Economics Letters 108 (2010) 65–68

Contents lists available at ScienceDirect

Economics Letters

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

Improved GMM estimation of the spatial autoregressive error model ☆

Matthias Arnold *, Dominik Wied

Institut für Wirtschafts- und Sozialstatistik, TU Dortmund, Dortmund, Germany

ARTICLE INFO

Article history:
Received 23 February 2009
Received in revised form 31 March 2010
Accepted 5 April 2010
Available online 13 April 2010

Keywords:
GMM estimation
Spatial autoregression
Regression residuals

JEL classification:
C13
C21

ABSTRACT

We suggest an improved GMM estimator for the autoregressive parameter of a spatial autoregressive error model by taking into account that unobservable regression disturbances are different from observable regression residuals.

1. Introduction and summary

Disturbances of regression models are typically not observable, so inference on the disturbances must rely on the regression residuals. It is well known that under general conditions, the residuals converge to the disturbances when the sample size increases, see e.g. Rao and Toutenburg (1995). However, the statistical properties of the disturbances and the residuals are different in finite samples.

This paper considers a linear regression model where the disturbances are generated by a spatial autoregressive model introduced by Cliff and Ord (1973) and where the parameter of main interest is the spatial autoregressive parameter.

Since the calculation of the maximum likelihood estimator can be computationally expensive, Kelejian and Prucha (1999) suggest a generalized method of moments (GMM) estimator, which uses three theoretical moments of the disturbances and equates them to the corresponding empirical moments of the residuals. This estimator has been applied to industrial specialization by Tingvall (2004), to microlevel data by Bell and Bockstael (2000) and to agricultural data by Schlenker et al. (2006) and Anselin et al. (2004). It has also been extended in several ways, for example to panel data by Druska and Horrace (2004) and to systems of simultaneous equations by Kelejian and Prucha (2004).

We suggest a variation of this estimator that is motivated by the following argument: If the empirical moments must rely on the residuals, the theoretical moments should be calculated in terms of the residuals, too. The computational costs are of the same order of magnitude for both estimators. Although both estimators coincide as sample size increases, our version is superior in finite samples, both in terms of bias and mean squared error. As a consequence, significance tests for the regression coefficients are less distorted because estimation of the corresponding covariance matrix is more accurate.

An empirical example illustrates our results. For a data set of Indonesian rice farms previously analyzed by Erwidodo (1990) and Druska and Horrace (2004), statistically significant effects of some of the covariates disappear if we implement the proposed modification.

In the following, we restrict ourselves to ordinary least squares regression in order to keep notation as simple as possible. The main idea however also applies to generalized least squares or nonlinear regression.

2. The model and the main result

We consider a linear regression model

$$y = X\beta + u, \tag{1}$$

where y is the $(n\times 1)$-vector of observations on the dependent variable, X is the nonstochastic $(n\times k)$-matrix on the explanatory variables and β

☆ We would like to thank Yongdai Kim and Walter Krämer for helpful comments and suggestions. Financial support from Deutsche Forschungsgemeinschaft (SFB 823, "Statistik nichtlinearer dynamischer Prozesse") is gratefully acknowledged.
* Corresponding author. Institut fur Wirtschafts- und Sozialstatistik, Fakultät Statistik, TU Dortmund, 44221 Dortmund, Germany. Tel.: +49 231 7553869; fax: +49 231 7555284.
E-mail address: arnold@statistik.tu-dortmund.de (M. Arnold).

0165-1765/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.econlet.2010.04.012

is the $(k\times 1)$-vector of unknown model parameters. We assume that u, the $(n\times 1)$-vector of disturbances, is generated by a spatial autoregressive model,

$$u = \rho W u + \varepsilon, \tag{2}$$

where W $(n\times n)$ is a weighting matrix of known constants, ρ is a scalar parameter and ε is an $(n\times 1)$-vector of innovations. We impose the following assumptions.

Assumption 1. (a) All diagonal elements of W are zero. (b) The row sums of W are equal to one, $\sum_{j=1}^{n} w_{ij}=1\,\forall i=1,\ldots,n$. (c) $|\rho|<1$.

Assumption 2. The innovations $\varepsilon_1,\ldots,\varepsilon_n$ are independently and identically distributed with zero mean and variance σ^2, where the variance is bounded by some positive constant b, $0<\sigma^2<b<\infty$. Additionally, $E(\varepsilon_i^4)<\infty\,\forall i\in 1,\ldots,n$.

Assumption 3. The elements of X are nonstochastic and bounded in absolute value by some $0<c_X<\infty$. Further, X has full column rank, and the matrix $Q_X = lim_{n\to\infty}\frac{1}{n}X'X$ is finite and nonsingular.

Assumption 1 ensures that the matrix $I-\rho W$ is nonsingular so that we have $u=(I-\rho W)^{-1}\varepsilon$ and

$$\mathrm{Cov}(u) = \sigma^2(I-\rho W)^{-1}\left(I-\rho W'\right)^{-1}. \tag{3}$$

Since u is not observable, estimation of ρ and σ^2 must rely on $\hat{u}$, the vector of regression residuals. For the case of OLS-regression, $\hat{u}$ is given by $\hat{u}=y-X\hat{\beta}=Mu$, where $M=I-X(X'X)^{-1}X'$, and $\hat{\beta}=(X'X)^{-1}X'y$ is the OLS-estimator of β.

In this situation, Kelejian and Prucha (1999) suggest a GMM estimator for ρ and σ^2 that uses three moments of ε, namely

$$E\left(\frac{1}{n}\varepsilon'\varepsilon\right) = \sigma^2, E\left(\frac{1}{n}\varepsilon'W'W\varepsilon\right) = \frac{\sigma^2}{n}tr\left(W'W\right), E\left(\frac{1}{n}\varepsilon'W'\varepsilon\right) = 0. \tag{4}$$

With the help of Eq. (2), the sample counterpart of Eq. (4) can be written as

$$G\left(\rho,\rho^2,\sigma^2\right)' - g = v\left(\rho,\sigma^2\right),$$

where

$$G = \begin{pmatrix} \frac{2}{n}\hat{u}'W\hat{u} & -\frac{1}{n}\hat{u}'W'W\hat{u} & 1 \\ \frac{2}{n}\hat{u}'W'WW\hat{u} & -\frac{1}{n}\hat{u}'W'W'WW\hat{u} & \frac{1}{n}tr\left(W'W\right) \\ \frac{1}{n}\hat{u}'\left[W+W'\right]W\hat{u} & -\frac{1}{n}\hat{u}'W'WW\hat{u} & 0 \end{pmatrix}$$

and

$$g = \left(\frac{1}{n}\hat{u}'\hat{u}, \frac{1}{n}\hat{u}'W'W\hat{u}, \frac{1}{n}\hat{u}'W\hat{u}\right)'.$$

The nonlinear least squares estimator of Kelejian and Prucha (1999) for ρ and σ^2 is defined as

$$\left(\hat{\rho}_{KP},\hat{\sigma}^2_{KP}\right) = \arg\min\left\{v\left(\rho,\sigma^2\right)'v\left(\rho,\sigma^2\right) : \rho\in[-a,a], \sigma^2\in[0,b]\right\}, \tag{5}$$

where $a\geq 1$ and $b<\infty$.

Our version proceeds as follows: If the unobservable disturbances u have to be replaced by the regression residuals $\hat{u}$, why not calculate the moment conditions (4) also in terms of $\hat{\varepsilon}=M\varepsilon=Mu-\rho MWu$ instead of ε? Therefore, we suggest an estimator that is based on the moments of $M\varepsilon$ corresponding to Eq. (4):

$$E\left(\frac{1}{n}(M\varepsilon)'M\varepsilon\right) = \frac{\sigma^2}{n}tr(M), \tag{6}$$

$$E\left(\frac{1}{n}(WM\varepsilon)'WM\varepsilon\right) = \frac{\sigma^2}{n}tr\left(MW'W\right), \tag{7}$$

$$E\left(\frac{1}{n}(WM\varepsilon)'M\varepsilon\right) = \frac{\sigma^2}{n}tr(WM), \tag{8}$$

where we use the fact that M is an orthogonal projector. If we multiply Eq. (2) by M and WM, respectively, we get

$$M\varepsilon = Mu - \rho MWu, \tag{9}$$

$$WM\varepsilon = WMu - \rho WMWu. \tag{10}$$

Plugging Eqs. (9) and (10) into the moment conditions (6)–(8) yields

$$\begin{aligned} &\frac{1}{n}E(u'Mu) - \frac{2\rho}{n}E(u'MWu) + \frac{\rho^2}{n}E(u'W'MWu) \\ &= \frac{\sigma^2}{n}tr(M), \frac{1}{n}E\left(u'MW'WMu\right) - \frac{2\rho}{n}E\left(u'W'WMWMu\right) \\ &+ \frac{\rho^2}{n}E\left(u'W'MW'WMWu\right) = \frac{\sigma^2}{n}tr\left(MW'W\right), \frac{1}{n}E\left(u'MW'Mu\right) \\ &- \frac{\rho}{n}E\left(u'M\left[W+W'\right]MWu\right) + \frac{\rho^2}{n}E\left(u'W'MWMWu\right) = \frac{\sigma^2}{n}tr(WM). \end{aligned}$$

Finally, for every $(n\times n)$-matrix A, the theoretical moments E $(u'Au)$ are replaced by their sample counterparts $\hat{u}'A\hat{u}$. Since $Mu=\hat{u}$ and $tr(M) = \frac{n-k}{n}$, the sample counterpart to Eqs. (6)–(8) can be written as

$$H\left(\rho,\rho^2,\sigma^2\right)' - h = w\left(\rho,\sigma^2\right),$$

where

$$H = \begin{pmatrix} \frac{2}{n}\hat{u}'W\hat{u} & -\frac{1}{n}\hat{u}'W'MW\hat{u} & \frac{n-k}{n} \\ \frac{2}{n}\hat{u}'W'WMW\hat{u} & -\frac{1}{n}\hat{u}'W'MW'WMW\hat{u} & \frac{1}{n}tr\left(MW'W\right) \\ \frac{1}{n}\hat{u}'\left[W+W'\right]MW\hat{u} & -\frac{1}{n}\hat{u}'W'MWMW\hat{u} & \frac{1}{n}tr(WM) \end{pmatrix}$$

and $h=g$. Our nonlinear least squares estimator for ρ and σ^2 is defined as

$$\left(\hat{\rho}_{RB},\hat{\sigma}^2_{RB}\right) = \arg\min\left\{w\left(\rho,\sigma^2\right)'w\left(\rho,\sigma^2\right) : \rho\in[-a,a], \sigma^2\in[0,b]\right\}, \tag{11}$$

where $a\geq 1$ and $b<\infty$.

The following theorem states the asymptotic equivalence of $\left(\hat{\rho}_{KP},\hat{\sigma}^2_{KP}\right)$ and $\left(\hat{\rho}_{RB},\hat{\sigma}^2_{RB}\right)$.

Theorem 1. *Under Assumptions 1–3, for $n\to\infty$*

$$\left(\hat{\rho}_{RB},\hat{\sigma}^2_{RB}\right) - \left(\hat{\rho}_{KP},\hat{\sigma}^2_{KP}\right) \xrightarrow{P} 0.$$

Proof. Because of Assumption 3, for large n the elements of $X(X'X)^{-1}X'$ are bounded by the corresponding elements of $\frac{kc_X^2}{n}Q_X^{-1} \xrightarrow{n\to\infty} 0$ so that $M=I-X(X'X)^{-1}X' \xrightarrow{n\to\infty} I$ and thus $H\xrightarrow{P}G$ as $n\to\infty$. Since