the spatial lag coefficient as the N (T) increases with T (N) fixed. Until a new estimator that allows to account for the endogeneity of the lagged dependent variable, spatial lag and other potentially endogenous variables is found, applied researchers can apply extended GMM to estimate "time-space simultaneous" models. However, one should be careful when the sample is relatively small. This is especially true if one expects to obtain a small time autoregressive coefficient.

References

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

[20] 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.

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

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

[23] 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

[24] 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 Econometrics109, 107-150.

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

[26] 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.

[27] 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.

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

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

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

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

[32] 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

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

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

[35] Yu, J., R. de Jong and L. Lee, 2006, "Quasi-Maximum Likelihood Estimators For Spatial Dynamic Panel Data With Fixed Effects When Both n and T Are Large", University of Rochester. http://www.econ.ohio-state.edu/lee/wp/Yu-deJong-Lee-2006.pdf

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

5 Appendices

5.A Monte Carlo Results for ρ

Time	Country	Phi	Rho	bias				RMSE			
				IV 1	IV 2	IV 3	IV 4	IV 1	IV 2	IV 3	IV 4
5	5	0	0	0.038	0.016	0.036	0.037	0.373	0.208	0.216	0.469
20	5	0	0	0.008	0.024	0.004	0.043	0.289	0.103	0.115	0.222
30	5	0	0	0.028	0.011	0.019	0.051	0.320	0.056	0.087	0.188
40	5	0	0	0.059	0.014	0.016	0.023	0.336	0.083	0.085	0.168
5	5	0.25	0	0.047	0.012	0.011	0.112	0.314	0.218	0.209	0.620
20	5	0.25	0	0.114	0.004	0.016	0.009	0.267	0.064	0.133	0.243
30	5	0.25	0	0.056	0.005	0.007	0.004	0.299	0.068	0.087	0.129
40	5	0.25	0	0.038	0.004	0.001	0.012	0.291	0.059	0.068	0.130
5	5	0.5	0	0.075	0.012	0.006	0.060	0.471	0.167	0.184	0.217
20	5	0.5	0	0.003	0.007	0.006	0.028	0.185	0.081	0.098	0.173
30	5	0.5	0	0.020	0.023	0.014	0.016	0.118	0.075	0.076	0.123
40	5	0.5	0	0.034	0.006	0.009	0.003	0.205	0.072	0.067	0.114
5	5	0.75	0	0.152	0.019	0.098	0.044	0.300	0.141	0.200	0.273
20	5	0.75	0	0.019	0.010	0.011	0.010	0.168	0.048	0.081	0.088
30	5	0.75	0	0.019	0.013	0.013	0.008	0.110	0.051	0.072	0.082
40	5	0.75	0	0.035	0.001	0.012	0.007	0.143	0.039	0.070	0.056
5	5	0	0.25	0.181	0.109	0.095	0.453	0.319	0.242	0.243	0.549
20	5	0	0.25	0.098	0.074	0.009	0.202	0.322	0.123	0.083	0.268
30	5	0	0.25	0.066	0.073	0.041	0.234	0.182	0.098	0.080	0.291
40	5	0	0.25	0.075	0.080	0.012	0.203	0.325	0.096	0.093	0.231
5	5	0.25	0.25	0.068	0.048	0.028	0.163	0.245	0.210	0.196	0.365
20	5	0.25	0.25	0.071	0.066	0.021	0.215	0.266	0.089	0.107	0.265
30	5	0.25	0.25	0.015	0.073	0.014	0.211	0.283	0.094	0.090	0.256
40	5	0.25	0.25	0.085	0.061	0.012	0.223	0.218	0.086	0.076	0.257
5	5	0.5	0.25	0.062	0.070	0.017	0.161	0.365	0.212	0.252	0.398
20	5	0.5	0.25	0.095	0.066	0.029	0.207	0.186	0.096	0.078	0.254
30	5	0.5	0.25	0.017	0.045	0.011	0.144	0.163	0.069	0.079	0.173
40	5	0.5	0.25	0.021	0.067	0.027	0.186	0.111	0.086	0.058	0.211
5	5	0.75	0.25	0.074	0.022	0.055	0.070	0.269	0.110	0.178	0.216
20	5	0.75	0.25	0.019	0.014	0.022	0.048	0.121	0.042	0.083	0.093
30	5	0.75	0.25	0.005	0.029	0.009	0.061	0.082	0.055	0.067	0.089
40	5	0.75	0.25	0.001	0.016	0.011	0.052	0.072	0.033	0.044	0.084
5	5	0	0.5	0.210	0.060	0.105	0.339	0.338	0.150	0.169	0.393
20	5	0	0.5	0.152	0.102	0.023	0.310	0.265	0.125	0.078	0.320
30	5	0	0.5	0.154	0.116	0.020	0.310	0.260	0.125	0.059	0.314
40	5	0	0.5	0.173	0.109	0.013	0.307	0.233	0.117	0.061	0.313
5	5	0.25	0.5	0.151	0.132	0.106	0.311	0.284	0.198	0.204	0.376
20	5	0.25	0.5	0.116	0.090	0.027	0.309	0.190	0.102	0.065	0.328
30	5	0.25	0.5	0.052	0.081	0.004	0.266	0.176	0.095	0.070	0.279
40	5	0.25	0.5	0.031	0.084	0.003	0.293	0.100	0.093	0.066	0.302
5	5	0.5	0.5	0.069	0.082	0.052	0.262	0.216	0.155	0.191	0.314
20	5	0.5	0.5	0.025	0.077	0.032	0.203	0.111	0.088	0.061	0.214
30	5	0.5	0.5	0.003	0.049	0.002	0.158	0.076	0.062	0.047	0.173
40	5	0.5	0.5	0.001	0.051	0.006	0.172	0.105	0.059	0.059	0.178
5	5	0	0.75	0.085	0.070	0.069	0.243	0.118	0.095	0.101	0.247
20	5	0	0.75	0.093	0.085	0.026	0.230	0.130	0.090	0.055	0.231
30	5	0	0.75	0.068	0.079	0.009	0.222	0.108	0.084	0.046	0.223
40	5	0	0.75	0.074	0.083	0.017	0.221	0.116	0.086	0.040	0.222
5	5	0.25	0.75	0.101	0.069	0.079	0.208	0.144	0.091	0.113	0.221
20	5	0.25	0.75	0.070	0.073	0.019	0.187	0.105	0.079	0.046	0.189
30	5	0.25	0.75	0.036	0.065	0.012	0.178	0.063	0.073	0.047	0.179
40	5	0.25	0.75	0.017	0.061	0.005	0.182	0.054	0.064	0.035	0.183

Time	Country	Phi	Rho	bias				RMSE			
				IV 1	IV 2	IV 3	IV 4	IV 1	IV 2	IV 3	IV 4
5	20	0	0	0.129	0.033	0.005	0.038	0.400	0.181	0.125	0.355
20	20	0	0	0.017	0.010	0.013	0.020	0.190	0.056	0.053	0.122
30	20	0	0	0.008	0.008	0.041	0.061	0.310	0.059	0.056	0.107
40	20	0	0	0.065	0.027	0.003	0.065	0.240	0.048	0.047	0.124
5	20	0.25	0	0.144	0.003	0.002	0.023	0.336	0.253	0.141	0.190
20	20	0.25	0	0.060	0.024	0.024	0.045	0.196	0.081	0.064	0.132
30	20	0.25	0	0.049	0.011	0.010	0.035	0.181	0.033	0.048	0.128
40	20	0.25	0	0.012	0.010	0.002	0.021	0.133	0.045	0.023	0.104
5	20	0.5	0	0.012	0.036	0.024	0.025	0.210	0.149	0.162	0.310
20	20	0.5	0	0.045	0.003	0.012	0.046	0.126	0.061	0.052	0.106
30	20	0.5	0	0.024	0.019	0.018	0.022	0.113	0.044	0.041	0.089
40	20	0.5	0	0.013	0.003	0.004	0.016	0.091	0.022	0.029	0.053
5	20	0.75	0	0.094	0.026	0.023	0.030	0.160	0.092	0.090	0.246
20	20	0.75	0	0.027	0.010	0.013	0.046	0.065	0.054	0.047	0.120
30	20	0.75	0	0.003	0.001	0.000	0.006	0.036	0.027	0.036	0.049
40	20	0.75	0	0.002	0.011	0.000	0.002	0.062	0.022	0.024	0.059
5	20	0	0.25	0.121	0.103	0.026	0.301	0.278	0.174	0.107	0.352
20	20	0	0.25	0.084	0.114	0.029	0.275	0.298	0.129	0.063	0.287
30	20	0	0.25	0.262	0.107	0.016	0.271	0.430	0.113	0.028	0.275
40	20	0	0.25	0.181	0.093	0.004	0.231	0.340	0.104	0.035	0.239
5	20	0.25	0.25	0.059	0.063	0.015	0.268	0.277	0.184	0.085	0.349
20	20	0.25	0.25	0.052	0.088	0.007	0.238	0.143	0.114	0.036	0.262
30	20	0.25	0.25	0.039	0.116	0.016	0.247	0.180	0.121	0.036	0.252
40	20	0.25	0.25	0.005	0.092	0.016	0.256	0.117	0.099	0.028	0.269
5	20	0.5	0.25	0.069	0.097	0.011	0.202	0.191	0.143	0.109	0.271
20	20	0.5	0.25	0.057	0.058	0.017	0.193	0.096	0.073	0.039	0.202
30	20	0.5	0.25	0.023	0.072	0.000	0.214	0.099	0.094	0.039	0.230
40	20	0.5	0.25	0.009	0.059	0.005	0.193	0.056	0.069	0.040	0.202
5	20	0.75	0.25	0.089	0.066	0.023	0.078	0.156	0.116	0.074	0.258
20	20	0.75	0.25	0.000	0.030	0.004	0.153	0.049	0.045	0.043	0.163
30	20	0.75	0.25	0.010	0.035	0.014	0.141	0.076	0.042	0.036	0.152
40	20	0.75	0.25	0.005	0.033	0.005	0.151	0.029	0.039	0.027	0.164
5	20	0	0.5	0.169	0.153	0.035	0.395	0.220	0.175	0.078	0.411
20	20	0	0.5	0.094	0.123	0.006	0.376	0.268	0.128	0.031	0.380
30	20	0	0.5	0.166	0.127	0.011	0.339	0.214	0.130	0.043	0.346
40	20	0	0.5	0.246	0.152	0.010	0.375	0.320	0.155	0.026	0.377
5	20	0.25	0.5	0.092	0.169	0.014	0.399	0.224	0.188	0.087	0.408
20	20	0.25	0.5	0.090	0.128	0.015	0.368	0.137	0.134	0.028	0.369
30	20	0.25	0.5	0.018	0.127	0.003	0.360	0.101	0.130	0.014	0.362
40	20	0.25	0.5	0.031	0.106	0.004	0.357	0.075	0.115	0.027	0.359
5	20	0.5	0.5	0.077	0.050	0.007	0.321	0.121	0.086	0.060	0.334
20	20	0.5	0.5	0.031	0.064	0.009	0.263	0.068	0.072	0.035	0.269
30	20	0.5	0.5	0.024	0.061	0.000	0.257	0.055	0.072	0.032	0.261
40	20	0.5	0.5	0.010	0.065	0.011	0.272	0.046	0.068	0.026	0.274
5	20	0	0.75	0.083	0.097	0.018	0.291	0.127	0.108	0.065	0.299
20	20	0	0.75	0.141	0.107	0.009	0.302	0.189	0.110	0.034	0.302
30	20	0	0.75	0.120	0.098	0.000	0.291	0.158	0.099	0.023	0.292
40	20	0	0.75	0.110	0.113	0.010	0.287	0.129	0.114	0.025	0.287
5	20	0.25	0.75	0.076	0.080	0.009	0.296	0.116	0.118	0.052	0.298
20	20	0.25	0.75	0.050	0.081	0.010	0.252	0.062	0.085	0.038	0.253
30	20	0.25	0.75	0.029	0.080	0.008	0.247	0.078	0.082	0.020	0.248
40	20	0.25	0.75	0.013	0.068	0.001	0.244	0.026	0.069	0.011	0.245

Time	Country	Phi	Rho	bias				RMSE			
				IV 1	IV 2	IV 3	IV 4	IV 1	IV 2	IV 3	IV 4
5	50	0	0	0.010	0.058	0.005	0.023	0.307	0.151	0.034	0.129
20	50	0	0	0.015	0.030	0.008	0.001	0.287	0.078	0.044	0.135
30	50	0	0	0.141	0.004	0.000	0.007	0.323	0.066	0.034	0.089
40	50	0	0	0.009	0.029	0.001	0.007	0.259	0.070	0.024	0.051
5	50	0.25	0	0.095	0.072	0.010	0.036	0.236	0.182	0.089	0.325
20	50	0.25	0	0.031	0.010	0.004	0.011	0.122	0.041	0.032	0.077
30	50	0.25	0	0.024	0.016	0.003	0.008	0.093	0.072	0.031	0.095
40	50	0.25	0	0.011	0.010	0.001	0.018	0.098	0.036	0.026	0.051
5	50	0.5	0	0.009	0.008	0.012	0.015	0.174	0.106	0.061	0.222
20	50	0.5	0	0.002	0.011	0.001	0.006	0.097	0.079	0.050	0.085
30	50	0.5	0	0.014	0.002	0.006	0.020	0.049	0.069	0.019	0.064
40	50	0.5	0	0.003	0.000	0.002	0.007	0.066	0.030	0.017	0.051
5	50	0.75	0	0.014	0.018	0.013	0.042	0.117	0.078	0.056	0.125
20	50	0.75	0	0.012	0.009	0.016	0.015	0.051	0.025	0.047	0.081
30	50	0.75	0	0.012	0.008	0.005	0.007	0.041	0.033	0.016	0.030
40	50	0.75	0	0.002	0.001	0.002	0.006	0.036	0.018	0.020	0.028
5	50	0	0.25	0.012	0.069	0.023	0.268	0.397	0.218	0.097	0.310
20	50	0	0.25	0.042	0.086	0.005	0.244	0.417	0.102	0.021	0.248
30	50	0	0.25	0.105	0.097	0.012	0.192	0.169	0.109	0.027	0.203
40	50	0	0.25	0.037	0.095	0.010	0.236	0.311	0.120	0.020	0.243
5	50	0.25	0.25	0.042	0.053	0.001	0.258	0.192	0.148	0.064	0.307
20	50	0.25	0.25	0.002	0.056	0.016	0.206	0.102	0.079	0.034	0.222
30	50	0.25	0.25	0.027	0.080	0.000	0.226	0.109	0.093	0.029	0.230
40	50	0.25	0.25	0.051	0.080	0.001	0.243	0.106	0.087	0.020	0.245
5	50	0.5	0.25	0.023	0.079	0.007	0.238	0.227	0.141	0.082	0.287
20	50	0.5	0.25	0.014	0.056	0.001	0.203	0.088	0.076	0.033	0.211
30	50	0.5	0.25	0.000	0.049	0.002	0.231	0.026	0.051	0.021	0.235
40	50	0.5	0.25	0.003	0.048	0.001	0.200	0.058	0.061	0.018	0.204
5	50	0.75	0.25	0.013	0.030	0.008	0.204	0.095	0.090	0.050	0.218
20	50	0.75	0.25	0.004	0.015	0.014	0.183	0.031	0.029	0.041	0.187
30	50	0.75	0.25	0.003	0.019	0.004	0.152	0.024	0.022	0.019	0.155
40	50	0.75	0.25	0.005	0.022	0.000	0.174	0.025	0.027	0.015	0.178
5	50	0	0.5	0.094	0.156	0.006	0.369	0.196	0.202	0.061	0.374
20	50	0	0.5	0.160	0.140	0.009	0.360	0.231	0.152	0.024	0.361
30	50	0	0.5	0.076	0.132	0.000	0.365	0.286	0.137	0.016	0.366
40	50	0	0.5	0.144	0.140	0.002	0.364	0.274	0.141	0.017	0.365
5	50	0.25	0.5	0.125	0.109	0.036	0.388	0.195	0.143	0.066	0.395
20	50	0.25	0.5	0.021	0.101	0.004	0.360	0.080	0.106	0.029	0.362
30	50	0.25	0.5	0.031	0.094	0.003	0.357	0.098	0.101	0.029	0.358
40	50	0.25	0.5	0.007	0.100	0.001	0.352	0.068	0.102	0.013	0.353
5	50	0.5	0.5	0.060	0.087	0.039	0.360	0.171	0.142	0.064	0.365
20	50	0.5	0.5	0.026	0.058	0.007	0.297	0.064	0.067	0.020	0.302
30	50	0.5	0.5	0.023	0.056	0.012	0.285	0.037	0.065	0.022	0.287
40	50	0.5	0.5	0.015	0.040	0.002	0.276	0.023	0.042	0.014	0.278
5	50	0	0.75	0.131	0.094	0.033	0.313	0.169	0.119	0.047	0.315
20	50	0	0.75	0.148	0.105	0.002	0.306	0.190	0.108	0.017	0.306
30	50	0	0.75	0.074	0.109	0.005	0.306	0.123	0.110	0.013	0.306
40	50	0	0.75	0.118	0.098	0.006	0.303	0.125	0.099	0.014	0.304
5	50	0.25	0.75	0.044	0.058	0.003	0.292	0.082	0.074	0.028	0.294
20	50	0.25	0.75	0.004	0.060	0.002	0.266	0.035	0.061	0.011	0.266
30	50	0.25	0.75	0.019	0.063	0.004	0.262	0.039	0.064	0.012	0.262
40	50	0.25	0.75	0.005	0.056	0.004	0.253	0.018	0.057	0.009	0.253

Time	Country	Phi	Rho	bias				RMSE			
				IV 1	IV 2	IV 3	IV 4	IV 1	IV 2	IV 3	IV 4
5	70	0	0	0.080	0.014	0.012	0.011	0.389	0.177	0.064	0.150
20	70	0	0	0.020	0.021	0.005	0.015	0.269	0.054	0.019	0.080
30	70	0	0	0.014	0.010	0.002	0.008	0.327	0.068	0.020	0.057
40	70	0	0	0.094	0.007	0.004	0.004	0.344	0.045	0.013	0.051
5	70	0.25	0	0.076	0.014	0.006	0.015	0.209	0.135	0.043	0.154
20	70	0.25	0	0.001	0.019	0.002	0.000	0.120	0.094	0.014	0.066
30	70	0.25	0	0.002	0.016	0.001	0.042	0.078	0.058	0.018	0.069
40	70	0.25	0	0.010	0.011	0.003	0.015	0.067	0.050	0.022	0.033
5	70	0.5	0	0.001	0.061	0.012	0.063	0.198	0.122	0.070	0.186
20	70	0.5	0	0.021	0.008	0.004	0.006	0.070	0.054	0.017	0.060
30	70	0.5	0	0.019	0.007	0.000	0.024	0.050	0.023	0.021	0.057
40	70	0.5	0	0.009	0.007	0.005	0.012	0.035	0.035	0.018	0.043
5	70	0.75	0	0.039	0.039	0.029	0.030	0.154	0.089	0.057	0.092
20	70	0.75	0	0.024	0.017	0.020	0.010	0.035	0.025	0.037	0.049
30	70	0.75	0	0.002	0.000	0.004	0.008	0.033	0.016	0.009	0.030
40	70	0.75	0	0.016	0.011	0.004	0.014	0.033	0.024	0.017	0.063
5	70	0	0.25	0.106	0.023	0.007	0.252	0.375	0.241	0.041	0.265
20	70	0	0.25	0.123	0.087	0.001	0.231	0.237	0.101	0.029	0.240
30	70	0	0.25	0.003	0.078	0.008	0.238	0.311	0.090	0.026	0.245
40	70	0	0.25	0.012	0.094	0.007	0.250	0.263	0.108	0.023	0.254
5	70	0.25	0.25	0.128	0.124	0.003	0.219	0.200	0.155	0.070	0.253
20	70	0.25	0.25	0.041	0.044	0.009	0.225	0.091	0.068	0.035	0.232
30	70	0.25	0.25	0.021	0.058	0.004	0.235	0.055	0.071	0.020	0.241
40	70	0.25	0.25	0.005	0.053	0.002	0.225	0.071	0.072	0.013	0.227
5	70	0.5	0.25	0.081	0.035	0.002	0.247	0.178	0.129	0.025	0.272
20	70	0.5	0.25	0.019	0.042	0.001	0.236	0.055	0.049	0.040	0.242
30	70	0.5	0.25	0.005	0.024	0.001	0.193	0.042	0.032	0.012	0.196
40	70	0.5	0.25	0.005	0.035	0.001	0.218	0.036	0.044	0.019	0.219
5	70	0.75	0.25	0.001	0.004	0.001	0.262	0.056	0.058	0.026	0.275
20	70	0.75	0.25	0.001	0.014	0.002	0.187	0.038	0.027	0.028	0.192
30	70	0.75	0.25	0.007	0.013	0.005	0.157	0.028	0.021	0.015	0.161
40	70	0.75	0.25	0.001	0.012	0.002	0.173	0.018	0.018	0.013	0.175
5	70	0	0.5	0.162	0.130	0.025	0.363	0.314	0.158	0.057	0.367
20	70	0	0.5	0.132	0.139	0.004	0.370	0.245	0.143	0.031	0.373
30	70	0	0.5	0.190	0.140	0.001	0.353	0.391	0.144	0.020	0.355
40	70	0	0.5	0.149	0.136	0.002	0.358	0.367	0.139	0.017	0.359
5	70	0.25	0.5	0.101	0.081	0.012	0.373	0.166	0.126	0.052	0.379
20	70	0.25	0.5	0.017	0.094	0.004	0.351	0.082	0.097	0.026	0.351
30	70	0.25	0.5	0.008	0.086	0.001	0.346	0.065	0.088	0.015	0.347
40	70	0.25	0.5	0.001	0.098	0.005	0.351	0.057	0.102	0.014	0.352
5	70	0.5	0.5	0.063	0.077	0.001	0.344	0.102	0.122	0.029	0.354
20	70	0.5	0.5	0.010	0.035	0.002	0.293	0.048	0.047	0.011	0.294
30	70	0.5	0.5	0.003	0.037	0.005	0.281	0.038	0.045	0.012	0.283
40	70	0.5	0.5	0.009	0.042	0.003	0.280	0.030	0.043	0.011	0.280
5	70	0	0.75	0.154	0.121	0.007	0.325	0.170	0.127	0.033	0.326
20	70	0	0.75	0.095	0.102	0.002	0.301	0.139	0.105	0.018	0.301
30	70	0	0.75	0.100	0.100	0.008	0.304	0.162	0.102	0.018	0.304
40	70	0	0.75	0.070	0.103	0.003	0.310	0.113	0.103	0.011	0.310
5	70	0.25	0.75	0.057	0.063	0.013	0.309	0.093	0.091	0.037	0.311
20	70	0.25	0.75	0.013	0.051	0.003	0.269	0.041	0.055	0.016	0.270
30	70	0.25	0.75	0.003	0.056	0.000	0.265	0.031	0.059	0.008	0.265
40	70	0.25	0.75	0.006	0.056	0.003	0.263	0.023	0.057	0.007	0.263

Note: see page 13-14 for explanation of the different instrument structures.