GMM estimation of spatial panels

Moscone, Francesco and Tosetti, Elisa

Brunel University

17 April 2009

GMM estimation of spatial panels*

F. Moscone[†]
Brunel Business School

E. Tosetti[‡]
University of Cambridge

Abstract

We consider Generalized Method of Moments (GMM) estimation of a regression model with spatially correlated errors. We propose some new moment conditions, and derive the asymptotic distribution of the GMM based on them. The analysis is supported by a small Monte Carlo exercise.

Keywords: Generalized Method of Moments, spatial econometrics.
JEL Code: C2, C5.

1 Introduction

GMM estimation of spatial regression models has been originally advanced by Kelejian and Prucha (1999). They suggested three moment conditions that exploit the properties of disturbances implied by a standard set of assumptions. Substantial work has followed their original study. Druska and Horrace (2004) have considered GMM estimation of a panel regression with time dummies and time-varying spatial weights. Lee and Liu (2006a) suggested a set of linear and quadratic moment conditions in the errors with inner matrices satisfying certain regularity properties; Lee and Liu (2006b) have extended this framework to estimate regression models with higher-order spatial lags. Fingleton (2008a) and Fingleton (2008b) proposed a GMM estimator for spatial regression models with an endogenous spatial lag and moving average errors. Kelejian and Prucha (2008) have generalized their original work to allow heteroskedasticity and spatial lags in the dependent variable. This has been extended by Kapoor et al. (2007) to estimate a spatial panel regression with individual-specific error components.

We focus on GMM estimation of a regression model where the error follows a spatial autoregressive (SAR) process. We show that there are more moments than those currently exploited in the literature, and derive the asymptotic distribution of the GMM based on such moments. We perform a small Monte Carlo exercise to compare the properties of GMM estimators based on different sets of moments.

2 The framework

Consider the model expressed in stacked form

$$\mathbf{y}_t = \mathbf{X}_t\boldsymbol{\beta} + \mathbf{u}_t, \quad t = 1, ..., T, \tag{1}$$

$$\mathbf{u}_t = \delta\mathbf{S}\mathbf{u}_t + \boldsymbol{\varepsilon}_t, \tag{2}$$

where $\mathbf{y}_t = (y_{1t}, ..., y_{Nt})'$,$\mathbf{X}_t = (\mathbf{x}_{1t}, ..., \mathbf{x}_{Nt})'$,$\mathbf{u}_t = (u_{1t}, ..., u_{Nt})'$, $\boldsymbol{\varepsilon}_t = (\varepsilon_{1t}, ..., \varepsilon_{Nt})'$ and $\mathbf{S}$ is $N \times N$ spatial weights matrix. We assume:

*The authors acknowledge financial support from ESRC (ref. no. RES-061-25-0317). We have benefited from comments by the partecipants of the III World Spatial Econometrics Association.
[†]e-mail: francesco.moscone@brunel.ac.uk.
[‡]e-mail: et268@cam.ac.uk.

Assumption 1: $\varepsilon_{it} \sim IIDN(0, \sigma^2)$, with $\sigma^2 \le K < \infty$, for $i = 1, ..., N, t = 1, ..., T$.

Assumption 2: $\mathbf{X}_t$ and $\boldsymbol{\varepsilon}_{t'}$ are independently distributed for all t, t'. As N and/or $T \to \infty \frac{1}{NT}\sum_{t=1}^{T} \mathbf{X}_t'\mathbf{X}_t \to \mathbf{M}$, where $\mathbf{M}$ is finite and non-singular.

Assumption 3: $\mathbf{S}$ has zero diagonal elements; $\mathbf{S}$ and $(\mathbf{I}_N - \delta\mathbf{S})^{-1}$ have bounded row and column norms.

Assumption 4: $\delta \in \left(-\frac{1}{\rho_{\mathbf{S}}}, \frac{1}{\rho_{\mathbf{S}}}\right)$, where $\rho_{\mathbf{S}} = \max_{1 \le i \le N} \{|\lambda_i(\mathbf{S})|\}$.

Normality and constant variance of ε_{it} stated in Assumption 1 are only taken for ease of exposition, and our results can be readily extended to the case of non-normal, heteroskedastic variables (see Kelejian and Prucha (2008) on this). Assumption 4 implies that (2) can be rewritten as $\mathbf{u}_t = \mathbf{R}\boldsymbol{\varepsilon}_t$,where $\mathbf{R} = (\mathbf{I}_N - \delta\mathbf{S})^{-1}$. Let $\hat{\boldsymbol{\beta}}$ be the OLS estimator of $\boldsymbol{\beta}$. Under Assumptions 1-3, $\hat{\boldsymbol{\beta}}$ is consistent for $\boldsymbol{\beta}$, as N and/or T tends to infinity, but, for $\delta \neq 0$, is not efficient. Efficient estimation of $\boldsymbol{\beta}$ can be achieved by estimating the parameters in equation (2), namely δ and σ^2, and then apply feasible GLS.

3 GMM estimation of SAR processes

Let $\boldsymbol{\theta}_0 = \left(\delta_0, \sigma_0^2\right)'$ be the true parameter vector for (2). Kelejian and Prucha (1999) suggest the following moments for estimating $\boldsymbol{\theta}_0$

$$\mathcal{M}_1(\boldsymbol{\theta}) = E\left(\frac{1}{NT}\sum_{t=1}^{T} \boldsymbol{\varepsilon}_t'\boldsymbol{\varepsilon}_t\right) - \sigma^2 = 0, \tag{3}$$

$$\mathcal{M}_2(\boldsymbol{\theta}) = E\left(\frac{1}{NT}\sum_{t=1}^{T} \boldsymbol{\varepsilon}_t'\mathbf{S}'\mathbf{S}\boldsymbol{\varepsilon}_t\right) - \sigma^2\frac{1}{N}Tr(\mathbf{S}'\mathbf{S}) = 0, \tag{4}$$

$$\mathcal{M}_3(\boldsymbol{\theta}) = E\left(\frac{1}{NT}\sum_{t=1}^{T} \boldsymbol{\varepsilon}_t'\mathbf{S}\boldsymbol{\varepsilon}_t\right) = 0. \tag{5}$$

Moment (3) is implied by the constant variance of $\boldsymbol{\varepsilon}_t$; (4) exploits the variance of the spatial lag, $\mathbf{S}\boldsymbol{\varepsilon}_t$; (5) is based on the covariance between $\boldsymbol{\varepsilon}_t$ and $\mathbf{S}\boldsymbol{\varepsilon}_t$. From (2), the following additional moments can be suggested:

$$\mathcal{M}_4(\boldsymbol{\theta}) = E\left(\frac{1}{NT}\sum_{t=1}^{T} \mathbf{u}_t'\mathbf{u}_t\right) - \sigma^2\frac{1}{N}Tr\left(\mathbf{R}\mathbf{R}'\right) = 0, \tag{6}$$

$$\mathcal{M}_5(\boldsymbol{\theta}) = E\left(\frac{1}{NT}\sum_{t=1}^{T} \mathbf{u}_t'\mathbf{S}'\mathbf{S}\mathbf{u}_t\right) - \sigma^2\frac{1}{N}Tr\left(\mathbf{R}'\mathbf{S}'\mathbf{S}\mathbf{R}\right) = 0, \tag{7}$$

$$\mathcal{M}_6(\boldsymbol{\theta}) = E\left(\frac{1}{NT}\sum_{t=1}^{T} \mathbf{u}_t'\mathbf{S}\mathbf{u}_t\right) - \sigma^2\frac{1}{N}Tr\left(\mathbf{R}'\mathbf{S}\mathbf{R}\right) = 0, \tag{8}$$

$$\mathcal{M}_7(\boldsymbol{\theta}) = E\left(\frac{1}{NT}\sum_{t=1}^{T} \mathbf{u}_t'\boldsymbol{\varepsilon}_t\right) - \sigma^2\frac{1}{N}Tr(\mathbf{R}) = \mathbf{0}, \tag{9}$$

$$\mathcal{M}_8(\boldsymbol{\theta}) = E\left(\frac{1}{NT}\sum_{t=1}^{T} \mathbf{u}_t'\mathbf{S}'\mathbf{S}\boldsymbol{\varepsilon}_t\right) - \sigma^2\frac{1}{N}Tr\left(\mathbf{R}'\mathbf{S}'\mathbf{S}\right) = 0, \tag{10}$$

$$\mathcal{M}_9(\boldsymbol{\theta}) = E\left(\frac{1}{NT}\sum_{t=1}^{T} \mathbf{u}_t'\mathbf{S}\boldsymbol{\varepsilon}_t\right) - \sigma^2\frac{1}{N}Tr(\mathbf{R}'\mathbf{S}) = 0. \tag{11}$$

Moments (6)-(7) exploit the variance of $\mathbf{u}_t$ and $\mathbf{S}\mathbf{u}_t$, respectively; (8), (9) and (11) are based on the covariance of $\mathbf{u}_t$ with $\mathbf{S}\mathbf{u}_t$, $\boldsymbol{\varepsilon}_t$ and $\mathbf{S}\boldsymbol{\varepsilon}_t$, respectively; (10) exploits the covariance between the spatial lags $\mathbf{S}\mathbf{u}_t$ and $\mathbf{S}\boldsymbol{\varepsilon}_t$.

Remark 1 *Under $\delta_0 = 0$, moment (3) would be identical to (6) and (9); (4) would coincide with (7) and (11), and (5) would be the same as (8) and (10). Hence, when δ_0 is zero or close to zero we expect the additional moments (6)-(11) to be redundant.*

In this paper we intend to study the properties of the GMM estimator based on subsets of conditions (3)-(11). We first observe that conditions (3)-(11) contain the following expressions

$$\frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\mathbf{A}_\ell\boldsymbol{\varepsilon}_t,\quad \ell=1,...,9, \tag{12}$$

$$\begin{aligned}\mathbf{A}_1 &= \mathbf{I}_N,\quad \mathbf{A}_2=\mathbf{S}'\mathbf{S},\quad \mathbf{A}_3=\mathbf{S},\mathbf{A}_4=\mathbf{R}'\mathbf{R},\quad \mathbf{A}_5=\mathbf{R}'\mathbf{S}'\mathbf{S}\mathbf{R},\\ \mathbf{A}_6 &= \mathbf{R}'\mathbf{S}\mathbf{R},\mathbf{A}_7=\mathbf{R}',\quad \mathbf{A}_8=\mathbf{R}'\mathbf{S}'\mathbf{S},\quad \mathbf{A}_9=\mathbf{R}'\mathbf{S},\end{aligned}$$

$\mathbf{A}_\ell$ having bounded row and column norms. Under Assumption 1, the mean and variance of (12) are[1]

$$E\left(\frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\mathbf{A}_\ell\boldsymbol{\varepsilon}_t\right)=\frac{1}{N}\sigma^2Tr\left(\mathbf{A}_\ell\right),\quad \ell=1,...,r, \tag{13}$$

$$Var\left(\frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\mathbf{A}_\ell\boldsymbol{\varepsilon}_t\right)=\frac{1}{N^2T}\sigma^4Tr\left(\mathbf{A}_\ell^2+\mathbf{A}_\ell'\mathbf{A}_\ell\right), \tag{14}$$

$$Cov\left[\left(\frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\mathbf{A}_\ell\boldsymbol{\varepsilon}_t\right)\left(\frac{1}{NT}\sum_{t=1}^{T}\boldsymbol{\varepsilon}_t'\mathbf{A}_h\boldsymbol{\varepsilon}_t\right)\right]=\frac{1}{N^2T}\sigma^4Tr\left(\mathbf{A}_\ell\mathbf{A}_h+\mathbf{A}_\ell'\mathbf{A}_h\right),\ \ell\neq h, \tag{15}$$

Let $\hat{u}_{it}=y_{it}-\hat{\boldsymbol{\beta}}'\mathbf{x}_{it}$ and $\hat{\varepsilon}_{it}=\hat{u}_{it}-\delta\sum_{j=1}^{N}s_{ij}\hat{u}_{jt}$. The sample analogues of (3)-(11) can be obtained by replacing $\boldsymbol{\varepsilon}_t$ by $\hat{\boldsymbol{\varepsilon}}_t$ and $\mathbf{u}_t$ by $\hat{\mathbf{u}}_t$. Let $\mathcal{M}\left(\boldsymbol{\theta}\right)=\left[\mathcal{M}_1\left(\boldsymbol{\theta}\right),...,\mathcal{M}_r\left(\boldsymbol{\theta}\right)\right]'$ be a vector containing $r\leq 9$ moments among (3)-(11), and $\mathbf{M}_{NT}\left(\boldsymbol{\theta},\hat{\boldsymbol{\varepsilon}}\right)=\left[M_{NT,1}\left(\boldsymbol{\theta},\hat{\boldsymbol{\varepsilon}}\right),...,M_{NT,r}\left(\boldsymbol{\theta},\hat{\boldsymbol{\varepsilon}}\right)\right]'$ be the corresponding sample moments. Given that $\hat{u}_{it}$ (and hence $\hat{\varepsilon}_{it}$) is based on a consistent estimate of $\boldsymbol{\beta}$, under Assumptions 1-4 the hypotheses of Theorem 1 in Kelejian and Prucha (2001), and Theorem A1 and Lemma C1 in Kelejian and Prucha (2008) are satisfied and

$$\left[\mathbf{M}_{NT}\left(\boldsymbol{\theta},\boldsymbol{\varepsilon}\right)-E\left[\mathbf{M}_{NT}\left(\boldsymbol{\theta},\boldsymbol{\varepsilon}\right)\right]\right]\xrightarrow{p}\mathbf{0},\ \text{as } N \text{ and/or } T \rightarrow \infty \tag{16}$$

$$\left(NT\right)^{1/2}\left[\mathbf{M}_{NT}\left(\boldsymbol{\theta},\hat{\boldsymbol{\varepsilon}}\right)-\mathbf{M}_{NT}\left(\boldsymbol{\theta},\boldsymbol{\varepsilon}\right)\right]\xrightarrow{p}\mathbf{0},\ \text{as } N \text{ and/or } T \rightarrow \infty \tag{17}$$

$$\left(NT\right)^{1/2}\mathbf{V}_r\left(\boldsymbol{\theta}\right)^{-1/2}\mathbf{M}_{NT}\left(\boldsymbol{\theta},\boldsymbol{\varepsilon}\right)\overset{a}{\sim}N(\mathbf{0},\mathbf{I}_r),\ \text{as } N \text{ and/or } T \rightarrow \infty \tag{18}$$

where $\mathbf{V}_r\left(\boldsymbol{\theta}\right)=E\left[\mathbf{M}_{NT}\left(\boldsymbol{\theta},\boldsymbol{\varepsilon}\right)\mathbf{M}_{NT}\left(\boldsymbol{\theta},\boldsymbol{\varepsilon}\right)'\right]$ is assumed to be non-singular, i.e. $\lambda_r\left(\mathbf{V}_r\left(\boldsymbol{\theta}\right)\right)\geq K>0$. $\mathbf{V}_r\left(\boldsymbol{\theta}\right)$ has (14) on its main diagonal, and (15) as off-diagonal elements. The above results hold for N and/or T going to infinity. The asymptotic in T can be proved by applying standard multivariate law of large numbers and central limit theorem, since under Assumptions 1-2 $\hat{\boldsymbol{\varepsilon}}_t'\mathbf{A}_\ell\hat{\boldsymbol{\varepsilon}}_t$ (and $\boldsymbol{\varepsilon}_t'\mathbf{A}_\ell\boldsymbol{\varepsilon}_t$), for $t=1,...,T$, are IID. The above results have been used by Kelejian and Prucha (2008) and Kelejian and Prucha (1999) to prove asymptotic normality of the GMM based on (3)-(5). We now show that this result continues to hold if the GMM is based on any subsets of (3)-(11) such that the covariance matrix of the corresponding sample moments is non-singular.

3.1 Estimation

Suppose we select r moments from (3)-(11), such that $\mathbf{V}_r\left(\boldsymbol{\theta}\right)$ is non-singular, and let $\mathbf{M}_{NT}(\boldsymbol{\theta},\hat{\boldsymbol{\varepsilon}})$ be the vector of their sample analogues. The GMM estimator $\hat{\boldsymbol{\theta}}=\left(\hat{\delta},\hat{\sigma}^2\right)'$ is the solution to the following optimization problem

$$\hat{\boldsymbol{\theta}}=\underset{\boldsymbol{\theta}\in\boldsymbol{\Theta}}{\arg\min}\left\{\mathbf{M}_{NT}(\boldsymbol{\theta},\hat{\boldsymbol{\varepsilon}})'\mathbf{Q}_{NT}\mathbf{M}_{NT}(\boldsymbol{\theta},\hat{\boldsymbol{\varepsilon}})\right\}, \tag{19}$$

where $\boldsymbol{\Theta}$ is the parameter space[2], and $\mathbf{Q}_{NT}$ is a $r\times r$, positive-definite weighting matrix satisfying $\mathbf{Q}_{NT}\xrightarrow{p}\mathbf{Q}$.The following theorem establishes the asymptotic distribution of $\hat{\boldsymbol{\theta}}$.

[1] See Ullah (2004). These results hold under normality of ε_{it}, but they can be easily extended to the non-normal case.

[2] Notice that under Assumption 1 and 4 $\boldsymbol{\Theta}$ is a compact interval.

Theorem 2 *Under Assumptions 1-4, $\hat{\boldsymbol{\theta}}$ in (19) is consistent for $\boldsymbol{\theta}_0$ and, as N and/or $T \to \infty$,*

$$(NT)^{1/2}\left(\hat{\boldsymbol{\theta}}-\boldsymbol{\theta}_0\right) \overset{a}{\sim} N\left(\mathbf{0},\left(\mathbf{D}'\mathbf{Q}\mathbf{D}\right)^{-1}\mathbf{D}'\mathbf{Q}\mathbf{Q}^*\left(\boldsymbol{\theta}_0\right)\mathbf{Q}\mathbf{D}\left(\mathbf{D}'\mathbf{Q}\mathbf{D}\right)^{-1}\right), \tag{20}$$

where $\mathbf{D}=\mathbf{D}\left(\boldsymbol{\theta}_0,\boldsymbol{\varepsilon}\right)=p\lim \frac{\partial}{\partial\boldsymbol{\theta}}\mathbf{M}_{NT}\left(\boldsymbol{\theta}_0,\boldsymbol{\varepsilon}\right)$.

The efficient GMM estimator can be obtained by imposing $\mathbf{Q}=\mathbf{Q}^*\left(\boldsymbol{\theta}_0\right)^{-1}$, where

$$\mathbf{Q}^*\left(\boldsymbol{\theta}_0\right)=\left\{E\left[NT\ \mathbf{M}_{NT}(\boldsymbol{\theta}_0,\boldsymbol{\varepsilon})\mathbf{M}'_{NT}(\boldsymbol{\theta}_0,\boldsymbol{\varepsilon})\right]\right\} \tag{21}$$

is the optimal weighting matrix. Notice that, under Assumption 1, $\mathbf{Q}^*\left(\boldsymbol{\theta}\right)=NT\mathbf{V}\left(\boldsymbol{\theta}\right)$, and therefore $\mathbf{Q}^*\left(\boldsymbol{\theta}\right)$ has as (ℓ,h) element expression (15) multiplied by NT. In practise, $\mathbf{Q}$ and $\mathbf{D}$ are evaluated at point estimates, $\mathbf{Q}^*\left(\hat{\boldsymbol{\theta}}\right)^{-1}$ and $\mathbf{D}\left(\hat{\boldsymbol{\theta}},\hat{\varepsilon}\right)$. In the Appendix we sketch the proof of consistency of $\hat{\boldsymbol{\theta}}$ and derive $\mathbf{D}$, and refer to Kelejian and Prucha (2008), Kelejian and Prucha (1999) for further details on consistency of GMM estimators of spatial models. We do not report the proof of asymptotic normality of $\hat{\boldsymbol{\theta}}$ since, once established (16)-(18), this is identical to that in Kelejian and Prucha (2008). When conditions (3)-(5) are employed in (19) $\mathbf{Q}^*\left(\boldsymbol{\theta}_0\right)$ is

$$\mathbf{Q}^{*KP}_{NT}\left(\boldsymbol{\theta}\right)=\frac{\sigma^4}{N}\left(\begin{array}{lll} 2N & 2Tr\left(\mathbf{S}'\mathbf{S}\right) & 0 \\ 2Tr\left(\mathbf{S}'\mathbf{S}\right) & 2Tr\left[\left(\mathbf{S}'\mathbf{S}\right)^2\right] & 2Tr\left(\mathbf{S}'\mathbf{S}^2\right) \\ 0 & 2Tr\left(\mathbf{S}'\mathbf{S}^2\right) & Tr\left(\mathbf{S}^2+\mathbf{S}'\mathbf{S}\right) \end{array}\right).$$

Since σ^2 enters in $\mathbf{Q}^{*KP}_{NT}\left(\boldsymbol{\theta}\right)$ only as a scale factor, we can compute $\hat{\boldsymbol{\theta}}$ in a single step by minimizing (19). However, in general, δ and σ^2 do enter in the formula for $\mathbf{Q}^*_{NT}\left(\boldsymbol{\theta}\right)$. In this case estimation can proceed adopting a two-stage iterative procedure where in the first stage we minimize (19) using $\mathbf{Q}=\mathbf{I}_r$, and OLS residuals $\hat{u}_{it}$, and in the second stage, we employ $\hat{\boldsymbol{\theta}}$ to compute $\mathbf{Q}^*_{NT}\left(\hat{\boldsymbol{\theta}}\right)$ and use it in (19). Once estimated $\boldsymbol{\theta}$, efficient estimation of $\boldsymbol{\beta}$ can be obtained by applying feasible GLS. We next run a Monte Carlo exercise to evaluate and compare the small sample properties of GMM estimators based on subsets of (3)-(11).

4 Monte Carlo experiments

We consider:

$$\begin{aligned} y_{it} &= 1+x_{1,it}+x_{2,it}+u_{it}, i=1,...,N; t=1,...,T, \\ x_{\ell,it} &= 0.6x_{\ell,it-1}+v_{\ell it}, v_{\ell,it}\sim IIDN(0,1-0.6^2), \\ u_{it} &= \delta\sum_{j=1}^{N}s_{ij}u_{jt}+\varepsilon_{it}, \quad \varepsilon_{it}\sim IIDN(0,1). \end{aligned}$$

The values of $x_{\ell,it}$ and u_{it} are drawn for each i and t, and at each replication. $\mathbf{S}$ is a regular lattice where each unit has two adjacent neighbours and set $s_{ij}=1$ if i and j are adjacent and $s_{ij}=0$ otherwise; $\mathbf{S}$ is row-standardized. We experimented with $\delta=0.0, 0.4, 0.8$, and provide results for the following estimators of $\boldsymbol{\theta}=\left(\delta,\sigma^2\right)'$ (adopting (21)): $\hat{\boldsymbol{\theta}}^{KP}_{GMM}$, based on (3)-(5); $\hat{\boldsymbol{\theta}}^{(1)}_{GMM}$, based on (6)-(8); $\hat{\boldsymbol{\theta}}^{(2)}_{GMM}$, based on (9)-(11); and $\hat{\boldsymbol{\theta}}^{(3)}_{GMM}$, based on (3)-(11). Estimation of $\boldsymbol{\theta}$ is performed on $\hat{u}_{it}=y_{it}-\hat{\alpha}-\hat{\beta}_1x_{1,it}-\hat{\beta}_2x_{2,it}$. We assess the performance of estimators by computing their bias, RMSE, size and power (at 5% significance level). We ran $1,000$ replications for all pairs $N=10,20,50; T=5,10$.

Table 1 shows results for estimators of δ. For purpose of comparison, we also provide the quasi-ML estimator of δ, $\hat{\delta}_{ML}$. The bias and RMSE of $\hat{\delta}^{KP}_{GMM}$ decrease as N and/or T get large, for all values of δ. The size of $\hat{\delta}^{KP}_{GMM}$ is close to the nominal 5% level for $\delta=0.0, 0.4$, for all N,T larger than 10; while it deviates from the 5% level when $T=5$. When $\delta=0.8$, the empirical rejection frequencies are slightly larger than the nominal 5% level. We observe that, for a given pair of N and T, larger values of δ are associated to smaller RMSEs and higher power of $\hat{\delta}_{GMM}^{KP}$. A similar pattern can be observed for the GMM estimator based on other sets of conditions (i.e., $\hat{\delta}_{1,GMM}$, $\hat{\delta}_{2,GMM}$ and $\hat{\delta}_{3,GMM}$), and for $\hat{\delta}_{ML}$. However, some important differences in the performance of these estimators can be noted. First, $\hat{\delta}_{2,GMM}$ performs overall better than $\hat{\delta}_{GMM}^{KP}$: its bias (in absolute value) and RMSE are lower than those for $\hat{\delta}_{GMM}^{KP}$, for all values of δ, and the size is very close to 5%, for $\delta = 0.0, 0.4$. In the case $\delta = 0.8$, $\hat{\delta}_{2,GMM}$ is slightly oversized when N and T are small. Notice that results for $\hat{\delta}_{2,GMM}$ are very close to outputs for $\hat{\delta}_{ML}$ in all cases considered. The performance of $\hat{\delta}_{1,GMM}$ and $\hat{\delta}_{3,GMM}$ is similar to that of $\hat{\delta}_{2,GMM}$ when the $\delta = 0.0, 0.4$. For $\delta = 0.8$, $\hat{\delta}_{1,GMM}$ presents some distortions and is characterized by rejections frequencies larger than 5%, ranging between 6.10% and 14.80%. $\hat{\delta}_{3,GMM}$ has bias and RMSE similar to $\hat{\delta}_{2,GMM}$, while its size deviates from the 5% level. An explanation for this result is that some moments used in computing $\hat{\delta}_{3,GMM}$ might be highly correlated, leading to a nearly-singular $\mathbf{Q}^*$ matrix.

5 Conclusions

We have introduced new moments in a GMM estimation of a spatial regression model. Given that when $\delta = 0$ some of the suggested moments are redundant, we have proposed to use only a subset of the moments in the estimation procedure. Our Monte Carlo experiments point at conditions (9)-(11) as those that yield the best performance of the GMM estimator.

References

Druska, V. and W. C. Horrace (2004). Generalized moments estimation for spatial panels: indonesian rice farming. *American Journal of Agricultural Economics 86*, 185–198.

Fingleton, B. (2008a). A generalized method of moments estimator for a spatial model with moving average errors, with application to real estate prices. *Empirical Economics 34*, 35Ű57.

Fingleton, B. (2008b). A generalized method of moments estimator for a spatial panel model with an endogenous spatial lag and spatial moving average errors. *Spatial Economic Analysis 3*, 27–44.

Kapoor, M., H. H. Kelejian, and I. Prucha (2007). Panel data models with spatially correlated error components. *Journal of Econometrics 140*, 97–130.

Kelejian, H. H. and I. Prucha (1999). A generalized moments estimator for the autoregressive parameter in a spatial model. *International Economic Review 40*, 509–533.

Kelejian, H. H. and I. Prucha (2001). On the asymptotic distribution of the moran i test with applications. *Journal of Econometrics 104*, 219–257.

Kelejian, H. H. and I. Prucha (2008). Specification and estimation of spatial autoregressive models with autoregressive and heteroskedastic disturbances. Forthcoming, Journal of Econometrics.

Lee, L. F. and X. Liu (2006a). Efficient GMM estimation of a SAR model with autoregressive disturbances. Mimeo.

Lee, L. F. and X. Liu (2006b). Efficient GMM estimation of high order spatial autoregressive models. Mimeo.

Ullah, A. (2004). *Finite Sample Econometrics*. Oxford: Oxford University Press.

6 Appendix

Let $\mathcal{M}(\boldsymbol{\theta}) = [\mathcal{M}_1(\boldsymbol{\theta}), ..., \mathcal{M}_r(\boldsymbol{\theta})]'$ and $\mathbf{M}_{NT}(\boldsymbol{\theta}, \hat{\boldsymbol{\varepsilon}}) = [M_{NT,1}(\boldsymbol{\theta}, \hat{\boldsymbol{\varepsilon}}), ..., M_{NT,r}(\boldsymbol{\theta}, \hat{\boldsymbol{\varepsilon}})]'$, and

$$R(\boldsymbol{\theta}, \hat{\boldsymbol{\varepsilon}}) = \mathbf{M}_{NT}(\boldsymbol{\theta}, \hat{\boldsymbol{\varepsilon}})' \mathbf{Q}_{NT}^*(\boldsymbol{\theta}) \mathbf{M}_{NT}(\boldsymbol{\theta}, \hat{\boldsymbol{\varepsilon}}), \qquad Z(\boldsymbol{\theta}) = \mathcal{M}(\boldsymbol{\theta})' \mathbf{Q}_{NT}^*(\boldsymbol{\theta}_0) \mathcal{M}(\boldsymbol{\theta}).$$

Consistency of the GMM can be showed by proving:

(I) *Identification uniqueness*: for all N, T, and for $K > 0 : \inf_{\boldsymbol{\theta}:\|\boldsymbol{\theta}-\boldsymbol{\theta}_0\|_2 \geq K} |Z(\boldsymbol{\theta}) - Z(\boldsymbol{\theta}_0)| > 0$.

(II) *Uniform convergence*: $\lim_{N,T\to\infty} \sup_{\boldsymbol{\theta}} |R(\boldsymbol{\theta}, \hat{\boldsymbol{\varepsilon}}) - Z(\boldsymbol{\theta})| = 0$

To prove (I), note that

$$\mathcal{M}(\boldsymbol{\theta}) = \boldsymbol{\Gamma}\boldsymbol{\varphi}(\boldsymbol{\theta}) - \mathbf{D}\boldsymbol{\gamma}, \qquad \mathbf{M}_{NT}(\boldsymbol{\theta}, \hat{\boldsymbol{\varepsilon}}) = \mathbf{G}_{NT}(\hat{\boldsymbol{\varepsilon}})\,\boldsymbol{\varphi}(\boldsymbol{\theta}) - \mathbf{D}\mathbf{g}_{NT}(\hat{\boldsymbol{\varepsilon}}),$$

where (here we provide $\boldsymbol{\Gamma}, \boldsymbol{\varphi}(\boldsymbol{\theta}), \mathbf{D}, \boldsymbol{\gamma}$ when $\mathcal{M}(\boldsymbol{\theta})$ contains (3)-(11), but we recall that our analysis is based on subsets of these moments)

$$\boldsymbol{\Gamma} = \begin{bmatrix} -2E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{S}\mathbf{u}_t\right) & E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t\mathbf{S}'\mathbf{S}\mathbf{u}_t\right) & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{S}'\mathbf{S}^2\mathbf{u}_t\right) & E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{S}'^2\mathbf{S}^2\mathbf{u}_t\right) & -\frac{1}{N}Tr\left(\mathbf{S}'\mathbf{S}\right) & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{S}'\mathbf{S}\mathbf{u}_t\right) & E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{S}'^2\mathbf{S}\mathbf{u}_t\right) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ -E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{S}\mathbf{u}_t\right) & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ -E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{S}'\mathbf{S}^2\mathbf{u}_t\right) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ -E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{S}^2\mathbf{u}_t\right) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

$$\boldsymbol{\varphi}(\boldsymbol{\theta}) = \begin{bmatrix} \delta & \delta^2 & \sigma^2 & \frac{\sigma^2}{N}Tr\left(\mathbf{R}\mathbf{R}'\right) & \frac{\sigma^2}{N}Tr(\mathbf{R}'\mathbf{S}'\mathbf{S}\mathbf{R}) & \frac{\sigma^2}{N}Tr\left(\mathbf{R}'\mathbf{S}\mathbf{R}\right) & \frac{\sigma^2}{N}Tr(\mathbf{R}) & \frac{\sigma^2}{N}Tr\left(\mathbf{R}'\mathbf{S}'\mathbf{S}\right) & \frac{\sigma^2}{N}Tr(\mathbf{R}'\mathbf{S}) \end{bmatrix}',$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{I}_3 & \mathbf{I}_3 & \mathbf{I}_3 \end{bmatrix}', \boldsymbol{\gamma} = \begin{bmatrix} E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{u}_t\right) & E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{S}'\mathbf{S}\mathbf{u}_t\right) & E\left(\frac{1}{NT}\sum_{t=1}^{T}\mathbf{u}_t'\mathbf{S}\mathbf{u}_t\right) \end{bmatrix}',$$

and $\mathbf{G}_{NT}(\hat{\boldsymbol{\varepsilon}})$, $\mathbf{g}_{NT}(\hat{\boldsymbol{\varepsilon}})$ are the sample analogues of $\boldsymbol{\Gamma}$, $\boldsymbol{\gamma}$. Following Kelejian and Prucha (1999), the proof of consistency requires the following assumption:

Assumption 5: $\boldsymbol{\Gamma}'\boldsymbol{\Gamma}$ is non-singular, i.e. its smallest eigenvalue $\lambda_r(\boldsymbol{\Gamma}'\boldsymbol{\Gamma}) > 0$.

We have, for $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|_2 \geq K > 0$,

$$\begin{aligned} &\mathcal{M}(\boldsymbol{\theta})'\,\mathbf{Q}_{NT}^*(\boldsymbol{\theta}_0)\,\mathcal{M}(\boldsymbol{\theta}) - \mathcal{M}(\boldsymbol{\theta}_0)'\,\mathbf{Q}_{NT}^*(\boldsymbol{\theta}_0)\,\mathcal{M}(\boldsymbol{\theta}_0) = \left[\boldsymbol{\varphi}(\boldsymbol{\theta}) - \boldsymbol{\varphi}(\boldsymbol{\theta_0})\right]'\boldsymbol{\Gamma}'\mathbf{Q}_{NT}^*(\boldsymbol{\theta}_0)\,\boldsymbol{\Gamma}\left[\boldsymbol{\varphi}(\boldsymbol{\theta}) - \boldsymbol{\varphi}(\boldsymbol{\theta_0})\right] \\ \geq\ & \lambda_r\left(\mathbf{Q}_{NT}^*(\boldsymbol{\theta}_0)\right)\lambda_r\left(\boldsymbol{\Gamma}\boldsymbol{\Gamma}'\right)\left[\boldsymbol{\varphi}(\boldsymbol{\theta}) - \boldsymbol{\varphi}(\boldsymbol{\theta_0})\right]'\left[\boldsymbol{\varphi}(\boldsymbol{\theta}) - \boldsymbol{\varphi}(\boldsymbol{\theta_0})\right] > 0 \end{aligned}$$

If moments (6)-(8) alone are included in the analysis, we need to take the following identificability conditions:

Assumption 6**:** $\varphi_r(\boldsymbol{\theta}) - \varphi_r(\boldsymbol{\theta_0}) \neq 0$, for all $\boldsymbol{\theta}$ such that $\|\boldsymbol{\theta} - \boldsymbol{\theta}_0\|_2 \geq K > 0$, and $r = 4, ..., 9$.

To prove (II), let $\boldsymbol{\Pi} = \left[\mathbf{G}_{NT}(\hat{\boldsymbol{\varepsilon}}), -\mathbf{D}\mathbf{g}_{NT}(\hat{\boldsymbol{\varepsilon}})\right]$ and $\mathbf{P} = [\boldsymbol{\Gamma}, -\mathbf{D}\boldsymbol{\gamma}]$, and notice that

$$R(\boldsymbol{\theta}, \hat{\boldsymbol{\varepsilon}}) - Z(\boldsymbol{\theta}) \leq \left[\boldsymbol{\Pi}'\mathbf{Q}_{NT}^*(\boldsymbol{\theta})\,\boldsymbol{\Pi} - \mathbf{P}'\mathbf{Q}_{NT}^*(\boldsymbol{\theta})\,\mathbf{P}\right]\|\boldsymbol{\varphi}(\boldsymbol{\theta})\|_2 \xrightarrow{p} 0$$

since, under Assumptions 1-4, and given (16), $\boldsymbol{\Pi} \xrightarrow{p} \mathbf{P}$, and the elements of $\boldsymbol{\varphi}(\boldsymbol{\theta})$ are bounded.

6.1 The D matrix

$$\mathbf{D} = \begin{bmatrix} -\frac{2}{NT}\sum_{t=1}^{T}\hat{\mathbf{u}}_t'\mathbf{S}'\hat{\boldsymbol{\varepsilon}}_t & -1 \\ -\frac{2}{NT}\sum_{t=1}^{T}\hat{\mathbf{u}}_t'\mathbf{S}^{2\prime}\mathbf{S}\hat{\boldsymbol{\varepsilon}}_t & -\frac{1}{N}Tr\left(\mathbf{S}'\mathbf{S}\right) \\ -\frac{1}{NT}\sum_{t=1}^{T}\hat{\mathbf{u}}_t'\mathbf{S}'\mathbf{S}\hat{\boldsymbol{\varepsilon}}_t - \frac{1}{N}\hat{\boldsymbol{\varepsilon}}_t'\mathbf{S}^2\hat{\mathbf{u}}_t & 0 \\ -\frac{\sigma^2}{N}Tr\left(\mathbf{R}\mathbf{S}\mathbf{R}\mathbf{R}' + \mathbf{R}\mathbf{R}'\mathbf{S}'\mathbf{R}'\right) & -\frac{1}{N}Tr\left(\mathbf{R}\mathbf{R}'\right) \\ -\frac{\sigma^2}{N}Tr\left(\mathbf{R}'\mathbf{S}'\mathbf{R}'\mathbf{S}'\mathbf{S}\mathbf{R} + \mathbf{R}'\mathbf{S}'\mathbf{S}\mathbf{R}\mathbf{S}\mathbf{R}\right) & -\frac{1}{N}Tr(\mathbf{R}'\mathbf{S}'\mathbf{S}\mathbf{R}) \\ -\frac{\sigma^2}{N}Tr\left(\mathbf{R}'\mathbf{S}'\mathbf{R}'\mathbf{S}\mathbf{R} + \mathbf{R}'\mathbf{S}\mathbf{R}\mathbf{S}\mathbf{R}\right) & -\frac{1}{N}Tr\left(\mathbf{R}'\mathbf{S}\mathbf{R}\right) \\ -\frac{1}{NT}\sum_{t=1}^{T}\hat{\mathbf{u}}_t'\mathbf{S}\hat{\mathbf{u}}_t - \frac{\hat{\sigma}^2}{N}Tr\left(\mathbf{R}\mathbf{S}\mathbf{R}\right) & -\frac{1}{N}Tr(\mathbf{R}) \\ -\frac{1}{NT}\sum_{t=1}^{T}\hat{\mathbf{u}}_t'\mathbf{S}^2\hat{\mathbf{u}}_t - \frac{\hat{\sigma}^2}{N}Tr\left(\mathbf{R}'\mathbf{S}'\mathbf{R}'\mathbf{S}\right) & -\frac{1}{N}Tr(\mathbf{R}'\mathbf{S}) \\ -\frac{1}{NT}\sum_{t=1}^{T}\hat{\mathbf{u}}_t'\mathbf{S}'\mathbf{S}^2\hat{\mathbf{u}}_t - \frac{\hat{\sigma}^2}{N}Tr\left(\mathbf{R}'\mathbf{S}'\mathbf{R}'\mathbf{S}'\mathbf{S}\right) & -\frac{1}{N}Tr\left(\mathbf{R}'\mathbf{S}'\mathbf{S}\right) \end{bmatrix}$$

where $\mathbf{R}$ is evaluated at $\hat{\delta}$.

Table 1: Small sample properties (X100) of GMM and ML estimates of δ

		$\delta = 0.0$				$\delta = 0.4$				$\delta = 0.8$			
N	T	Bias	RMSE	Size	Power	Bias	RMSE	Size	Power	Bias	RMSE	Size	Power
ML estimation													
10	5	-1.58	13.66	5.40	9.90	-4.49	13.00	5.70	9.00	-6.72	10.06	8.70	15.10
10	10	-0.94	9.57	4.60	14.30	-2.38	8.81	4.50	16.30	-3.53	6.06	7.70	37.60
20	5	-1.69	9.98	5.60	12.70	-2.94	9.27	6.00	15.30	-3.76	6.31	8.80	34.40
20	10	-0.19	7.12	4.20	28.80	-0.99	6.32	4.80	35.70	-1.80	3.81	6.80	70.30
50	5	-0.78	6.42	6.00	31.00	-1.33	5.71	5.70	37.70	-1.67	3.37	8.20	80.60
50	10	-0.16	4.26	4.70	60.40	-0.46	3.78	4.70	70.20	-0.74	2.09	6.00	99.00
$\hat{\delta}^{KP}_{GMM}$													
10	5	-1.45	14.44	7.30	11.00	-3.80	14.29	7.50	11.00	-6.09	10.16	8.60	17.30
10	10	-0.88	9.70	5.00	15.20	-2.19	8.85	4.60	17.90	-3.35	6.05	7.10	36.80
20	5	-1.76	10.12	5.60	12.90	-2.88	9.36	6.00	15.60	-3.62	6.36	8.90	34.00
20	10	0.17	7.18	4.60	28.70	-0.93	6.37	4.90	35.30	-1.74	3.89	6.60	68.90
50	5	-0.78	6.44	5.80	31.50	-1.29	5.74	5.60	38.60	-1.60	3.45	7.90	78.80
50	10	-0.16	4.26	5.00	60.40	-0.45	3.79	5.20	69.80	-0.70	2.13	6.10	99.00
$\hat{\delta}_{1,GMM}$													
10	5	-1.56	13.33	4.60	9.10	-5.33	13.08	6.00	7.60	-8.99	12.4	14.80	11.30
10	10	-0.92	9.45	4.50	14.10	-2.85	8.89	4.60	14.90	-4.83	7.36	12.00	30.90
20	5	-1.67	9.86	5.40	12.30	-3.36	9.38	6.10	13.90	-4.75	7.40	12.40	29.20
20	10	-0.18	7.08	3.80	28.50	-1.25	6.34	4.60	34.00	-2.27	4.16	8.90	65.90
50	5	-0.78	6.38	5.80	30.90	-1.51	5.73	5.60	36.10	-1.97	3.61	9.70	76.60
50	10	-0.16	4.25	4.70	60.50	-0.55	3.78	4.70	69.10	-0.88	2.16	6.10	99.00
$\hat{\delta}_{2,GMM}$													
10	5	-1.58	13.91	6.10	10.20	-4.21	13.05	6.00	10.10	-6.42	9.87	7.80	16.50
10	10	-0.92	9.63	4.70	14.50	-2.22	8.79	4.60	17.00	-3.34	5.95	7.30	38.50
20	5	-1.73	10.04	6.00	12.60	-2.84	9.27	5.90	16.00	-3.53	6.53	9.10	35.40
20	10	-0.16	7.15	4.40	28.60	-0.91	6.33	4.70	36.20	-1.73	3.80	7.00	70.90
50	5	-0.77	6.43	6.00	31.50	-1.28	5.71	5.60	38.60	-1.63	3.35	7.60	80.60
50	10	-0.17	4.26	4.80	60.60	-0.44	3.78	4.80	70.30	-0.72	2.09	5.90	98.90
$\hat{\delta}_{3,GMM}$													
10	5	-1.18	12.79	4.60	7.60	-7.44	14.74	8.10	8.40	-7.91	11.62	11.50	15.40
10	10	-0.82	9.01	3.70	10.50	-4.78	10.41	8.10	12.60	-3.95	6.77	10.80	35.30
20	5	-1.51	9.01	4.30	9.60	-5.12	10.82	9.60	13.60	-4.21	6.95	11.70	33.20
20	10	-0.08	7.27	5.10	26.00	-2.57	7.39	9.20	28.60	-2.19	4.15	9.20	66.80
50	5	-0.90	7.47	4.80	26.80	-2.68	6.81	10.70	31.20	-1.93	3.61	9.20	77.30
50	10	0.04	5.53	4.40	66.00	-1.40	4.43	9.60	60.90	-1.04	2.30	8.80	98.40