Spatial panel multinomial logit models for large samples

Abdourahmane Diallo, Ghislain Geniaux

Spatial Panel Multinomial Logit Models for large samples : An application to Land Use Change Models

[1] A. DIALLO

[1]INRA SAD Ecodeveloppement Avignon (UR 767)

SEW 2012

Plan

Panel Data in Spatial Econometrics Models

Benefits

- Modeling heterogeneity
- More specifications.

Costs

- More complex estimation procedures(Incidental parameters problem).
- Large data sets.

Introduction

- The objectives of our article are:
  1. to propose a full GMM for spatial panel multinomial choice model,
  2. to derive a linerized version of our GMM estimator.

GMM

The strutural model :

$$y^*_{jt} = \alpha_{jt} I_n + \lambda_t W y^*_{jt} + X_t \beta_{jt} + \epsilon_{jt} \qquad (1)$$

$$y_{ijt} = \begin{cases} 1 & \text{if} \quad y^*_{ijt} \geq y^*_{ikt} \quad \mathrm{k} \neq j \in A \\ 0 & \text{otherwise} \end{cases}$$

Under the assumptions in Kelejian and Prucha (1999), we can rewrite the spatial latent model :

$y^*_{jt} = (I - \lambda_t W)^{-1} \alpha l_n + (I - \lambda_t W)^{-1} X_t \beta_j + (I - \lambda_t W)^{-1} \epsilon_{jt}$

We obtain that the variance-covariance matrix is proportional to: $V = I_t \otimes [(I - \lambda_t W)^{-1}((I - \lambda_t W)^{-1})']$

Let us denote

$\sigma^2_V(\lambda_t) = diag(V_t) = (\sigma^2_1(\lambda_t), \sigma^2_2(\lambda_t), ..., \sigma^2_n(\lambda_t))'$,

$V_t = (I - \lambda_t W)^{-1}((I - \lambda_t W)^{-1})$

$X^*_t = \frac{(I-\lambda_t W)^{-1} X}{\sigma_V(\lambda_t)}$ and $W^*_t = \frac{(I-\lambda_t W)^{-1} l_n}{\sigma_V(\lambda_t)}$, A set of alternatives , J the cardinal of A and A^- is a subset of A with cardinal $J - 1$.

Suppose that unobserved components are independently identically distributed across individuals and time-periods and drawn from the type *I* extreme value distribution, then the choice probability for the individual i to select an alternative j at period t is:

$$P_{ijt} = \frac{\exp\left(X_{it}^{*}\beta_j + W_{it}^{*}\alpha_{jt}\right)}{1 + \sum_{k \in A^{-}} \exp\left(X_{it}^{*}\beta_k + W_{it}^{*}\alpha_{jt}\right)}$$

The previous model implies both heteroskedasticity and autocorrelation $\Longrightarrow$ multinomial panel logit estimation method is inconsistent.

We propose, like in Pinkse and Slade (1998), Klier and McMillen (2008) and Diallo and Geniaux (2012), a GMM estimator based on the moment conditions implied by the likelihood function for a spatial panel multinomial logit model.

$$\tilde{u}_{ijt} = y_{ijt} - p_{ijt}$$

is a generalized residual.

Let us denote $\theta = [\beta = (\beta_1, ..., \beta_J), \lambda = (\lambda_1, ..., \lambda_T), \alpha = (\alpha_{11}, ..., \alpha_{JT})] \in \Theta$, where Θ is the parameter open space.

$Q(\theta) = {S_{nT}}'(\theta)\Sigma_{nT}S_{nT}(\theta)$

where $S_{nT}(\theta) = \frac{1}{nT}Z'\tilde{u}$, Σ_{nT} some positive definite matrix

and Z is a matrix of instruments $Z = \begin{pmatrix} Z_1 & 0 & . & . & 0 \\ 0 & Z_2 & . & . & . \\ . & . & . & . & 0 \\ 0 & . & . & 0 & Z_T \end{pmatrix}$.

Where $Z_t = \Big[X_t, WX_t, W^2X_t, W^3X_t],\ t = 1, 2, ..., T.$

$$
\begin{aligned}
(G_{\alpha_j})_{it} &= p_{ijt}(1-p_{ijt})(((I-\lambda_t W)^{-1} l_n)_i \\
(G_{\beta_j})_{it} &= p_{ijt}(1-p_{ijt})X^*_{it} \\
(G_{\lambda t})_{it} &= p_{ijt}\Big[\sum_{k\in A} p_{ikt}H_{it}(\beta_j-\beta_k) - \frac{X^*_{it}(\beta_j-\beta_k)\Lambda_{ii}}{\sigma_i^2} \\
&+ \sum_{k\in A} p_{ikt}H^*_{it}(\alpha_{jt}-\alpha_{kt}) - \frac{(W^* l_n)_{it}(\alpha_{jt}-\alpha_{kt})\Lambda_{ii}}{\sigma_i^2}\Big]
\end{aligned}
$$

Where $H_t = (I-\lambda_t W)^{-1}WX^*_t$, $H^*_t = (I-\lambda_t W)^{-1}WW^*_t$ and Λ is a the digonal matrix of the following matrix $\Big((I-\lambda_t W)^{-1}W(I-\lambda_t W)^{-1}(I-\lambda_t W)^{-1'}\Big)$.

It is important to note that, when $\lambda_t = 0$ and $\alpha_{jt} = 0$,

$$\begin{aligned} G_{\alpha_j} &= p_j(1-p_j)[I_t \otimes I_n] \\ G_{\beta_j} &= p_j(1-p_j)X \\ G_{\lambda_t} &= p_j\left[\sum_{k\in A} p_k WX_t(\beta_j-\beta_k)\right] \end{aligned}$$

$$X = \begin{pmatrix} X_1 \\ . \\ X_T \end{pmatrix}$$

As Klier and Klier and McMillen (2008) and Diallo and Geniaux (2012), we develop a linearized version of our GMM

LGMM

$$\tilde{u}_{ijt} = y_{ijt} - p_{ijt}$$

$$\tilde{u}(\beta, \lambda, \alpha) - \tilde{u}(\beta^{SMNlogit}, \lambda = 0, \alpha = 0) = G(\theta - \theta_0)$$

where $G(.)$ is a gradient term

$$\tilde{u}(\beta, \lambda, \alpha) = G(\theta - \theta_0) + \tilde{u}(\beta^{SMNlogit}, \lambda = 0, \alpha = 0)$$

where $\theta_0 = (\beta^{SMNlogit}, \lambda = 0, \alpha = 0)$.

linearized procedure

We estimate $\beta^{SMNlogit}$ by standard multinomial logit estimation method with $\lambda = 0$ and $\alpha = 0$ [1] and we calculate $u(\beta^{SMNplogit}, 0, 0)$, we obtain the following gradient terms:

$$
\begin{aligned}
G_{\alpha_j}(\beta^{SMNPlogit}, \lambda = 0, \alpha = 0) &= p_j(1 - p_j)[I_t \otimes I_n] \\
G_{\beta_j}(\beta^{SMNPlogit}, \lambda = 0, \alpha = 0) &= p_j(1 - p_j)X \\
G_{\lambda_t}(\beta^{SMNPlogit}, \lambda = 0, \alpha = 0) &= p_j\left[\sum_{k \in A} p_k WX_t(\beta_j - \beta_k)\right]
\end{aligned}
$$

[1] spatial autocorellation, heteroscedastic and heterogeneity are ignored

linearized procedure

Regress G_α ,G_β and G_λ on Z. The predicted values are $\hat{G}_\alpha, \hat{G}_\beta$ and $\hat{G}_\lambda$. Finally regress $u(\beta^{SMNPlogit}, 0, 0) + G_\beta' \beta^{SMNlogit}$ on $\hat{G}_\alpha, \hat{G}_\beta$ and $\hat{G}_\lambda$. The coefficients of this last regression provides the estimated values of α, β and λ..

This algorithm provides accurate estimators of λ and α, and an biased estimator of β.

Adjusting coeffient

To correct the downward bias in β, we will need to calculate the marginal effects [2]. The general coefficient to adjust the downward bias in β is given by

$$\frac{Tr((\sigma_V(\lambda))}{Tr((A))}$$

where Tr is the trace operator of matrix

[2] see (Diallo and Geniaux, 2012)

- We present the results of Monte Carlo analysis to evaluate the performance of full GMM and linearized multinomial logit GMM estimators with a spatial panel data framework.
- we generate the data according to:

$$P_{ijt} = \frac{\exp\left(X_{it}^{*}\beta_j + W_{it}^{*}\alpha_{jt}\right)}{1 + \sum_{k\in A^{-}} \exp\left(X_{it}^{*}\beta_k + W_{it}^{*}\alpha_{jt}\right)}$$

MC settings

- n=1000, $i = 1, .., n$ individuals, $j = 1, ..., 4$ alternatives and $t = 1, ..., 3$ time-period.
- X_t is drawn from bivariate $U(-1, 1)$
- $\lambda = (\lambda_1, \lambda_2, \lambda_3) \in [0, 0.9] \times [0, 0.9] \times [0, 0.9]$
- For alternative 0 $\beta_{10} = \beta_{20} = 0$,for alternative 1 $\beta_{11} = 0.5; \beta_{21} = 1$,for alternative 2 $\beta_{12} = 0; \beta_{22} = -1$,for alternative 3 $\beta_{13} = 0; \beta_{23} = 0.7$.

MC settings

- The fixed time-varying preferential effects are genarated as $\alpha_{jt} = (1/n) \sum_{i=1}^{n} X_{it} \beta_j$
- $\alpha_{0t} = 0$
- The exogenous spatial weight matrix for the experiments is created as follows: firstly we draw two random numbers from the uniform distribution $U(0, 1)$ for each observation. These numbers are used to specify the coordinates of each observation in the $[0, 1] \times [0, 1]$ plane.

MC settings

- To work in large samples, we use the following approximation $(I-\lambda_t W)^{-1} = I + \lambda_t W + \lambda_t^2 W^2 + \lambda_t^3 W^3$ and we suppose that each location have two neighbors.
- $Z = \begin{pmatrix} Z_1 & 0 & . & . & 0 \\ 0 & Z_2 & . & . & . \\ . & . & . & . & 0 \\ 0 & . & . & 0 & Z_T \end{pmatrix}$.

  Where $Z_t = \left[X_t, WX_t, W^2X_t, W^3X_t\right],\ t = 1, 2, ..., T$.

Results of Monte Carlo experiment with 1000 repetitions, n=1000

			Standard Multinomial Logit						LGMM								
			b_{11}	b_{21}	b_{12}	b_{22}	b_{13}	b_{23}	b_{11}	b_{21}	b_{12}	b_{22}	b_{13}	b_{23}			
λ_{01}	λ_{02}	λ_{03}	0.5	1	0	-1	0	0.7	0.5	1	0	-1	0	0.7	λ_1	λ_2	λ_3
0	0	0	0.505	1.005	-0.007	-1.008	-0.007	0.701	0.498	1.001	0.001	-0.998	-0.003	0.707	0.001	0.002	0.001
	RMSE		0.082	0.081	0.082	0.083	0.080	0.081	0.076	0.081	0.075	0.087	0.077	0.080	0.119	0.124	0.120
0.1	0.2	0.3	0.496	0.993	-0.002	-1.001	-0.002	0.694	0.505	1.001	0.005	-1.019	-0.005	0.697	0.101	0.200	0.302
	RMSE		0.089	0.088	0.087	0.090	0.090	0.091	0.076	0.081	0.075	0.083	0.075	0.078	0.122	0.126	0.129
0.7	0.8	0.9	0.405	0.800	-0.004	-0.818	-0.002	0.581	0.501	0.997	0.004	-1.049	-0.010	0.697	0.775	0.885	1.025
	RMSE		0.097	0.095	0.098	0.096	0.099	0.101	0.085	0.095	0.085	0.096	0.080	0.083	0.142	0.156	0.161

References I

Diallo, A., Geniaux, G., 2012. Spatial multinomial logit model for large samples: An application for land use model. In: SEA2012 meeting on Spatial Econometrics.

Kelejian, H., Prucha, I., 1999. A generalized moments estimator for the autoregressive parameter in a spatial model. International economic review 40 (2), 509–533.

Klier, T., McMillen, D., 2008. Clustering of auto supplier plants in the united states. Journal of Business and Economic Statistics 26 (4), 460–471.

Pinkse, J., Slade, M., 1998. Contracting in space: An application of spatial statistics to discrete-choice models. Journal of Econometrics 85 (1), 125–154.