Regional Research Institute Publications and Working Papers

Regional Research Institute

10-27-2011

A Note on Partitioning Effects Estimates Over Space

Christa D. Jensen

Donald J. Lacombe

Lett Spat Resour Sci (2012) 5:47–53
DOI 10.1007/s12076-011-0070-z

ORIGINAL PAPER

A note on partitioning effects estimates over space

Christa D. Jensen · Donald J. Lacombe

Received: 17 August 2011 / Accepted: 7 October 2011 / Published online: 27 October 2011

Abstract In this paper we provide an applied example for calculating the so-called *effects estimates* of LeSage and Pace (Introduction to Spatial, Econometrics, CRC Press, Boca Raton, 2009) for partitions of the impacts over space. While the partitioning of the impacts by orders of neighbors over space for the spatial autoregressive model is a relatively straightforward procedure, care must be taken in the case of the spatial Durbin model. The results from our empirical application regarding calculation of the partitioned effects over space for the spatial Durbin model corrects an error in the LeSage and Pace (Introduction to Spatial Econometrics, CRC Press, Boca Raton, 2009) text. We provide an illustration of these calculations for both models using a widely available data set on voter turnout for the 1980 United States presidential election.

Keywords Spatial econometrics · Interpretation · Marginal/partitioned effects · Effects estimates

JEL Classification C31 · C18

1 Introduction

In their recent book, *Introduction to Spatial Econometrics*, LeSage and Pace (2009) develop a methodology for correctly interpreting and summarizing the marginal im-

C.D. Jensen (✉)
Regional Research Institute, Department of Economics, West Virginia University,
886 Chestnut Ridge Road, P.O. Box 6825, Morgantown, WV 26506-6825, USA
e-mail: christa.jensen@mail.wvu.edu

D.J. Lacombe
Department of Economics, Department of Agricultural and Resource Economics,
Regional Research Institute, 886 Chestnut Ridge Road, P.O. Box 6825, Morgantown,
WV 26506-6825, USA
e-mail: Donald.Lacombe@mail.wvu.edu

pact of explanatory variables on the dependent variable in spatial econometric models. They refer to these marginal impacts as *effects estimates* and devote a substantial portion of their book to the theoretical and interpretational issues surrounding them. Spatial econometric models that include a spatial lag of the dependent variable, such as the spatial autoregressive (SAR) and spatial Durbin (SDM) models, require great care when making inferences based on estimated coefficients because the traditional assumption of independence (as in the case of ordinary least squares, or OLS) is not satisfied.[1] In other words, the partial derivative of the dependent variable with respect to an independent variable is not simply the coefficient estimate β, but a more complicated expression involving a *matrix* of effects estimates.

By design, spatial econometric models exploit complicated dependence structures that exist in geographically referenced data, such as counties, states, regions, or countries. This spatial dependence results in parameter estimates that contain a wealth of information on the relationship between changes in explanatory variables and how these changes affect the dependent variable. As noted by LeSage and Pace (2009, p. 33), "... the ability of spatial econometric models to capture these interactions represents an important aspect of spatial econometric modeling." One aspect of these effects estimates that may prove useful to applied practitioners is the ability to calculate how they vary over orders of neighbors. As noted in LeSage and Pace (2009, p. 40), "[i]t should be clear that impacts arising from a change in the explanatory variables will influence low-order neighbors more than higher-order neighbors." In other words, the effect of a change in an explanatory variable declines as we move over space.

The purpose of this paper is to point out the differences in the calculation of partitioned effects over space for the SAR and SDM models and to correct an error regarding the calculation of these effects over space for the SDM model in the LeSage and Pace (2009) text. Section 2 provides a concise theoretical background for the SAR and SDM models and shows how effects estimates are calculated and interpreted for each case. Section 3 discusses how these effects estimates can be partitioned over orders of neighbors and Section 4 provides results from an applied example using data on voter participation in the 1980 presidential election. Section 5 concludes.

2 Spatial autoregressive and spatial Durbin model

The SAR model is used when theoretical considerations or the results of diagnostic tests (e.g. Lagrange multiplier tests) suggest that spatial autocorrelation is present in the dependent variable. This type of model can be represented as follows:

$$\begin{aligned} y &= \rho W y + X\beta + \varepsilon \\ \varepsilon &\sim MVN(0, \sigma^2 I_n) \end{aligned} \tag{1}$$

where n is the number of observations, y is an $n \times 1$ vector of observations on the dependent variable, X is an $n \times k$ matrix of independent variables, ε is an $n \times 1$ vector

[1] In contrast, the coefficients from a spatial error model, or SEM, are interpreted in a similar manner as in the OLS case.

of i.i.d. errors, ρ is a scalar spatial autocorrelation parameter, β is a $k \times 1$ vector of regression parameters, and W is an $n \times n$ row stochastic spatial weight matrix.

To obtain the reduced form of the SAR model (i.e. the data generating process), where the y term does not appear on the right hand side of the equation, we take the following steps:

$$\begin{aligned} y &= \rho W y + X\beta + \varepsilon \\ y - \rho W y &= X\beta + \varepsilon \\ (I_n - \rho W)y &= X\beta + \varepsilon \\ y &= (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1}\varepsilon \end{aligned} \tag{2}$$

Now, if we examine the partial derivative of y with respect to X, we obtain:

$$\partial y / \partial X = (I_n - \rho W)^{-1}\beta \tag{3}$$

Note that $(I_n - \rho W)^{-1}$ is a matrix, which implies that we will get a matrix of effects estimates for each of our independent variables. We can also formally express the above ideas following the notation used in LeSage and Pace (2009):

$$(I_n - \rho W)y = X\beta + \iota_n \alpha + \varepsilon \tag{4}$$

$$y = \sum_{r=1}^{k} S_r(W)x_r + V(W)\iota_n\alpha + V(W)\varepsilon \tag{5}$$

where r represents the rth explanatory variable, $S_r(W) = V(W)(I_n\beta_r)$, $V(W) = (I_n - \rho W)^{-1}$, $\iota_n\alpha$ is the intercept term, and ε is the usual error term. We can express the above idea in the following matrix:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} S_r(W)_{11} & S_r(W)_{12} & \cdots & S_r(W)_{1n} \\ S_r(W)_{21} & S_r(W)_{22} & \cdots & S_r(W)_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_r(W)_{n1} & S_r(W)_{n2} & \cdots & S_r(W)_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \tag{6}$$

Given the above matrix representation of the effects estimates, we can now define the own-partial derivative and the cross-partial derivative:

$$\partial y_i / \partial x_{ir} = S_r(W)_{ii} \tag{7}$$

$$\partial y_i / \partial x_{jr} = S_r(W)_{ij} \tag{8}$$

The own-partial derivative is how a change in an explanatory variable in location i affects the dependent variable in location i and the cross-partial derivative is how a change in an explanatory variable in location j affects the dependent variable in location i (where $i \neq j$).

The calculations of the matrix of effects estimates for the SDM model and the resulting own- and cross-partial derivatives follow the same mathematical derivation with the exception that the $S_r(W)$ term is replaced with the following:

$$S_r(W) = V(W)(I_n\beta_r + W\theta_r) \tag{9}$$

where $V(W) = (I_n - \rho W)^{-1}$ still holds. The obvious difference is that we must take into account the spatially weighted explanatory variables when calculating the effects estimates for the SDM model.

Further details regarding the calculation of these effects estimates and their scalar summaries are discussed by LeSage and Pace (2009, Chap. 2). Elhorst (2010, pp. 18–22) provides a simplified example of how the effects estimates are calculated. We now turn our attention to the calculation of effects estimates over orders of neighbors.

3 Calculation of effects over orders of neighbors

LeSage and Pace (2009, p. 40) note that we can now rewrite the expression for $(I_n - \rho W)^{-1}$ as follows:

$$(I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \rho^3 W^3 + \cdots + \rho^q W^q \tag{10}$$

Thus, using the above expansion, we can calculate the impact associated with each power of W for the SAR model:

$$(I_n - \rho W)^{-1}\beta_r = \underbrace{I_n\beta_r}_{W^0 \text{ or } I_n} + \underbrace{\rho W\beta_r}_{W^1} + \underbrace{\rho^2 W^2\beta_r}_{W^2} + \cdots + \rho^q W^q \beta_r \tag{11}$$

The impact associated with each power of W can be calculated for the case of the SDM model in a similar manner:

$$(I_n - \rho W)^{-1}(I_n\beta_r + W\theta_r) \tag{12}$$

$$= (I_n + \rho W + \rho^2 W^2 + \cdots)(I_n\beta_r + W\theta_r) \tag{13}$$

$$= (I_n\beta_r + W\theta_r) + \rho W(I_n\beta_r + W\theta_r) + \rho^2 W^2(I_n\beta_r + W\theta_r) + \cdots \tag{14}$$

$$= \underbrace{(I_n\beta_r + W\theta_r)}_{W^0 \text{ or } I_n} + \underbrace{\left(\rho W I_n\beta_r + \rho W^2\theta_r\right)}_{W^1} + \underbrace{\left(\rho^2 W^2 I_n\beta_r + \rho^2 W^3\theta_r\right)}_{W^2} + \cdots \tag{15}$$

The powers of the weight matrix within the expansion of $(I_n - \rho W)^{-1}$ correspond to the different orders of neighbors: zero-order neighbors (or the I_n term), first-order neighbors (W), second-order neighbors (W^2), and so on. Using this expansion, we can examine how the impacts of explanatory variables manifest themselves over space for each of these two models.

Table 1 Summary of SAR effects estimates over space for *Education*

W-order	Direct	Indirect	Total
W^0 (I_n)	0.2210	0.0000	0.2210
W^1	0.0000	0.1295	0.1295
W^2	0.0126	0.0633	0.0759
W^3	0.0024	0.0421	0.0445
W^4	0.0018	0.0243	0.0261
W^5	0.0007	0.0146	0.0153
W^6	0.0004	0.0086	0.0090
W^7	0.0002	0.0051	0.0052
W^8	0.0001	0.0030	0.0031
W^9	0.0001	0.0018	0.0018

4 An applied example

The calculation of the effects estimates will differ based on the type of spatial econometric model an applied researcher uses. We illustrate the calculation of the effects estimates over space using an example from LeSage and Pace (2009) that contains information for 3,107 counties in the United States on voter participation in the 1980 presidential election.[2] The dependent variable is defined as those voting as a logged proportion of those eligible to vote. The explanatory variables include the proportion of the population over age 18 (*Voting Pop*), the proportion of the population with college degrees (*Education*), the proportion of the population that own homes (*Homeowners*), and median household income (*Income*). Note that all of the variables used in this analysis are logged, which means that the computed marginal effects can be interpreted as elasticities.

Since our main concern is the calculation of the effects estimates over space, we present in Table 1 the partitioned effects for one of the explanatory variables, *Education*. We expect *ex ante* that increases in the proportion of the population with college degrees will lead to increases in voter turnout and that is indeed the case for this particular explanatory variable. Results in Table 1 show that for all of the effects estimates, the effect of this covariate on the dependent variable declines as the order of neighbors increases. Of particular note is the pattern of the effects estimates. For the zero-order neighbors, the indirect effect is zero, which is always the case for the SAR model. From the series expansion in (11), note that, by definition, the first term (i.e. the "W^0 or I_n" term) contains zeros on the off-diagonal elements. For this reason, the indirect effects for the zero-order term will always be equal to zero for the SAR model. Similarly, the direct effects for the "W^1" term, or the neighbors as defined by the spatial weight matrix W, are also zero because by definition the spatial weight matrix W contains zeros on the main diagonal.

The SDM model, due to its increased complexity relative to the standard SAR model, exhibits much different behavior in terms of the effects estimates over space.

[2]This data set originally appeared in a paper by Pace and Barry (1997) and is available in Jim LeSage's Econometrics Toolbox.

Table 2 Summary of SDM effects estimates over space for *Education*

W-order	Direct	Indirect	Total
W^0 (I_n)	0.1352	0.0998	0.2351
W^1	0.0110	0.1455	0.1565
W^2	0.0124	0.0919	0.1043
W^3	0.0042	0.0653	0.0696
W^4	0.0028	0.0437	0.0464
W^5	0.0014	0.0296	0.0310
W^6	0.0008	0.0199	0.0207
W^7	0.0005	0.0134	0.0139
W^8	0.0003	0.0090	0.0093
W^9	0.0002	0.0061	0.0062

Table 2 contains the partitioned effects for the *Education* variable using the SDM model for estimation. As before, we expect *ex ante* that increases in the proportion of the population with college degrees will lead to increases in voter turnout. The results in Table 2 confirm this hypothesis as the direct, indirect, and total effects estimates are all positive.

As in the case of the SAR model, we wish to pay particular attention to the direct and indirect effects over space for the own (i.e. zero-order or W^0) and first-order neighbor (i.e. W^1) relationships. Notice now that the indirect effects are *not equal to zero* for the zero-order neighbors and that the direct effects for the first-order neighbors are also *not equal to zero*. The reason for this discrepancy between the models can be explained by examining the power series expansion of the $S_r(W)$ term for the SDM model as shown in (15).

The first term in (15), labeled "W^0 or I_n", provides an expression for the zero-order neighbor relationship. Instead of only having elements on the main diagonal of this term, there are now non-zero values in the off-diagonal elements as well. This is due to the presence of the W matrix associated with the spatially weighted explanatory variables, which is an integral part of the SDM model. As can be seen in Table 2, the indirect effect of the zero-order neighbors now contains a non-zero value. Likewise, the first-order neighbors now exhibit direct effects because the second term in (15), labeled "W^1", contains a W^2 term. When the spatial weight matrix is taken to the power of 2, elements on the main diagonal cease to be equal to zero. Therefore, there are now direct effects associated with first-order neighbors.

LeSage and Pace (2009, p. 72), in their Table 3.5 entitled "Marginal spatial partitioning of impacts", show that the partitioned effects from an SDM model contain no indirect effects for the zero-order neighbors and no direct effects for the first-order neighbors. LeSage and Pace (2009, p. 72) note that "[d]irect effects for W^1 will equal zero and the indirect effects for W^0 equal zero as discussed in Chap. 2." However, this is only true if the SDM model can be reduced to the SAR model, i.e. the spatially weighted explanatory variables as a whole are not statistically different from zero. In general, the zero-order indirect effects and the first-order direct effects for the SDM model take on values that are different from zero. The magnitude of these effects will depend on several factors as pointed out in LeSage and Pace (2009).

5 Conclusion

Spatial econometric models such as the spatial autoregressive (SAR) and the spatial Durbin models (SDM) allow for a much richer interpretation of empirical results than previously thought. The work of LeSage and Pace (2009) has provided theoretical as well as practical results regarding the so-called *effects estimates* and their interpretation. LeSage and Pace (2009) also illustrate how one can calculate the marginal, or spatially partitioned effects estimates from either a SAR or SDM model. These partitioned effects estimates show how the effects of changes in explanatory variables change over space. The LeSage and Pace (2009) text notes that there are no differences in the calculation of effects estimates over space for the SAR and SDM model. However, our mathematical and empirical analysis illustrates that there will be differences in these calculations and corrects the record regarding these calculations. The work of LeSage and Pace (2009) and LeSage and Dominguez (2010) has demonstrated that care must be taken when interpreting the results from various spatial econometric models. In this short paper, we show that the marginal, or spatially partitioned effects estimates, must also be interpreted with care.

References

Elhorst, J.P.: Applied spatial econometrics: raising the bar. Spatial Econ. Anal. **5**(1), 9–28 (2010)

LeSage, J.P., Dominguez, M.: The importance of modeling spatial spillovers in public choice analysis. Public Choice (2010). doi:10.1007/s11127-010-9714-6

LeSage, J.P., Pace, R.K.: Introduction to Spatial Econometrics. CRC Press, Boca Raton (2009)

Pace, R.K., Barry, R.P.: Quick computation of spatial autoregressive estimators. Geogr. Anal. **29**, 232–246 (1997)