Regional Science and Urban Economics 41 (2011) 487–497

Contents lists available at ScienceDirect

Regional Science and Urban Economics

journal homepage: www.elsevier.com/locate/regec

GMM estimation of spatial panels with fixed effects and unknown heteroskedasticity☆

F. Moscone *, E. Tosetti

Brunel University, United Kingdom

ARTICLE INFO

Article history:
Received 2 November 2010
Received in revised form 13 March 2011
Accepted 20 March 2011
Available online 26 March 2011

JEL classifications:
C2
C5

Keywords:
Spatial panels
Fixed effects
GMM

ABSTRACT

In this paper we consider the estimation of a panel data regression model with spatial autoregressive disturbances, fixed effects and unknown heteroskedasticity. Following the work by Kelejian and Prucha (1999), Lee and Liu (2006a) and others, we adopt the Generalized Method of Moments (GMM) and consider moments as a set of linear quadratic conditions in the disturbances. As in Lee and Liu (2006a), we assume that the inner matrices in the quadratic forms have zero diagonal elements to robustify moments against unknown heteroskedasticity. We derive the asymptotic distribution of the GMM estimator based on such conditions. Hence, we carry out some Monte Carlo experiments to investigate the small sample properties of GMM estimators based on various sets of moment conditions.

1. Introduction

In recent years, there has been a growing literature in economics and econometrics, both applied and theoretical, dealing with spatial issues. One important area of research is estimation and inference in the context of regression models with spatially correlated disturbances. Estimation of these models can be achieved by applying parametric methods such as the maximum likelihood (ML) approach (Anselin, 1988; Mardia and Marshall, 1984) and the generalized method of moments (GMM) (Kelejian and Prucha, 2010; Kelejian and Prucha, 1999; Conley, 1999), or semi-parametric techniques like the spatial HAC estimator by (Kelejian and Prucha, 2007).

GMM estimation of spatial regression models in a single cross sectional setting has been originally advanced by (Kelejian and Prucha, 1999). They focused on a regression equation with spatial autoregressive (SAR) disturbances, and suggested the use of three moment conditions that exploit the properties of disturbances implied by a standard set of assumptions. Estimation consists of solving a non-linear optimization problem, which yields a consistent estimator under a number of regularity conditions. Recently, considerable work has been carried out to extend the procedure advanced by Kelejian and Prucha in various directions. Liu et al. (2006) and Lee and Liu (2006a) suggested a set of moments that encompass Kelejian and Prucha conditions as special cases. They considered a vector of linear and quadratic conditions in the error term, where the matrices appearing in the linear and quadratic forms have bounded row and column norms (see also Lee, 2007). Hence, they focused on the problem of selecting the matrices appearing in the vector of linear and quadratic moment conditions, in order to obtain the lowest variance for the GMM estimator. Lin and Lee (2010) also showed that these moments can be made robust against unknown heteroskedasticity by imposing that the diagonal elements of inner matrices are zero. Lee and Liu (2006b) have extended this framework to estimate the SAR model with higher-order spatial lags. Kelejian and Prucha (2010) have generalized their original work to include spatial lags in the dependent variable as well as allowing for heteroskedastic disturbances. This setting has been extended by Kapoor et al. (2007) to estimate a spatial panel regression model with individual-specific error components. Druska and Horrace (2004) have introduced the Keleijan and Prucha GMM within the framework of a panel with SAR disturbances, time dummies and time-varying spatial weights, while Fingleton (2008a, 2008b) have extended it to the case of a regression model with spatial moving average disturbances.

In this paper, we focus on GMM estimation of a panel data regression model with fixed effects, unknown heteroskedasticity, and spatial autoregressive (SAR) errors. Little work exists on estimation of spatial panels with fixed effects. Quasi-maximum likelihood (ML) estimation of a panel with fixed effects and spatial lags both in the dependent variable and in the disturbances, under homoskedastic errors, has been developed by Lee and Yu (2010). The authors propose a transformation approach to eliminate the fixed effects that yields consistent estimators for regression parameters when either N or T are large. Yu et al. (2008) and Yu et al. (2007) have investigated the properties of the quasi-ML estimator of a

☆ The authors acknowledge financial support from the ESRC (Ref. no. RES-061-25-0317).
* Corresponding author.
E-mail address: francesco.moscone@brunel.ac.uk (F. Moscone).

0166-0462/$ – see front matter © 2011 Published by Elsevier B.V.
doi:10.1016/j.regsciurbeco.2011.03.007

spatial dynamic panel with fixed effects, possibly non-stationary. Mutl and Pfaffermayr (2011) consider GMM estimation of fixed effects vs random effects spatial panel specifications. Hence, they propose a spatial Hausman test that compares the two models, accounting for spatial autocorrelation in the disturbances.

Following the work by Kelejian and Prucha (1999) and Lee and Liu (2006a), in this paper we adopt the GMM and consider as moments a set of quadratic conditions in the disturbances. As in Lee and Liu (2006a), we assume that the inner matrices in the quadratic forms have zero diagonal elements to robustify moments against unknown heteroskedasticity. To eliminate the fixed effects, we transform the data by applying the demeaning operator. We contribute to the existing literature on GMM estimator of spatial models in two ways. First, we investigate the statistical properties of the estimated spatial autoregressive parameter when data have been transformed by demeaning operator to get rid of the fixed effects. We show that consistency and asymptotic normality is achieved for N and/or T going to infinity. We then carry a small Monte Carlo study to compare the small sample properties of GMM estimators based on alternative choices of the inner matrices in the moment conditions and the quasi-ML estimator. Our results show that the GMM estimator has good small sample properties when compared to the performance of the quasi-ML, especially when T is relatively small, and when the spatial parameter is close to 1. Further, when adopting the inner matrix suggested by Liu et al. (2006), the GMM estimator performs better respect to the same estimator based on other quadratic conditions.

In the following, Section 2 sets out the framework of a regression model with SAR disturbances; Section 3 introduces the GMM estimator; Section 4 carries a small Monte Carlo exercise; Section 4.2 concludes.

2. The framework

Consider the panel data regression model

$$y_{it} = \alpha_i + \beta' \mathbf{x}_{it} + u_{it}, \quad i = 1, 2, \ldots, N, t = 1, 2, \ldots, T \tag{1}$$

where α_i are fixed unknown parameters, and errors are assumed to follow the SAR process

$$u_{it} = \delta \sum_{j=1}^{N} s_{ij} u_{jt} + \varepsilon_{it} \tag{2}$$

and s_{ij} is the $(i,j)^{th}$ element of an $N \times N$ spatial weights matrix, $\mathbf{S}$. In matrix form,

$$\mathbf{y} = (\mathbf{1}_T \otimes \boldsymbol{\alpha}) + \mathbf{X}\beta + \mathbf{u}, \tag{3}$$

$$\mathbf{u} = \delta(\mathbf{I}_T \otimes \mathbf{S})\mathbf{u} + \varepsilon, \tag{4}$$

where $\mathbf{y} = (\mathbf{y}'_1, \ldots, \mathbf{y}'_T)'$, $\mathbf{X} = (\mathbf{X}'_1, \ldots, \mathbf{X}'_T)'$, $\mathbf{u} = (\mathbf{u}'_1, \ldots, \mathbf{u}'_T)'$, and $\varepsilon = (\varepsilon'_1, \ldots, \varepsilon'_T)'$ with $\mathbf{y}_t = (y_{1t}, \ldots, y_{Nt})'$,$\mathbf{X}_t = (\mathbf{x}_{1t}, \ldots, \mathbf{x}_{Nt})'$, $\mathbf{u}_t = (u_{1t}, \ldots, u_{Nt})'$, and $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$. $\mathbf{1}_T$ is a T-dimensional vector of ones and $\otimes$ is the Kronecker product. The OLS estimator applied to Eq. (1) yields the fixed effects (FE) estimator (or within estimator) of β

$$\hat{\beta} = \left[\mathbf{X}'(\mathbf{M} \otimes \mathbf{I}_N)\mathbf{X}\right]^{-1} \mathbf{X}'(\mathbf{M} \otimes \mathbf{I}_N)\mathbf{y} \tag{5}$$

where $\mathbf{M} = \mathbf{I}_T - \mathbf{1}_T\left(\mathbf{1}'_T \mathbf{1}_T\right)^{-1} \mathbf{1}_T$ is the matrix that converts y_{it} and $\mathbf{x}_{it}$ in deviations from their individual-specific means.

We make use of the following assumptions:

Assumption 1. ε_{it} are independently distributed random variables with zero mean, variance $0 < E(\varepsilon_{it}^2) = \sigma_i^2 \le \sigma_{\max}^2 < \infty$, and such that $E|\varepsilon_{it}|^{4+\eta} \le K < \infty$ for some $\eta > 0$ and for $i = 1, 2, \ldots, N; t = 1, 2, \ldots, T$.

Assumption 2. $\mathbf{X}_t$ and $\varepsilon_{t'}$ are independently distributed for all t and t'. The elements of $\mathbf{X}$ are uniformly bounded constants; as N and/or T go to infinity the matrix $\frac{1}{NT}\mathbf{X}'(\mathbf{M} \otimes \mathbf{I}_N)\mathbf{X}$ exists and is nonsingular matrix.

Assumption 3. The main diagonal elements of $\mathbf{S}$ are zero. The row and column norms of the matrices $\mathbf{S}$ and $\mathbf{R} = (\mathbf{I}_N - \delta\mathbf{S})^{-1}$ are bounded.

Assumption 4. $\delta \in [c_l, c_u]$, with $-\infty < c_l, c_u < \infty$, and $(\mathbf{I}_N - \delta\mathbf{S})^{-1}$ is non-singular for all $\delta \in [c_l, c_u]$.

The existence of moments of order higher than four stated in Assumption 1 is needed for applicability of the central limit theorem by Kelejian and Prucha (2001). Assumption 2 implies strict exogeneity of regressors. This assumption rules out the presence of spatial or temporal lags of the dependent variable among the regressors. However, our approach can be extended to allow for these cases, by adopting an instrumental variable approach (Anderson and Hsiao, 1981; Mutl and Pfaffermayr, 2011). Assumption 4 allows rewriting Eq. (4) as:

$$\mathbf{u} = (\mathbf{I}_T \otimes \mathbf{R})\varepsilon, \tag{6}$$

where $\mathbf{R} = (\mathbf{I}_N - \delta\mathbf{S})^{-1}$. Under Assumptions 1–4 estimator (5) is unbiased since

$$E\left(\hat{\beta} - \beta\right) = \left[\frac{\mathbf{X}'(\mathbf{M} \otimes \mathbf{I}_N)\mathbf{X}}{NT}\right]^{-1} \frac{\mathbf{X}'(\mathbf{M} \otimes \mathbf{I}_N)E(\mathbf{u})}{NT} = 0,$$

and has variance

$$Var\left(\hat{\beta}\right) = \left[\mathbf{X}'(\mathbf{M} \otimes \mathbf{I}_N)\mathbf{X}\right]^{-1} \mathbf{X}'\left(\mathbf{M} \otimes \mathbf{R}\, diag\left\{\sigma_1^2, \ldots, \sigma_N^2\right\}\mathbf{R}'\right) \times \mathbf{X}\left[\mathbf{X}'(\mathbf{M} \otimes \mathbf{I}_N)\mathbf{X}\right]^{-1}.$$

Noting that, under Assumption 3, $\lambda_1(\mathbf{RR}')$ is bounded in N, we have

$$Var\left(\hat{\beta}\right) = \frac{1}{NT}\left[\frac{\mathbf{X}'(\mathbf{M} \otimes \mathbf{I}_N)\mathbf{X}}{NT}\right]^{-1} \frac{\mathbf{X}'\left(\mathbf{M} \otimes \mathbf{R}\, diag\left\{\sigma_1^2, \ldots, \sigma_N^2\right\}\mathbf{R}'\right)\mathbf{X}}{NT} \left[\frac{\mathbf{X}'(\mathbf{M} \otimes \mathbf{I}_N)\mathbf{X}}{NT}\right]^{-1} \le \frac{K}{NT}\left[\frac{\mathbf{X}'(\mathbf{M} \otimes \mathbf{I}_N)\mathbf{X}}{NT}\right]^{-1} = O\left(\frac{1}{NT}\right),$$

and the within estimator, $\hat{\beta}$, is $\sqrt{NT}$-consistent. However, when $\delta \neq 0$, $\hat{\beta}$ is in general not efficient since the covariance of errors (Eq. (4)) is non-diagonal and the elements along its main diagonal are not constant. Efficient estimation of the slope coefficients β can be achieved by estimating the parameters of Eq. (6) and then computing the feasible fixed-effects Generalized Least Squares (GLS) of the slope coefficients (see Qian and Schmidt, 2003 on this). In this paper we are concerned with consistent estimation of δ via GMM.

In the following, in order to distinguish the true parameters from other possible values in the parameter space, we denote by β_0, δ_0, and σ_{0i}^2 the true parameters, which generate an observed sample.

3. GMM estimation of SAR error models

3.1. Moment conditions

Following Kelejian and Prucha (1999), Lee and Liu (2006a), and others, for GMM estimation we consider a set of r linear quadratic conditions in the error term. In a panel data setting, in the presence of fixed effects, the ℓ th population moment condition is

$$\mathcal{M}_\ell(\delta) = \frac{1}{NT}E\left[\varepsilon^*(\delta)'\left(\mathbf{I}_T \otimes \mathbf{A}_\ell\right)\varepsilon^*(\delta)\right], \quad \ell = 1, 2, \ldots, r \tag{7}$$