Local and Global Spatial Effects in Hierarchical Models

Donald J. Lacombe[a], Stuart G. McIntyre[b]

[a]*Regional Research Institute, West Virginia University, 886 Chestnut Ridge Road, PO BOX 6825, Morgantown, WV, 26506-6825,+1 (304) 293-8543; Fax: +1 (304) 293-6699, donald.lacombe@mail.wvu.edu*
[b]*Department of Economics, University of Strathclyde, Sir William Duncan Building, 130 Rottenrow, Glasgow G4 0GE, Scotland, UK; Phone: +44 (0)141 548 3858, s.mcintyre@strath.ac.uk*

Abstract

Hierarchical models have a long history in empirical applications; recognition of the fact that many datasets of interest to applied econometricians are nested; counties within states, pupils within school, regions within countries, etc. Just as many datasets are characterized by nesting, many are also characterized by the presence of spatial dependence or heterogeneity. Significant advances have been made in developing econometric techniques and models to allow applied econometricians to address this spatial dimension to their data. This paper fuses these two literatures together and combines a hierarchical random intercept model with the two *general* spatial econometric models.

1. Introduction

Hierarchical econometric models have a long pedigree in applied econometrics, and several textbook treatments are available (see for instance: Raudenbush and Bryk (2002); Franzese (2005); Gelman and Hill (2007)). More recently, attention has shifted to the development of hierarchical spatial econometric models (see for instance: Corrado and Fingleton (2012); Elhorst (2014)). In this paper we develop two hierarchical spatial econometric models; one representing a local spatial spillover process and the other a global spatial spillover process. We explain the reason for focusing on these two processes in more detail later in this paper. We develop these models

in the Bayesian paradigm following in the footsteps of early pioneers (e.g. Lindley and Smith (1972)) in this area.

Spatial econometric models have become increasingly popular in applied econometrics and beyond in the past two decades, and with this popularity has arisen challenges and evolutions in thinking about how best to incorporate spatial heterogeneity into econometric models. One influential school of thought in this area is reflected in a recent paper (LeSage, 2014), in which the author argues that the applied econometricians determining which cross sectional or panel spatial econometric model to estimate has only to ask themselves whether the spatial process they are seeking to model is one which is *global* or *local*.

Central to this discussion is the idea of spillovers. Following LeSage (2014) we can think of a spillover being where the rth characteristic of the ith entity (county say) X_i^r, has some influence upon the outcome Y of some neighbouring entity j, Y_j. A classic example would be that of cigarette taxes and the impact of the tax rate of one state (e.g. X_i^r) on the cigarette consumption Y_j of a neighbouring state j. A *local* spatial spillover process is where the spatial spillovers do not exhibit endogenous feedback effects. The spillovers in local models only affect the neighbouring observations as defined by the spatial weight matrix. This is represented by the commonly used spatial Durbin error model (SDEM). A *global* spillover in contrast is where there are endogenous feedback effects, and thus there is an impact of X_i^r on an outcome in all areas Y_j where j now includes higher order neighbours to i (i.e. neighbours to i's neighbours, etc). In this way, a change in X_i^r leads to system wide change and results in a new long-run equilibrium. This is represented using the popular spatial Durbin Model (SDM).

LeSage (2014) reviews the available, and often discussed, spatial econometric models to determine how the applied econometrician can best capture either a *global* or *local* spatial process in their model. He concludes that *global* spatial processes are best captured using the spatial Durbin model (SDM)

2

and *local* spatial processes are best captured using the spatial Durbin error model (SDEM). The central reason for this conclusion is that each of these models nests within it the other local or global models that are often discussed or used in applied work. Estimating one of these two *general* models allows the modeller to test the significance of the coefficients on the spatial terms and thereby test whether it is appropriate to move from a general to a more specific formulation (Elhorst (2014) neatly represents this general to specific approach for these spatial models).

2. Material and methods

In this section we begin by motivating our focus on hierarchical spatial econometric models, before extending the commonly used and recognised SDM and SDEM models to the hierarchical case.

Both the SDEM and SDM model are well established econometric models. Where little innovation has so far occurred is in developing these models in a hierarchical context in the Bayesian paradigm as we do in this paper. There are many reasons to believe that developing these models in this way presages significant econometric improvements. Many of the applications of the SDEM and SDM model in practice will involve using data that has some degree of nesting within it; for instance counties nested within states, school districts nested within counties, etc. The hierarchical extension of this class of models can take into account of the nested structure of the data as well as provide empirical estimates of spillovers.

One important advantage of hierarchical models is that they enable us to control for unobserved heterogeneity at the upper level of the hierarchy (which we refer to as Level 2) using indicator variables as well as including covariate information at that level, something which is generally not available in a standard fixed–effects model. Hierarchical models allow the researcher to control for unobserved heterogeneity via the use of indicator variables for each level 2 unit while also allowing for the use of covariates at this upper

level as well. What these spatial hierarchical models add, is to model any spatial spillovers in the covariates at this level in addition to the unobserved heterogeneity. To take the example of county and state taxes, while state level taxation is common across the lower level units (counties) nested within the upper level unit (states), there is a sizeable literature demonstrating that state level tax rates are not set and varied independently of the rates set in neighbouring states. Now, if we believe that there might be endogenous feedback in the setting of these state level taxes, i.e. a global spillover process, then the SDM model is the appropriate model to consider, likewise if we believe that the spillovers are not characterised by endogenous feedback then a *local* model may be more appropriate.

When we extend these hierarchical models to include spatial terms, we must also ensure that we calculate the correct marginal effects in the model in order to make sure that we are properly interpreting the marginal effects estimates[1]. The models outlined in this paper allow the applied econometrician to capture either *local* or *global* spillover effects, correctly interpret the marginal effects estimates, and to test for different spatial specifications within a general to specific framework.

These models have been developed in the Bayesian paradigm because recent evidence suggests that there are considerable advantages to be had in estimating hierarchical models using Bayesian rather than frequentist methods. Stegmueller (2013) compared these two approaches and showed that "under conditions considered in this study, Bayesian point estimates were biased at most 5%, whereas ML estimates reached 10 or 15%" (Stegmueller, 2013, 759). There appear to be real estimation advantages to Bayesian estimation methods, particularly where there are a small number of upper level

---

[1]It is well understood in the spatial econometrics literature that the partial derivatives on the explanatory variables in a spatial regression model with a lagged dependent variable are not equal to β and do not provide the true 'marginal' effects estimates given their matrix structure, see LeSage and Pace (2009) for more on this.

units and there is correlation between them. Having outlined the case for extending the SDEM and SDM model to the hierarchical case, in a Bayesian framework, we formally present the hierarchical SDM and SDEM models in the next section.

3. Theory/Calculation

3.1. Hierarchical SDEM model

The standard SDEM model takes the form:

$$
\begin{aligned}
y &= X\beta_1 + WX\beta_2 + u \\
u &= \rho W u + \epsilon \\
\epsilon &\sim N(0, \sigma_e^2 I_n)
\end{aligned}
\tag{1}
$$

where, y is the dependent variable, X are the covariates, W is the spatial weight matrix and WX are the spatially weighted covariates (one can think of these for each area i as the average value of each of the covariates in neighbouring areas j, where $i \neq j$). The hierarchical SDEM model takes into account potential spatial autocorrelation in the residuals of the upper (Level 2) model. The hierarchical SDEM model can be expressed as follows:

Level 1:

$$
\begin{aligned}
y_{ij} &= \alpha_j + \beta x_{ij} + \varepsilon_{ij} \\
\varepsilon_{ij} &\sim N(0, \sigma^2)
\end{aligned}
\tag{2}
$$

Level 2:

$$
\begin{aligned}
\alpha_j &= Z_j \gamma + W Z_j \theta + \delta_j \\
\delta_j &= \lambda W \delta_j + u_j \\
u_j &\sim N\left(0, \tau^2\right)
\end{aligned}
\tag{3}
$$

where Equation 2 represent the first level of the hierarchy and Equation 3

represents the upper–level of the hierarchy. We can also write the hierarchical SDEM model in matrix form as follows[2]:

$$y = X\beta + \Delta\alpha + \varepsilon \tag{4}$$

$$\alpha = Z\gamma + WZ\theta + u \tag{5}$$

Given the above information, we can now define the posterior distribution as follows:

$$\pi\left(\theta, \alpha \,|y\right) \propto \sigma^{-N} \exp\left\{-\frac{1}{2}(y - \Delta\alpha - X\beta)'C_\beta^{-1}\left(y - \Delta\alpha - X\beta\right)\right\} \tag{6}$$

$$\times |A| \; \tau^{-J} \exp\left\{-\frac{1}{2}(A\alpha - AZ^*\gamma)'C_\gamma^{-1}\left(A\alpha - AZ^*\gamma\right)\right\} \tag{7}$$

$$\times \sigma^{-1} \exp\left\{-\frac{1}{2\sigma^2}V_{\sigma_0^2}S_{\sigma_0^2}^2\right\} \tag{8}$$

$$\times \exp\left\{-\frac{1}{2}(\beta - c)'T_\beta^{-1}\left(\beta - c\right)\right\} \tag{9}$$

$$\times \tau^{-1} \exp\left\{-\frac{1}{2\tau^2}V_{\tau_0^2}S_{\tau_0^2}^2\right\} \tag{10}$$

$$\times \exp\left\{-\frac{1}{2}(\gamma - d)'T_\gamma^{-1}\left(\gamma - d\right)\right\} \tag{11}$$

$$\times U\left(-\frac{1}{\lambda_{\min}}, \frac{1}{\lambda_{\max}}\right) \tag{12}$$

where $C_\beta^{-1} \equiv (\sigma^2 I_n)^{-1}$, $C_\gamma^{-1} \equiv (\tau^2 I_j)^{-1}$, $A \equiv (I_n - \lambda W)$, $\lambda_{\min}$ is the smallest eigenvalue of the spatial weight matrix W, and $\lambda_{\max}$ is the largest eigenvalue of the spatial weight matrix W, which equals 1 if the spatial weight matrix is row–normalized. The row–standardization of the spatial weight

---

[2]The Δ term in the matrix form of the model is a matrix that assigns each Level 1 unit to its corresponding Level 2 group. Another way of thinking about this matrix is that it is the "dummy variable matrix" that one would use in a standard fixed–effects model.

matrix W is usually performed which would render the prior distribution for λ in equation (12) as $U\left(-\frac{1}{\lambda_{\min}}, 1\right)$.

The above expression for the posterior distribution of the parameters results from the multiplication of the likelihood times the priors for the various parameters. Equation 8 is the prior for the level–one error variance σ^2 while 9 is the prior distribution for the β parameters at level 1. Moving on to the level 2 priors, Equation 10 is the prior distribution for the level 2 error variance τ while Equation 11 is the prior distribution for the regression parameters at level 2, namely γ. Finally, Equation (12) is the prior distribution for the spatial autoregressive parameter λ. The full conditional distribution for the spatial autoregressive error parameter, λ, does not fall into any recognizable distributional form and thus a random–walk Metropolis–Hastings algorithm is utilized to obtain inferences for this parameter. Details regarding the use of a random–walk Metropolis–Hastings algorithm to draw inferences for spatial autocorrelation parameters is contained in LeSage and Pace (2009).

3.2. Hierarchical SDM model

The standard Spatial Durbin model (SDM) (see Elhorst (2014) or LeSage and Pace (2009) for a textbook treatment) takes the form:

$$y = \rho W y + \alpha_{ln} + X\beta + WX\theta + \epsilon$$
$$\epsilon \sim N(0, \sigma_e^2 I_n) \tag{13}$$

where, y is the dependent variable, X are the covariates, W is the spatial weight matrix and WX are the spatially weighted covariates. Extending this to the hierarchical case gives:

Level 1:
$$y_{ij} = \alpha_j + \beta x_{ij} + \varepsilon_{ij}$$
$$\varepsilon_{ij} \sim N(0, \sigma^2) \tag{14}$$

Level 2:
$$\alpha_j = \rho W \alpha_j + Z_j \gamma + WZ_j \theta + u_j$$
$$u_j \sim N\left(0, \tau^2\right) \tag{15}$$

This model can also be written in matrix form and the derivation of the posterior distribution for the hierarchical SDM model follows closely the derivation of the SDEM model previously outlined.

4. Results & Discussion

For space reasons we cannot present a fuller empirical example, however we provide some evidence of the accuracy of our code. To test our code and ensure that it is able to recover model parameters appropriately, simulated data was generated for both the SDM and SDEM models. In the results tables below we report in column two the values of the parameters used to generate the data used in these estimations, column three reports the estimates our model produced for each parameter.

Table 1 contains the values of the parameters generated for the SDEM model. For the Level 1 explanatory variables, we generated random normal variates and used two explanatory variables at this level with a value of 3 for each coefficient. The covariates at Level 2 (i.e. a constant term and two regressors, X and WX) were generated from a normal distribution with mean 0 and variance of 10 and were also given coefficient values of 3. The error variance at Level 1 (i.e. σ^2) was set to 1 and the error variance at level 2 (i.e. τ^2) was set to 1. The spatial autoregressive parameter for the SDEM model, λ, was set to 0.8 which is a moderate–to–high level of spatial error correlation. This is the same as the value used for the autoregressive parameter in the SDM model.

As can be seen from the results in Table 1 for the SDEM model and Table 2-3 for the SDM model, the Gibbs sampling algorithm for both the SDM and SDEM models recovers the parameters well, indicating that the code works as expected. In the case of the SDM model, following LeSage and Pace (2009), since we have a spatially lagged dependent variable we calculate the scalar summaries for each of the covariates, these are reported in Table 3. The direct effect is the effect of the covariate in area i on the dependent

variable in area i, while the indirect effect is the impact of the covariate in neighbouring area $j \neq i$, on the dependent variable in area i.

Table 1: Simulation Exercise Results: SDEM Model

Parameter	True Parameter Value	SDEM Estimate
β_1	3.00	3.0053
β_2	3.00	2.9902
σ^2	1.00	1.0182
τ^2	1.00	1.0657
λ	0.80	0.8045
γ_1	3.00	3.0940
γ_2	3.00	2.9844
γ_3	3.00	2.8943

Table 2: Simulation Exercise Results: SDM Model

Parameter	True Parameter Value	SDM Estimate
β_1	3.00	3.0128
β_2	3.00	2.9766
σ^2	1.00	1.0000
τ^2	1.00	1.0216
ρ	0.80	0.8047
γ_1	3.00	2.8230
γ_2	3.00	3.1066
γ_3	3.00	2.8229
γ_4	3.00	2.9175
γ_5	3.00	3.0038

Table 3: Simulation Exercise Results, Scalar Summaries: SDM Model

	Effect	True value	Lower 95%	Mean	Upper 95%
γ_4	Direct	5.7974	5.5306	5.8927	6.2579
	Indirect	24.6840	22.7396	22.7396	27.4019
	Total	30.4813	28.3403	28.3403	33.5965
γ_5	Direct	5.7974	5.2056	5.5346	5.8580
	Indirect	24.6840	22.5608	24.3228	26.1731
	Total	30.4813	27.8539	29.8574	31.9322

5. Conclusions

This paper has extended the traditional hierarchical regression model to incorporate local and global spatial spillovers. These models have been developed using Bayesian methods to take advantage of the estimation performance advantages demonstrated in estimating hierarchical models in this way. Space consideration have precluded an "real" empirical demonstration of these models in this study, but future work will use these models in an applied setting. Nevertheless, using generated data we have demonstrated that these models are well coded, and we encourage applied researchers to use these models in their studies using the MATLAB code available from the authors.

Corrado, L., Fingleton, B., 2012. Where is the economics in spatial econometrics? Journal of Regional Science 52, 210–239.

Elhorst, J.P., 2014. Spatial Econometrics: From Cross-Sectional Data to Spatial Panels. Springer.

Franzese, R.J., 2005. Empirical strategies for various manifestations of multilevel data. Political Analysis 13, 430–446.

Gelman, A., Hill, J., 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press.

LeSage, J., Pace, R.K., 2009. Introduction to spatial econometrics. CRC press.

LeSage, J.P., 2014. What regional scientists need to know about spatial econometrics. Available at SSRN 2420725 .

Lindley, D.V., Smith, A.F., 1972. Bayes estimates for the linear model. Journal of the Royal Statistical Society. Series B (Methodological) 34, 1–41.

Raudenbush, S.W., Bryk, A.S., 2002. Hierarchical linear models: Applications and data analysis methods. volume 1. Sage.

Stegmueller, D., 2013. How many countries for multilevel modeling? a comparison of frequentist and bayesian approaches. American Journal of Political Science 57, 748–761.