A SHORT PRIMER ON SPATIAL ECONOMETRICS

LIAM ELBOURNE[†]

[†]UNIVERSITY OF OXFORD
E-mail address: liam.elbourne@economics.ox.ac.uk.
Date: November 2022.

1. Introduction: History & Motivation

Ordinary Least Squares is, in many settings, a point of entry for empirical work due to its simple implementation and interpretation, as well as its attractive properties under the Gauss-Markov Theorem. However, it is commonplace for empirical work to require more advanced methods, often motivated by violations of the theorem's assumptions. This is the case for instrumental variables methods, such as two-staged least squares, which seek to remedy the issue of endogeneity. This is a direct analogy to the development of spatial econometrics, which was formalized by Anselin in his seminal book (1988) and subsequent paper (Anselin et al., 1996) to treat two problems that lead to violations of the theorem: spatial dependence and spatial heterogeneity. These violations pertain to forms of spatial correlation, for which there are three in a linear model of the form $y_i = \sum_{k \in K} \beta_k x_{k,i} + u_i$ (for i = 1,...,N):

(1) $Corr(y_i, y_j) \neq 0$ for some $j \neq i$
(2) $Corr(y_i, \mathbf{x}_j) \neq 0$ for some $j \neq i$
(3) $Corr(u_i, u_j) \neq 0$ for some $j \neq i$

As we will see below, spatial econometrics controls for spatial correlation by introducing spatially lagged terms of y or X (Wy or WX), or by modelling spatial correlation in the error term ($u = \lambda W u + \epsilon$). This forms a more general framework which nests the standard linear model; however, before we begin specifying spatial models we must first define the neighbour relation.

2. Defining Neighbours: Spatial Weight Specifications

One of the defining features of spatial econometrics is its explicit modelling of the neighbour relation. The neighbour relation between n spatial units is represented by an n×n binary (or row-standardized) matrix W, for which there are several common approaches for irregular aerial entities:

- Distance-Based Neighbours: Neighbours are defined according to some condition on distance between centroids of geographic units.

  - K Nearest Neighbours: j is a neighbour to i if the centroid of j is one of the k nearest centroids to that of i (non-symmetric).
  - Inverse Distance Neighbours: All i have $(n-1)$ neighbours, each weighted by the inverse distance between the centroids of i and j.
  - Within Distance Neighbours: j is a neighbour to i if it is within some distance d_{ij} of i .
- Graph-Based Weight Matrices:
  - Delauney Triangulation: j is a neighbour to i if its Thiessen polygon shares a common border with that of i.

Spatial weight matrices W are typically row-standardized and interacted with dependent, independent, or error components of the model. This results in the nice property that the spatially lagged terms Wy and WX are the neighbour averages of y and X, respectively.

But how do we know which neighbour specification to use? This is an important lingering question and crucial shortcoming in the spatial econometrics literature (Leenders, 2012). An approach to estimate the spatial weights matrix W would be a welcome addition to the literature but has not currently been developed successfully. Common practice suggests investigating the robustness of one's model to alternative specifications of W (Ertur & Koch, 2007) and selecting the optimal W based on a goodness-of-fit criteria ex-post (Stakhovych & Bijmolt, 2009). This is the approach taken in this paper. An alternative approach, developed by LeSage & Pace (2009), is based on Bayesian posterior model probability. See their book for further details.

3. A Taxonomy of Spatial Econometric Models

Given a vector of dependent variables y, a matrix of independent variables X, and spatial weight matrix W, we can specify various spatial econometric models. The literature, thanks to the contributions of Anselin (1988), Anselin et al. (1996), LeSage & Pace (2009), Kelejian & Prucha (1998), and others, have developed a series of nested models which we now outline. We begin with the most general model - the Manski model - before imposing a series of linear restrictions which deliver the standard linear regression model as a special case.

- Manski Model: $y = \rho Wy + X\beta + WX\theta + u$ where $u = \lambda Wu + \epsilon$
  - The Manski Model is the most comprehensive approach to modelling spatial correlation, including all 3 forms.
- Kelejian-Prucha Model: $y = \rho Wy + X\beta + u$ where $u = \lambda Wu + \epsilon$
  - Restricts $\theta = 0$ in the Manski Model
  - Popularized by Kelejian & Prucha (1998)
- Spatial Durbin Model (SDM): $y = \rho Wy + X\beta + WX\theta + \epsilon$
  - Restricts $\lambda = 0$ in the Manski Model, focusing on types 1 and 2 spatial correlation.
- Spatial Durbin Error Model (SDEM): $y = X\beta + WX\theta + u$ where $u = \lambda Wu + \epsilon$
  - Restricts $\rho = 0$ in the Manski Model, focusing on types 2 and 3 spatial correlation.
- Spatial Lag Model (SLY): $y = \rho Wy + X\beta + \epsilon$
  - Restricts $\theta = 0$ and $\lambda = 0$ from the Manski model, focusing on just type 1 spatial correlation.
- Spatially Lagged X (SLX) Model: $y = X\beta + WX\theta + \epsilon$
  - Restricts $\rho = 0$ and $\lambda = 0$ from the Manski model, focusing on just type 2 spatial correlation.
  - This model is not commonly referenced in the literature.
- Spatial Error Model (SEM): $y = X\beta + u$ where $u = \lambda Wu + \epsilon$
  - Restricts either $\rho = 0$ & $\theta = 0$ or $\lambda = 0$ & $\theta = -\rho\beta$ in the Manski model, focusing on just type 3 spatial correlation (referred to as spatial autocorrelation).
- Standard Linear Regression Model: $y = X\beta + \epsilon$
  - Restricts $\rho = 0$, $\theta = 0$, and $\lambda = 0$, ignoring all forms of spatial correlation

Figure 2.1 illustrates the nesting nature of these models. Note that the standard linear regression model can also be obtained as a special case of the spatial models listed by imposing a neighbour relation such that W is a matrix of zeros. This would most easily be observed using a k-nearest neighbours specification in which k=0.

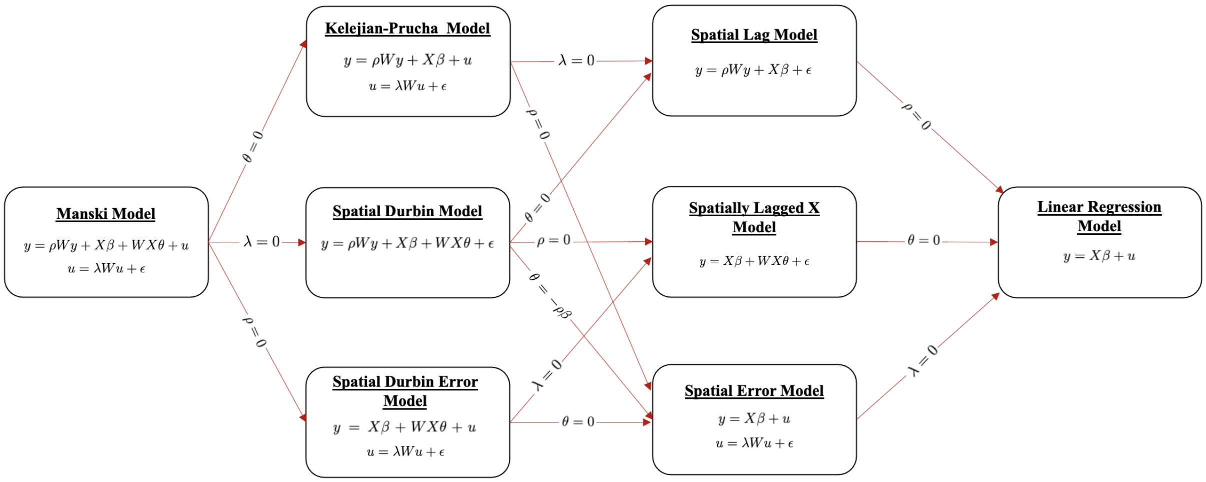

FIGURE 1. Illustration of the nested relationship of spatial econometric models, adapted from Elhorst (2010).

4. Estimation Procedure & Specification Tests

There are two contrasting viewpoints as to the optimal estimation procedure in spatial econometrics. The first, advocated for by LeSage & Pace (2009), takes a general-to-specific approach. They suggest beginning with the spatial Durbin model before undertaking likelihood-ratio (LR) tests to pare down to more specific models - namely, the spatial lag or spatial error models. Alternatively, Florax et al. (2003) argue that a specific-to-general approach is superior, begining with the linear regression model before expanding step-wise to spatial models if evidence of spatial autocorrelation is detected in the error term. This evidence is obtained from Anselin's classical (1988) or robust (1996) lagrange multiplier (LM) tests, or from Moran's I for regression residuals (Moran, 1950). Anselin's tests specify both the spatial lag model (SLY) and spatial error model (SEM) as explicit alternatives, and are as follows:

$$LM_{lag} = \frac{\left(\frac{u'Wy}{u'u/n}\right)^2}{D} \overset{a}{\sim} \chi_1^2 \tag{1}$$

$$LM_{err} = \frac{\left(\frac{u'Wu}{u'u/n}\right)^2}{T} \overset{a}{\sim} \chi_1^2 \tag{2}$$

$$RLM_{lag} = \frac{\left(\frac{u'Wy}{u'u/n} - \frac{u'Wu}{u'u/n}\right)^2}{D - T} \overset{a}{\sim} \chi_1^2 \tag{3}$$

$$RLM_{err} = \frac{\frac{u'Wu}{u'u/n} - TD^{-1}\frac{u'Wy}{u'u/n}}{T(1-TD)} \overset{a}{\sim} \chi^2_1 \tag{4}$$

where $D = [\frac{(WX\beta)'(I-X(X'X)^{-1}X')(WX\beta)}{(u'u/n)}] + tr(W^2 + W'W)$ and $T = tr(WW + W'W)$. These are calculated only for the standard linear regression specification outlined previously. By contrast, Moran's I for regression residuals proposes no alternatives and can be calculated for both the standard linear regression specification and all spatial regression specifications outlined previously. It is as follows:

$$Moran's\ I\ LM\ Test = \frac{n}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}} \frac{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(u_i - \bar{u})(u_j - \bar{u})}{\sum_{i=1}^{n}(u_i - \bar{u})^2} \tag{5}$$

Here, w_{ij} is the ijth element of W and u are regression residuals. The asymptotic distribution of Moran's I for regression residuals is standard normal after subtracting the mean of the regression residuals and dividing by the standard deviation (Cliff and Ord, 1981).

Faced with these contrasting approaches, Elhorst (2010) proposes the following hybrid procedure. First, estimate OLS and use Anselin's LM statistics (1988, 1996) to test whether the spatial lag or spatial error models are more appropriate. If OLS is rejected in favor of any of these alternatives, then proceed to the spatial Durbin model as advocated for by LeSage & Pace (2009). Once estimated, one can undertake LR tests to see whether the model should be pared down to the spatial lag model or spatial error model. The spatial lag or spatial error models should only be adopted if both the LM and LR tests point to its suitability. This is the estimation procedure we recommend.

5. Interpretation: Emanating Effects & Impact Measures

The interpretation of coefficients in spatial models differs from that in the classical linear regression setting. In the latter, the coefficient β_k is the partial derivative of y with respect to x_k and is thus interpreted as the effect of a unit increase in explanatory variable $x_{k,i}$ on outcome variable y_i, holding $x_{k,j}$ fixed for $j \neq i$. However, in spatial models with $\rho \neq 0$, coefficient interpretation is complicated by feedback, or emanating, effects resulting from the expansion of the information set to include neighbour effects. To see this mathematically, consider the spatial Durbin model:

$$y = \rho W y + X\beta + WX\theta + \epsilon$$

$$(I_n - \rho W)y = X\beta + WX\theta + \epsilon$$

$$y = \sum_{r=1}^{k} S_r(W)x_r + V(W)\epsilon$$

where $V(W) = (I_n - \rho W)^{-1}$ and $S_r(W) = V(W)[I_N\beta_r + W\theta_r]$. Then the partial derivative of y with respect to x_k takes the form:

$$\frac{\partial y}{\partial x_k} = S_k(W)$$

This necessitates an alternative procedure, for which LeSage & Pace (2006) develop a scalar summary measure. This measure decomposes the impact of a change in x_k on y into three parts: the average direct, average indirect, and average total effects. The average direct effect is the sum of the diagonal elements of the matrix $S_k(W)$ divided by n, the average total effect is the sum of all of the elements of $S_k(W)$ divided by n, and the average indirect effect is the difference between the two. Empirically, the matrix $S_k(W)$ can be approximated in R either analytically or by using the traces of powers of the spatial weight matrix as outlined by Bivand & Piras (2015) and Bivand et al. (2008).

6. Implementation in R

The spatial econometric methods described in this text can implemented in R using the *spdep* and *spatialreg* packages. Documentation for these methods can be found at *https://github.com/r-spatial/spdep/* and *https://r-spatial.github.io/spatialreg/* respectively, with additional applied discussion of the methods and package available in Bivand et al (2008).

We also note that the R source material does not make it clear whether the reported standard errors are classical or heteroskedasticity-robust, instead referring to them as 'asymptotic standard errors' which in general could refer to either. We are inclined to believe that they are the former based on the discussion in Piras (2010) and Bivand & Piras (2015), which suggest that none of the current ML-based implementation techniques for spatial models can accomodate heteroskedasticity-robust standard errors. Alternatively, it has come to our attention that the R package *sphet* has been developed by Piras (2010) to implement spatial regression models both with and without heteroskedastic errors using GMM, following Cliff & Ord (1981).

References

Anselin, L. (1988). Lagrange multiplier test diagnostics for spatial dependence and spatial heterogeneity. Geographical Analysis, 20(1), 1–17. https://doi.org/10.1111/j.1538-4632.1988.tb00159.x.

Anselin, L., Bera, A. K., Florax, R., & Yoon, M. J. (1996). Simple diagnostic tests for spatial dependence. Regional Science and Urban Economics, 26(1), 77–104. https://doi.org/10.1016/0166-0462(95)02111-6.

Bivand, R & Gianfranco Piras (2015). Comparing Implementations of Estimation Methods for Spatial Econometrics. Journal of Statistical Software, 23(18).

Bivand, R., Edzer J. Pebesma, & Virgilio Gomez-Rubio (2008). Applied Spatial Data Analysis with R. New York, NY: Springer.

Cliff, A.D. and Ord, J.K. (1981) Spatial processes: Models and applications. Pion Limited, London.

Elhorst, J. Paul (2010) Applied Spatial Econometrics: Raising the Bar, Spatial Economic Analysis, 5:1, 9-28, DOI: 10.1080/17421770903541772

Florax, Raymond & Folmer, Henk & Rey, Sergio. (2003). Specification Searches in Spatial Econometrics: The Relevance of Hendry's Methodology. Regional Science and Urban Economics. 33. 557-579. 10.1016/S0166-0462(03)00002-4.

Kelejian, H. H. & Prucha, I. R. (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, 17, 99-121.

LeSage, J. P. & Pace, R. K. (2009) Introduction to Spatial Econometrics, Boca Raton, Taylor & Francis.

Moran, P. A. P. "Notes on Continuous Stochastic Phenomena." Biometrika, vol. 37, no. 1/2, 1950, pp. 17–23. JSTOR, www.jstor.org/stable/2332142. Accessed 19 Mar. 2021.

Piras, Gianfranco (2010) sphet: Spatial Models with Heteroskedastic Innovations in R, Journal of Statistical Software June 2010, Volume 35, Issue 1.

Stakhovych, S. and Bijmolt, T.H. (2009), Specification of spatial models: A simulation study on weights matrices. Papers in Regional Science, 88: 389-408. https://doi.org/10.1111/j.1435-5957.2008.00213.x