Spatial versus non-spatial regression models – comparison on examples

Jan Hauke and Tomasz Kossowski

Adam Mickiewicz University, Poznań, Poland

Abstract

It is known that models connected with geographically distributed variables relies on sample data that is collected with reference to location measured as points in space. For data having location component we assume naturally that there exists so-called spatial effects identifying both spatial dependence between the observations and spatial heterogeneity occurring in the analysed relationships.

The standard linear regression models ignore these two effects. In fact these effects infect the traditional Gauss-Markov assumptions (like spherical disturbance of errors and terms). As noted in Anselin (1988), Gauss-Markov assumes that explanatory variables are fixed in repeated sampling, while spatial dependence violates this assumption. This gives rise to the need for alternative estimation approaches. Similarly, spatial heterogeneity violates the Gauss- Markov assumption that a single linear relationship exists across the sample data observations. If the relationship varies as we move across the spatial data sample, alternative estimation procedures are needed to successfully model this type of variation and draw appropriate inferences.

In the paper we compare the following models:
a) classical linear regression model with no spatial effects

$$\mathbf{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim N(0, \sigma^2\mathbf{I}), \tag{1}$$

b) spatial error model

$$\mathbf{y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} = \lambda\mathbf{W}\boldsymbol{\varepsilon} + \mathbf{u}, \qquad \mathbf{u} \sim N(0, \sigma^2\mathbf{I}), \tag{2}$$

c) mixed regressive-spatial autoregressive model

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim N(0, \sigma^2\mathbf{I}) \tag{3}$$

d) mixed regressive-spatial autoregressive model with spatial autoregressive error.

$$\mathbf{y} = \rho\mathbf{W}\mathbf{y} + \boldsymbol{X\beta} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} = \lambda\mathbf{W}\boldsymbol{\varepsilon} + \mathbf{u}, \qquad \mathbf{u} \sim N(0, \sigma^2\mathbf{I}), \tag{4}$$

The comparison is realized for chosen set of spatial data.

Keywords

OLS model, Spatial error model, Mixed regressive-spatial autoregressive model, Spatial dependence, Spatial heterogeneity.

References

Anselin, L. (1988). *Spatial Econometrics: Methods and Models.* Dordrecht: Kluwer.

Arbia, G. (2006). *Spatial econometrics: Statistical foundations and applications to regional convergence.* Advances in Spatial Science. Berlin: Springer.

Griffith, D.A. (2003). *Spatial Autocorrelation and Spatial Filtering.* Berlin, Heidelberg, New York: Springer

Olejnik, A. (2008). Using the spatial autoregressively distributed lag model in assessing the regional convergence of per-capita income in the EU25. *Papers in Regional Science 83 (3)*, 371–384.