5.3.4 Impact analysis

Direct and indirect impacts in the SAR model

Spatial autoregressive model (SAR):

$$\mathbf{y} = \mathbf{X\beta} + \rho\mathbf{W}\mathbf{y} + \boldsymbol{\varepsilon}$$

Data generating process (DGP) of SAR model:

$$\mathbf{y} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X\beta} + (\mathbf{I} - \rho\mathbf{W})^{-1}\boldsymbol{\varepsilon}$$

Example for direct and indirect impacts (LeSage/Pace, 2009):

Travel times (Y) to a central business district CBD in relation to population density (X_1) and distance (X_2)

Travel times (Y) | Population density (X_1) | Distance (X_2)

$$\mathbf{y} = \begin{pmatrix} 42 \\ 37 \\ 30 \\ 26 \\ 30 \\ 37 \\ 42 \end{pmatrix} \qquad \mathbf{X} = \begin{pmatrix} 10 & 30 \\ 20 & 20 \\ 30 & 10 \\ 50 & 0 \\ 30 & 10 \\ 20 & 20 \\ 10 & 30 \end{pmatrix}$$

ex-urban areas *R1*
far suburbs *R2*
near suburbs *R3*
CBD *R4*
near suburbs *R5*
far suburbs *R6*
ex-urban areas *R7*

Impact of a change in population density for a single region on travel time to CBD for all seven regions: doubling population density in region *R2*

Predictions from SAR model based on explanatory variables X_1 and X_2

$$\hat{\mathbf{y}} = (\mathbf{I} - \rho\mathbf{W})^{-1}\mathbf{X}\boldsymbol{\beta}$$

Predictions with original observation matrix **X**: $\hat{\mathbf{y}}^{(1)}$

Predictions with original observation matrix **X** after doubling population density in region R2: $\hat{\mathbf{y}}^{(2)}$

ML parameter estimates for the SAR model:

$$\hat{\beta}_{1,\mathrm{ML}} = 0.135, \quad \hat{\beta}_{2,\mathrm{ML}} = 0.561, \quad \hat{\rho}_{\mathrm{ML}} = 0.642$$

Table 1: Spatial spillovers form changes in Region R2 population density

Regions/Scenario	$\hat{\mathbf{y}}^{(1)}$	$\hat{\mathbf{y}}^{(2)}$	$\hat{\mathbf{y}}^{(2)} - \hat{\mathbf{y}}^{(1)}$
R1:	42.01	44.58	2.57
R2:	37.06	41.06	4.00
R3:	29.94	31.39	1.45
R4: CBD	26.00	26.54	0.53
R5:	29.94	30.14	0.20
R6:	37.06	37.14	0.07
R7:	42.01	42.06	0.05

- Direct effect: Increase of commuting time for residents in region R2 by 4 min.
- Indirect effects (spillover eff.): Increase of commuting times in region R1, R3- R7 by 4.87 min. (= 2.57 + 1.45 + 0.53 + 0.20 + 0.07 + 0.05)

Direct impacts: $\partial y_i / \partial x_{ij}$ (own derivative)

Effects on y_i resulting of a change in the j*th* explanatory variable X_j in the i*th* region

Indirect impacts: $\partial y_h / \partial x_{ij}$, $h \neq i$ (cross-partial derivatives)

Effects on y_h , $h \neq i$, resulting of a change in the j*th* explanatory variable X_j in the i*th* region

Restatement of the data generating process (DGP) of SAR model:

$$\mathbf{y} = \sum_{j=1}^{k} \mathbf{S}_j(\mathbf{W}) \cdot \mathbf{x}_j + (\mathbf{I}_n - \rho \mathbf{W})^{-1} \boldsymbol{\varepsilon})$$

with $\mathbf{S}_j(\mathbf{W}) = (\mathbf{I}_n - \rho \mathbf{W})^{-1} \beta_j$

”multiplier” matrix applies higher-order neighboring relations to $\mathbf{x}_j$

$\mathbf{x}_j$: nx1 vector of observations of the k*th* explanatory variable X_k

Role of the matrix $\mathbf{S}_j(\mathbf{W})$:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \sum_{j=1}^{k} \begin{pmatrix} S_j(\mathbf{W})_{11} S_j(\mathbf{W})_{12} \ldots S_j(\mathbf{W})_{1n} \\ S_j(\mathbf{W})_{21} S_j(\mathbf{W})_{22} \\ \vdots \quad \vdots \quad \ddots \\ S_j(\mathbf{W})_{n1} S_j(\mathbf{W})_{n2} \ldots S_j(\mathbf{W})_{nn} \end{pmatrix} \begin{pmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix} + (\mathbf{I}_n - \rho \cdot \mathbf{W})^{-1} \boldsymbol{\varepsilon}$$

Observation y_i dependent on the elements of the i*th* row of the matrix $\mathbf{S}_j(\mathbf{W})$:

$$y_i = \sum_{j=1}^{k} \left[S_j(\mathbf{W})_{i1} \cdot \mathbf{x}_{1j} + S_j(\mathbf{W})_{i2} \cdot \mathbf{x}_{2j} + ,\ldots, + S_j(\mathbf{W})_{in} \cdot \mathbf{x}_{nj} \right] + (\mathbf{I}_n - \rho \mathbf{W})_i^{-1} \boldsymbol{\varepsilon}$$

Series expansion of the matrix $(\mathbf{I}_n\text{-}\rho\mathbf{W})^{-1}$:

$$(\mathbf{I}_n - \rho \mathbf{W})^{-1} = \mathbf{I}_n + \rho \mathbf{W} + \rho^2 \mathbf{W}^2 + \rho^3 \mathbf{W}^3 + \ldots$$ (infinite power series)

Series expansion of the matrix $(\mathbf{I}_n\text{-}\rho\mathbf{W})^{-1}$ up to the q*th* order:

$$(\mathbf{I}_n - \rho \mathbf{W})^{-1} \approx \mathbf{I}_n + \rho \mathbf{W} + \rho^2 \mathbf{W}^2 + \ldots + \rho^q \mathbf{W}^q$$ (finite power series)

- Impacts of changes in x_{ij} on y_i:

$$\frac{\partial y_i}{\partial x_{ij}} = S_j(\mathbf{W})_{ii}$$ (own derivative for the i*th* region)

Standard regression model: $\frac{\partial y_i}{\partial x_{ij}} = \beta_j$

The expression $S_j(\mathbf{W})_{ii}$ measures the impact of a change in x_{ij} on the i*th* observation of the dependent variable y_i. This impact includes the effect of feedback loops where the observation i affects observation h which in turn also affects observation i. Also longer paths e.g. from i to h to l and back to i are captured by the expression.

- Impacts of changes in x_{hj} on y_i:

$$\frac{\partial y_i}{\partial x_{hj}} = S_j(\mathbf{W})_{ih}\,, \quad h \neq i$$ (cross-partial derivatives)

Standard regression model: $\frac{\partial y_i}{\partial x_{hj}} = 0$

The expression $S_j(\mathbf{W})_{ih}$ measures the impact of a change in x_{hj} on the i*th* observation of the dependent variable y_i. This impact likewise includes the effect of feedback loops where the observation h affects observation l which in turn affects observation i. Also longer paths e.g. from h to l to r and then to i are captured by the expression.

Summary measures of impacts

1. Average direct Impact

Summarizing the impacts of changes in the i*th* observation of X_j on y_i by averaging the elements $\mathbf{S}_j(\mathbf{W})_{ii}$

Average direct impact: $\overline{M}(x_j)_{dir} = \frac{1}{n} tr(\mathbf{S}_j(\mathbf{W}))$

Average response of dependent to independent variable.

2. Average total impact

Average total impact: $\overline{M}(x_j)_{tot} = \frac{1}{n} \cdot \boldsymbol{\iota}' \cdot \mathbf{S}_j(\mathbf{W}) \cdot \boldsymbol{\iota}$

$\boldsymbol{\iota}$ is an nx1 column vector of ones: $\underset{nx1}{\boldsymbol{\iota}} = (1\ 1\ \ldots\ 1)'$

2a) *Average Total Impact to an Observation*

The sum across the i*th* row of $\mathbf{S}_j(\mathbf{W})$ as the total impact from changing the j*th* explanatory variable X_j in all regions by the same amount δ, $\mathbf{x}_j + \delta$, on the individual observation y_i

The n sums are given by the column vector $\mathbf{c}_j$: $\mathbf{c}_j = \mathbf{S}_j(\mathbf{W}) \cdot \boldsymbol{\iota}$

Average of these total impacts: $\overline{M}(x_j)_{tot} = \frac{1}{n} \cdot \boldsymbol{\iota}' \cdot \mathbf{c}_j$

2b) *Average Total Impact from an Observation*

The sum down the h*th* column of $\mathbf{S}_j(\mathbf{W})$ as the total impact from changing the j*th* explanatory variable X_j in the h*th* region by the amount δ, $x_{hj}+\delta$, over all y_i

The n sums are given by the row vector $\mathbf{r}_j$: $\mathbf{r}_j = \boldsymbol{\iota}' \cdot \mathbf{S}_j(\mathbf{W})$

Average of these total impacts: $\overline{M}(x_j)_{tot} = \frac{1}{n} \cdot \mathbf{r}_j \cdot \boldsymbol{\iota}$

Both forms of average total impacts are equal.

3. Average indirect Impact

The average of all derivatives (=**average total impact**) **net of** the average of all own derivatives (=**average direct impact**) equals the average cross derivative (= **average indirect impact**).

Average indirect impact: $\overline{M}(x_j)_{indir} = \overline{M}(x_j)_{tot} - \overline{M}(x_j)_{dir}$

4. Measures of dispersion

Measures of dispersion for the impact estimates are needed for significance testing. They can be constructed from Monte Carlo simulation by which the sampling distribution of the average impact estimates $\overline{M}(x_j)$ is obtained.

Example (LeSage/Pace, 2009):

In a SAR model, **cumulative and spatially partitioned direct, indirect and total impacts** of an explanatory variable X_j are computed using the parameter values

$$\beta_j = 0.5 \text{ and } \rho = 0.7$$

The cumulative effects are spatialy partitioned up to the 9*th* order.

Cumulative direct effect: 0.5860 (0.05834 reached up to 9*th* order neighbours)
Involved feedback effects: 0.5860 – ($\beta_j = 0.5$) = 0.086:

Cumulative indirect effects: 1.0841
Higher order neighbourhood effects: 1.0841 – ($\beta_j \cdot \rho = 0.5 \cdot 0.7 = 0.35$) = 0.7341

for $\mathbf{W}^0$ (q=0: $\mathbf{W}^0 = \mathbf{I}$) : no indirect effect ($\mathbf{S}_j(\mathbf{W}) = \mathbf{I} \cdot \beta_j = 0.5 \cdot \mathbf{I}$)

for W^1 (q=1: $\mathbf{W}^1 = \mathbf{W}$): no direct feedback effect (only zero diagonal elements in **W**)

for $\mathbf{W}^2$: The term $\rho \cdot \mathbf{W}^2 \cdot \beta_j$ captures both direct and indirect effects as $\mathbf{W}^2$ beside zero diagonal elements in general also contains non-zero diagonal elements. In the latter case direct effects are generated by loops e.g. from i to h to i . Indirect effects come from paths e.g. from i to h to l.

Table 2: Spatial partitioning of direct, indirect and total impacts

	Cumulative Effects		
	Mean	Std. dev	*t*-statistic
Direct effect	0.5860	0.0148	39.6106
Indirect effect	1.08414	0.0587	18.4745
Total effect	1.6700	0.0735	22.7302
	Spatially Partitioned Effects		
Order of **W**	Total	Direct	Indirect
$\mathbf{W}^0$	0.5000	0.5000	0
$\mathbf{W}^1$	0.3500	0	0.3500
$\mathbf{W}^2$	0.2452	0.0407	0.2045
$\mathbf{W}^3$	0.1718	0.0144	0.1574
$\mathbf{W}^4$	0.1204	0.0114	0.1090
$\mathbf{W}^5$	0.0844	0.0066	0.0778
$\mathbf{W}^6$	0.0591	0.0044	0.0778
$\mathbf{W}^7$	0.0415	0.0028	0.0386
$\mathbf{W}^8$	0.0291	0.0019	0.0272
$\mathbf{W}^9$	0.0204	0.0012	0.0191
$\sum_{\ell=0}^{q=9}\mathbf{W}^{\ell}$	1.6220	0.5834	1.0386

Spatially partitioned indirect effects < direct effect (own-region effect) of order zero

Cumulative indirect effects > cumulative direct effect

From lower- to higher-order neighbours: decay of spatial spillover effects