1 Overview

The programs made available on this web page are sample programs for the computation of estimators introduced in Kapoor, Kelejian, and Prucha, (2003). In particular, we provide sample programs for the estimation of a linear panel data model with spatially correlated random effects. More specifically, we assume that in each time period $t = 1, \ldots, T$ the data are generated according to the following model:

$$\begin{aligned} y_N(t) &= X_N(t)\beta + u_N(t), & (1)\\ u_N(t) &= \rho W_N u_N(t) + \varepsilon_N(t), \qquad |\rho| < 1, & (2) \end{aligned}$$

where $y_N(t)$ is the $N \times 1$ vector of observations on the dependent variable in time period t, $X_N(t)$ is the $N \times k$ matrix of observation on k exogenous variables in period t, W_N is an $N \times N$ spatial weighting matrix of known constants which does not involve t, β is the $k \times 1$ vector of regression parameters, ρ is the spatial autoregressive parameter, $u_N(t)$ is the $N \times 1$ vector of regression disturbances, and $\varepsilon_N(t)$ is an $N \times 1$ vector of innovations in period t.

Stacking the observations in (1) and (2) we have

$$\begin{aligned} y_N &= X_N\beta + u_N, & (3)\\ u_N &= \rho(I_T \otimes W_N)u_N + \varepsilon_N, & (4) \end{aligned}$$

where

$$\begin{aligned} y_N &= [y_N'(1), \ldots, y_N'(T)]',\\ X_N &= [X_N'(1), \ldots, X_N'(T)]',\\ u_N &= [u_N'(1), \ldots, u_N'(T)]',\\ \varepsilon_N &= [\varepsilon_N'(1), \ldots, \varepsilon_N'(T)]'. \end{aligned}$$

We assume furthermore the following error component structure for the innovation vector ε_N :

$$\varepsilon_N = (e_T \otimes I_N)\mu_N + v_N, \tag{5}$$

where μ_N represents the $N \times 1$ vector of unit specific error components, and

$$v_N = [v_N'(1), \ldots, v_N'(T)]'$$

contains the error components that vary over both the cross-sectional units and time periods. The error components are assumed to satisfy:

$$\begin{aligned} E\mu_N = 0, \quad E\mu_N\mu_N' = \sigma_\mu^2 I_N,\\ Ev_N = 0, \quad Ev_N v_N' = \sigma_v^2 I_{NT.} \end{aligned}$$

A complete set of assumptions is given in the paper.

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 programs is analogous.

2 Data Files

The sample program involves two exogenous variables and an idealized spatial weighting matrix. 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 number of time periods is equal to 4. The actual estimation programs assume that the data for the dependent and exogenous variables and for the spatial weighting matrix are stored in files named VAR3.DAT and MMAT.DAT, respectively.

3 Estimation Programs

The main estimation program is contained in the file PROGRAM3.TSP. This program calls two "subroutines" contained in the files GMPROC3.TSP and GLSPROC3.TSP. Those subroutines compute the GM estimator for ρ and the feasible GLS estimator for β.

The program PROGRAM3.TSP first reads in the data for the dependent and exogenous variables and for the spatial weighting matrix from the files VAR3.DAT and MMAT.DAT. The actual estimation of the parameters of the model (3)-(5) is performed in three steps.

Step 1: In the first step we estimate the regression model in (3) using ordinary least squares (OLS) to obtain

$$\begin{aligned} \widehat{\beta}_{OLS} &= (X_N' X_N)^{-1} X_N' y_N, \\ \widehat{u}_N &= y_N - X_N \widehat{\beta}_{OLS}. \end{aligned}$$

Step 2: This second step is performed by GMPROC3.TSP. In this step we estimate the spatial autoregressive parameter ρ and the variance components σ_v^2 and σ_μ^2 (or equivalently σ_v^2 and $\sigma_1^2 = \sigma_v^2 + T\sigma_\mu^2$) in terms of the residuals obtained via the first step and the generalized moments procedure suggested in the paper. This estimation procedure is executed in two parts.

In the first part of Step 2 we compute initial GM estimators for ρ, σ_v^2 and σ_1^2. Let $\widetilde{\rho}_N$, $\widetilde{\sigma}_v^2$ and $\widetilde{\sigma}_1^2$ denote those estimators, respectively. Then $\widetilde{\rho}_N$ and $\widetilde{\sigma}_{v,N}^2$ are defined as the unweighted nonlinear least squares estimators, which minimize

$$\left[g_N^0 - G_N^0 \begin{bmatrix} \rho \\ \rho^2 \\ \sigma_v^2 \end{bmatrix} \right]' \left[g_N^0 - G_N^0 \begin{bmatrix} \rho \\ \rho^2 \\ \sigma_v^2 \end{bmatrix} \right], \tag{6}$$

where

$$\begin{aligned} G_N^0 &= \begin{bmatrix} \frac{2}{N(T-1)} \widehat{u}_N' Q_{0,N} \widehat{\overline{u}}_N & -\frac{1}{N(T-1)} \widehat{\overline{u}}_N' Q_{0,N} \widehat{\overline{u}}_N & 1 \\ \frac{2}{N(T-1)} \widehat{\overline{u}}_N' Q_{0,N} \widehat{\overline{\overline{u}}}_N & -\frac{1}{N(T-1)} \widehat{\overline{\overline{u}}}_N' Q_{0,N} \widehat{\overline{\overline{u}}}_N & \frac{1}{N} Tr(W_N' W_N) \\ \frac{1}{N(T-1)} (\widehat{u}_N' Q_{0,N} \widehat{\overline{\overline{u}}}_N + \widehat{\overline{u}}_N' Q_{0,N} \widehat{\overline{u}}_N) & -\frac{1}{N(T-1)} \widehat{\overline{u}}_N' Q_{0,N} \widehat{\overline{\overline{u}}}_N & 0 \end{bmatrix}, \\ g_N^0 &= \begin{bmatrix} \frac{1}{N(T-1)} \widehat{u}_N' Q_{0,N} \widehat{u}_N \\ \frac{1}{N(T-1)} \widehat{\overline{u}}_N' Q_{0,N} \widehat{\overline{u}}_N \\ \frac{1}{N(T-1)} \widehat{u}_N' Q_{0,N} \widehat{\overline{u}}_N \end{bmatrix}, \end{aligned}$$

with $\widehat{\overline{u}}_N = W_N \widehat{u}_N$ and $\widehat{\overline{\overline{u}}}_N = W_N^2 \widehat{u}_N$. Furthermore the estimator $\widetilde{\sigma}_1^2$ is defined as

$$\widetilde{\sigma}_1^2 = \frac{1}{N} (\widehat{u}_N - \widetilde{\rho}_N \widehat{\overline{u}}_N)' Q_{1,N} (\widehat{u}_N - \widetilde{\rho}_N \widehat{\overline{u}}_N). \tag{7}$$

The transformation matrices $Q_{0,N}$ and $Q_{1,N}$ are defined in the paper.

In the second part of Step 2 the program does one of three things, depending on which value the user has selected at the beginning of the program for the variable L.

If $L=1$ the program skips this part and passes on the initial GM estimates for ρ, σ_v^2 and σ_1^2 obtained in the first part to Step 2.

If $L=2$ the program computes the partially weighted GM estimators for ρ, σ_v^2 and σ_1^2, say $\check{\rho}_N$, $\check{\sigma}_{v,N}^2$ and $\check{\sigma}_{1,N}^2$ and passes those estimates on to Step 3. The estimators $\check{\rho}_N$, $\check{\sigma}_{v,N}^2$ and $\check{\sigma}_{1,N}^2$ are defined as the minimizing values of:

$$\left[G_N \begin{bmatrix} \rho \\ \rho^2 \\ \sigma_v^2 \\ \sigma_1^2 \end{bmatrix} - g_N\right]' \widetilde{\Upsilon}_N^{-1} \left[G_N \begin{bmatrix} \rho \\ \rho^2 \\ \sigma_v^2 \\ \sigma_1^2 \end{bmatrix} - g_N\right], \tag{8}$$

where

$$G_N = \begin{bmatrix} \frac{2}{N(T-1)}\widehat{u}_N' Q_{0,N}\widehat{\overline{u}}_N & -\frac{1}{N(T-1)}\widehat{\overline{u}}_N' Q_{0,N}\widehat{\overline{u}}_N & 1 & 0 \\ \frac{2}{N(T-1)}\widehat{\overline{u}}_N' Q_{0,N}\widehat{\overline{\overline{u}}}_N & -\frac{1}{N(T-1)}\widehat{\overline{\overline{u}}}_N' Q_{0,N}\widehat{\overline{\overline{u}}}_N & \frac{1}{N}Tr(W_N'W_N) & 0 \\ \frac{1}{N(T-1)}(\widehat{u}_N' Q_{0,N}\widehat{\overline{u}}_N + \widehat{\overline{\overline{u}}}_N' Q_{0,N}\widehat{\overline{\overline{u}}}_N) & -\frac{1}{N(T-1)}\widehat{\overline{u}}_N' Q_{0,N}\widehat{\overline{\overline{u}}}_N & 0 & 0 \\ \frac{2}{N}\widehat{u}_N' Q_{1,N}\widehat{\overline{u}}_N & -\frac{1}{N}\widehat{\overline{u}}_N' Q_{1,N}\widehat{\overline{u}}_N & 0 & 1 \\ \frac{2}{N}\widehat{\overline{u}}_N' Q_{1,N}\widehat{\overline{\overline{u}}}_N & -\frac{1}{N}\widehat{\overline{\overline{u}}}_N' Q_{1,N}\widehat{\overline{\overline{u}}}_N & 0 & \frac{1}{N}Tr(W_N'W_N) \\ \frac{1}{N}(\widehat{u}_N' Q_{1,N}\widehat{\overline{u}}_N + \widehat{\overline{\overline{u}}}_N' Q_{1,N}\widehat{\overline{\overline{u}}}_N) & -\frac{1}{N}\widehat{\overline{u}}_N' Q_{1,N}\widehat{\overline{\overline{u}}}_N & 0 & 0 \end{bmatrix},$$

$$g_N = \begin{bmatrix} \frac{1}{N(T-1)}\widehat{u}_N' Q_{0,N}\widehat{u}_N \\ \frac{1}{N(T-1)}\widehat{\overline{u}}_N' Q_{0,N}\widehat{\overline{u}}_N \\ \frac{1}{N(T-1)}\widehat{u}_N' Q_{0,N}\widehat{\overline{u}}_N \\ \frac{1}{N}\widehat{u}_N' Q_{1,N}\widehat{u}_N \\ \frac{1}{N}\widehat{\overline{u}}_N' Q_{1,N}\widehat{\overline{u}}_N \\ \frac{1}{N}\widehat{u}_N' Q_{1,N}\widehat{\overline{u}}_N \end{bmatrix},$$

and

$$\widetilde{\Upsilon}_N = \begin{bmatrix} \frac{1}{T-1}\widetilde{\sigma}_{v,N}^4 & 0 \\ 0 & \widetilde{\sigma}_{1,N}^4 \end{bmatrix} \otimes I_3.$$

If $L=3$ the program computes the (fully) weighted GM estimators for ρ, σ_v^2 and σ_1^2, say $\hat{\rho}_N$, $\hat{\sigma}_{v,N}^2$ and $\hat{\sigma}_{1,N}^2$, and passes those estimates on to Step 3. The difference between the partially weighted and the (fully) weighted GM estimators is that for the latter estimators in (8) the identity matrix I_3 in $\widetilde{\Upsilon}_N$ is replaced by the matrix T_W defined as:

$$T_W = \begin{bmatrix} 2 & 2Tr\left(\frac{W_N'W_N}{N}\right) & 0 \\ 2Tr\left(\frac{W_N'W_N}{N}\right) & 2Tr\left(\frac{W_N'W_NW_N'W_N}{N}\right) & Tr\left(\frac{W_N'W_N(W_N'+W_N)}{N}\right) \\ 0 & Tr\left(\frac{W_N'W_N(W_N'+W_N)}{N}\right) & Tr\left(\frac{W_NW_N+W_N'W_N}{N}\right) \end{bmatrix}.$$

Step 3: This third step is performed by GLSPROC3.TSP. In this step the regression model in (3)-(4) is reestimated in terms of a feasible GLS estimator. In abuse of notation, let $\check{\rho}_N$, $\check{\sigma}_{v,N}^2$ and $\check{\sigma}_{1,N}^2$ be the estimates computed and passed on by Step 2. Then the GLS estimator is defined as

$$\widehat{\beta}_{FGLS} = \left\{X_N^{*\prime}(\check{\rho}_N)[\Omega_{\varepsilon,N}^{-1}(\check{\sigma}_{v,N}^2,\check{\sigma}_{1,N}^2)]X_N^*(\check{\rho}_N)\right\}^{-1} X_N^{*\prime}(\check{\rho}_N)[\Omega_{\varepsilon,N}^{-1}(\check{\sigma}_{v,N}^2,\check{\sigma}_{1,N}^2)]y_N^*(\check{\rho}_N), \tag{9}$$

where $X_N^*(\check{\rho}_N) = [I_T \otimes (I_N - \check{\rho}_N W_N)]X_N$, $y_N^*(\check{\rho}_N) = [I_T \otimes (I_N - \check{\rho}_N W_N)]y_N$, and $\Omega_{\varepsilon,N}(\check{\sigma}_{v,N}^2,\check{\sigma}_{1,N}^2) = \check{\sigma}_{v,N}^2 Q_{0,N} + \check{\sigma}_{1,N}^2 Q_{1,N}$. This estimator can be computed as the OLS estimator applied to model (3) after premultiplication with $[I_T \otimes (I_N - \check{\rho}_N W_N)]$ and

$$\check{\sigma}_{v,N}\Omega_{\varepsilon,N}^{-1/2}(\check{\sigma}_{v,N}^2,\check{\sigma}_{1,N}^2) = Q_{0,N} + \check{\theta}Q_{1,N} = I_{NT} - (1-\check{\theta})Q_{1,N},$$

where $\check{\theta} = \check{\sigma}_{v,N}/\check{\sigma}_{1,N}$. Since $Q_{1,N}[I_T \otimes (I_N - \check{\rho}_N W_N)] = [I_T \otimes (I_N - \check{\rho}_N W_N)]Q_{1,N}$ it does not matter in which order the data are transformed.

References

[1] Kapoor, M. , H. H. Kelejian, and I. R. Prucha, (2003): "Panel Data Models with Spatially Correlated Error Components," *Working Paper,* University of Maryland, College Park.

[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.