1 Overview

The programs made available in this web page are sample programs for the computation of estimators introduced in Kelejian and Prucha (1998). In particular we provide sample programs for the estimation of the following cross-sectional (first-order) autoregressive spatial model with (first-order) autoregressive disturbances:

$$\begin{aligned} y_n &= X_n\beta + \lambda W_n y_n + u_n, \qquad |\lambda| < 1, \\ u_n &= \rho M_n u_n + \varepsilon_n, \qquad |\rho| < 1, \end{aligned} \tag{1}$$

where y_n is the $n \times 1$ vector of observations on the dependent variable, X_n is the $n \times k$ matrix of observation on k exogenous variables, M_n and W_n are $n \times n$ spatial weighting matrices of known constants, β is the $k \times 1$ vector of regression parameters, λ and ρ are scalar autoregressive parameters, u_n is the $n \times 1$ vector of regression disturbances, ε_n is an $n \times 1$ vector of innovations. The variables $W_n y_n$ and $M_n u_n$ are typically referred to as spatial lag of y_n and u_n, respectively.

For the following discussion it proves helpful to rewrite (1) more compactly as

$$\begin{aligned} y_n &= Z_n\delta + u_n, \\ u_n &= \rho M_n u_n + \varepsilon_n, \end{aligned} \tag{2}$$

where $Z_n = (X_n, W_n y_n)$ and $\delta = (\beta', \lambda)'$. Applying a Cochrane-Orcutt transformation to this model yields furthermore

$$y_{n*}(\rho) = Z_{n*}(\rho)\delta + \varepsilon_n, \tag{3}$$

where $y_{n*}(\rho) = y_n - \rho M_n y_n$ and $Z_{n*}(\rho) = Z_n - \rho M_n Z_n$.

Note: We provide two sets of sample programs, one for TSP and one for Stata. In the following we only describe the use of the TSP programs. The use of the Stata files is analogous.

2 Data Files

The sample program involves two exogenous variables and idealized spatial weighting matrices M_n and W_n. For simplicity M_n and W_n are taken to be equal. This matrix corresponds to the case where each unit has "one neighbor ahead and one neighbor behind" in a wrap around world, and the row sums of the weighting matrix are normalized to one; for a more detailed description of this idealized matrix see the Monte Carlo section of Kelejian and Prucha (1999). The sample size is taken to be 100. The actual estimation programs assume that the data for the exogenous variables and spatial weighting matrix are stored in files named VAR2.DAT and MMAT.DAT, respectively.

3 Estimation Programs

The main estimation program is contained in the file PROGRAM2.TSP. This program calls two "subroutines" contained in the files GMPROC1.TSP and TSLSPROC2.TSP. Those subroutines compute the GM estimator for ρ and the TSLS estimator for β and λ, respectively.

The program PROGRAM2.TSP first reads in the data for the exogenous variables and spatial weighting matrix from the files VAR2.DAT and MMAT.DAT. The actual estimation of the parameters of the model (1)-(2) is performed in three steps.

Step 1: In the first step we estimate the model in (2) by two-stage least squares estimator (TSLS):

$$\begin{aligned} \widetilde{\delta}_n &= (\widehat{Z}_n'\widehat{Z}_n)^{-1}\widehat{Z}_n' y_n, \\ \widehat{u}_n &= y_n - Z_n\widetilde{\delta}_n, \end{aligned} \tag{4}$$

where $\widehat{Z}_n = P_{H_n} Z_n = (X_n, \widehat{W_n y_n})$, where $\widehat{W_n y_n} = P_{H_n} W_n y_n$ and $P_{H_n} = H_n(H_n'H_n)^{-1}H_n'$, and where H_n is the matrix of instruments which is formed as a subset of linearly independent columns of $(X_n, W_nX_n, W_n^2X_n)$.

Step 2: In the second step the spatial autoregressive parameter ρ is estimated in terms of the residuals obtained via the first step and the generalized moments procedure suggested in Kelejian and Prucha (1998). More specifically the estimators of ρ and σ_ε^2, $\widetilde{\rho}_n$ and $\widetilde{\sigma}_{\varepsilon,n}^2$ respectively are defined as the nonlinear least squares estimators that minimizes

$$\left[g_n - G_n \begin{bmatrix} \rho \\ \rho^2 \\ \sigma_\varepsilon^2 \end{bmatrix} \right]' \left[g_n - G_n \begin{bmatrix} \rho \\ \rho^2 \\ \sigma_\varepsilon^2 \end{bmatrix} \right], \tag{5}$$

where

$$G_n = \frac{1}{n} \begin{bmatrix} 2\widehat{u}_n'\widehat{\overline{u}}_n & -\widehat{\overline{u}}_n'\widehat{\overline{u}}_n & n \\ 2\widehat{\overline{u}}_n'\widehat{\overline{\overline{u}}}_n & -\widehat{\overline{\overline{u}}}_n'\widehat{\overline{\overline{u}}}_n & Tr(M_n'M_n) \\ \widehat{u}_n'\widehat{\overline{\overline{u}}}_n + \widehat{\overline{u}}_n'\widehat{\overline{u}}_n & -\widehat{\overline{u}}_n'\widehat{\overline{\overline{u}}}_n & 0 \end{bmatrix}, \qquad g_n = \frac{1}{n} \begin{bmatrix} \widehat{u}_n'\widehat{u}_n \\ \widehat{\overline{u}}_n'\widehat{\overline{u}}_n \\ \widehat{u}_n'\widehat{\overline{u}}_n \end{bmatrix},$$

and $\widehat{\overline{u}}_n = M_n\widehat{u}_n$, and $\widehat{\overline{\overline{u}}}_n = M_n^2\widehat{u}_n$. The code for computing the GM estimators is contained in the file GMPROC1.TSP

Step 2: In the third step the transformed regression model in (3) is estimated by TSLS:

$$\widetilde{\delta}_n = \left[\widehat{Z}_{n*}'(\widetilde{\rho}_n)\widehat{Z}_{n*}(\widetilde{\rho}_n)\right]^{-1} \widehat{Z}_{n*}'(\widetilde{\rho}_n) y_{n*}(\widetilde{\rho}_n), \tag{6}$$

where $\widehat{Z}_{n*}(\widetilde{\rho}_n) = P_{H_n}Z_{n*}(\widetilde{\rho}_n)$, $Z_{n*}(\widetilde{\rho}_n) = Z_n - \widetilde{\rho}_n M_n Z_n$, $y_{n*}(\widetilde{\rho}_n) = y_n - \widetilde{\rho}_n M_n y_n$. This procedure is executed in TSLSPROC2.TSP

References

[1] 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,* Vol. 17:1, 99-121.

[2] Kelejian, H. H. and I. R. Prucha, (1999): "A Generalized Moments of Estimator for the Autoregressive Parameter in a Spatial Model," *International Economic Review,* 40, 509-533.