SOCIAL
SCIENCES

TECHNISCHE UNIVERSITÄT
KAISERSLAUTERN

Spatial Regression Models *

Georg-August-Universität Göttingen

Tobias Rüttenauer

Fachbereich für Sozialwissenschaften
TU Kaiserslautern

April 2, 2019

TU Kaiserslautern

* This course profited a lot from Scott Cook's Spatial Econometrics course at Essex

Outline I

Part I

Spatial Weights and Autocorrelation

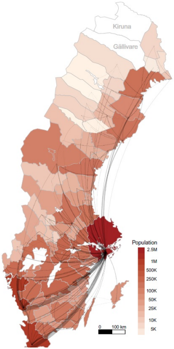
Introduction

Tobler's first law of geography

- 'everything is related to everything else, but near things are more related than distant things' (Tobler, 1970, p. 236)

In applied research

- Making a Place for Space (Logan, 2012)
- Context / peer / spillover effects
- Geo-coded micro data (e.g. Keuschnigg et al., 2019)
- Spatially aggregated data (e.g. Downey, 2006)
- Remote sensing (e.g. Weigand et al., 2019)



Spatial Regression

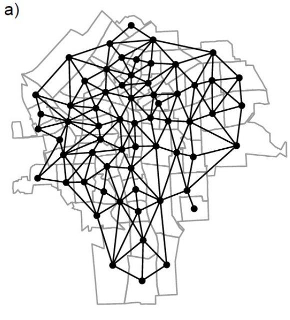
Aim

- Obtaining correct estimates with spatial data
- Obtaining correct inference with spatial data
- Investigating spatial relations

Different kinds of spatial relations

- Common exposure / Clustering on unobservables
- Common exposure / Clustering on observables
- Spillover / externalities from covariates
- Interdependence in outcomes / diffusion

W: Connectivity between units



(Bivand, 2018)

Spatial weights matrices

The spatial weights matrix $\boldsymbol{W}$ is an $N \times N$ dimensional matrix with elements w_{ij} specifying the relation or connectivity between each pair of units i and j.

$$\boldsymbol{W} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1n} \\ w_{21} & w_{22} & \dots & w_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1} & w_{n2} & \dots & w_{nn} \end{bmatrix} \tag{1}$$

Note: The diagonal elements $w_{i,i} = w_{1,1}, w_{2,2}, \dots, w_{n,n}$ of $\boldsymbol{W}$ are always zero. No unit is a neighbour of itself.

Spatial weights matrices

Multiplying any variable vector $\boldsymbol{x}$ with the spatial weights matrix $\boldsymbol{W}$ produces the 'spatially lagged' variable vector. For each unit i (each row), the spatially lagged variable vector $\boldsymbol{Wx}$ contains the spatially weighted values of the neighbouring units (as defined by $\boldsymbol{W}$).

$$\begin{aligned}\boldsymbol{Wx} &= \begin{bmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{bmatrix} \times \begin{bmatrix} 4 \\ 2 \\ 9 \end{bmatrix} \\ &= \begin{bmatrix} 0 \times 4 + 1 \times 2 + 0 \times 9 \\ 0.5 \times 4 + 0 \times 2 + 0.5 \times 9 \\ 0 \times 4 + 1 \times 2 + 0 \times 9 \end{bmatrix} = \begin{bmatrix} 2 \\ 6.5 \\ 2 \end{bmatrix}\end{aligned} \tag{2}$$

Spatial weights matrices

Things to consider

- Appropriate spatial representation
- Correct specification of connectivity (let theory decide?)
- Correct way of normalizing $\boldsymbol{W}$

$\Rightarrow$ In practical research: Very little guidance on correct choices

(LeSage and Pace, 2014; Neumayer and Plümper, 2016)

Contiguity weights

A very common type of spatial weights. Binary specification, taking the value 1 for neighbouring units (queens: sharing a common edge; rook: sharing a common border), and 0 otherwise.

- Contiguity weights $w_{i,j} = \begin{cases} 1, \text{if } i \text{ and } j \text{ neighbours} \\ 0, \text{otherwise} \end{cases}$

$$\boldsymbol{W} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

- Sparse matrices
- Problem of 'island' (units without neighbours)

Distance based weights

Another common type uses the distance d_{ij} between each unit i and j.

- Inverse distance weights $w_{i,j} = \frac{1}{d_{ij}}$

$$\boldsymbol{W} = \begin{bmatrix} 0 & \frac{1}{d_{ij}} & \frac{1}{d_{ij}} \\ \frac{1}{d_{ij}} & 0 & \frac{1}{d_{ij}} \\ \frac{1}{d_{ij}} & \frac{1}{d_{ij}} & 0 \end{bmatrix}$$

- Dense matrices
- Specifying thresholds may be useful

Normalization

Normalizing ensures that the parameter space of the spatial multiplier is restricted to $-1 < \rho > 1$, and the multiplier matrix is non-singular. Normalizing your weights matrix is always a good idea. Otherwise, the spatial parameters might blow up – if you can estimate the model at all.

Row normalization

Row-normalization divides each non-zero weight by the sum of all weights of unit i, which is the sum of the row.

$$\frac{w_{ij}}{\sum_j^n w_{ij}}$$

- Spatial lags are average values of neighbours
- Proportions between units (distance based) get lost
- Can induce asymmetries: $w_{ij} \neq w_{ji}$

(LeSage and Pace, 2014; Neumayer and Plümper, 2016)

Maximum eigenvalues normalization

Maximum eigenvalues normalization: Divide each non-zero weight by overall maximum eigenvalue λ_{max}. Each element of $\boldsymbol{W}$ is divided by the same scalar parameter.

$$\frac{\boldsymbol{W}}{\lambda_{max}}$$

- Interpretation may become more complicated
- Keeps proportions (relevant esp. for distance based $\boldsymbol{W}$)

(LeSage and Pace, 2014; Neumayer and Plümper, 2016)

Autocorrelation: Why care?

If spatially close observations are more likely to exhibit similar values, we cannot handle observations as if they were independent.

$$\mathrm{E}(\varepsilon_i \varepsilon_j) \neq \mathrm{E}(\varepsilon_i)\mathrm{E}(\varepsilon_j) = 0$$

This violates a basic assumption of the conventional OLS model. In consequence, ignoring spatial dependence can lead to

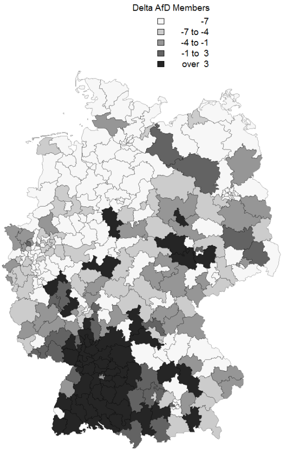
- biased inferencial statistics
- biased point estimates (depending on the DGP)

Detection of spatial dependence

Given the available dataset and a spatial weights matrix $\boldsymbol{W}$, how can we test for spatial autocorrelation?

- Visualisation
- Moran's I
- (Likelihood Ratio test)
- Lagrange Multiplier test

Visualisation



(Schulte-Cloos and Rüttenauer, 2018)

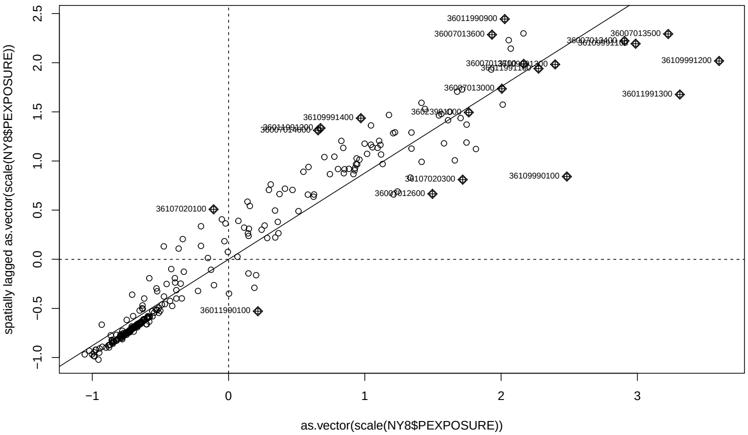
Moran's I

Global Moran's I test statistic:

$$I = \frac{N}{S_0} \frac{\sum_i \sum_j w_{ij}(y_i - \bar{y})(y_j - \bar{y})}{\sum_i (y_i - \bar{y})}, \text{where } S_0 = \sum_{i=1}^{N} \sum_{j=1}^{N} w_{ij} \qquad (3)$$

- Relation of the deviation from the mean value between unit i and neighbours of unit i.
- Negative values: negative autocorrelation
- Around zero: no autocorrelation
- Positive values: positive autocorrelation

Moran's I


spatially lagged as.vector(scale(NY8$PEXPOSURE))
as.vector(scale(NY8$PEXPOSURE))
−1.0
−0.5
0.0
0.5
1.0
1.5
2.0
2.5
−1
0
1
2
3
36011990900
36007013600
36007013500
36109991200
36007013000
36011991300
36109991400
36109990100
36107020300
36007012600
36107020100
36011990100

Lagrange Multiplier Test

Both methods – visualisation and Moran's I – can tell us that there is spatial autocorrelation. However, both method do not provide any information on why there is autocorrelation. Possible reasons:

- Interdependence (ρ)
- Clustering on unobservables (λ)
- Spillovers in covariates ($\boldsymbol{\theta}$)

In contrast, Lagrange Multiplier tests are directed:

- (Robust) test for spatial lag dependence LM_ρ^*
- (Robust) test for spatial error dependence LM_λ^*

(Anselin et al., 1996)

Lagrange Multiplier Test

Both methods – visualisation and Moran's I – can tell us that there is spatial autocorrelation. However, both method do not provide any information on why there is autocorrelation. Possible reasons:

- Interdependence (ρ)
- Clustering on unobservables (λ)
- Spillovers in covariates ($\boldsymbol{\theta}$)

In contrast, Lagrange Multiplier tests are directed:

- (Robust) test for spatial lag dependence LM_{ρ}^{*}
- (Robust) test for spatial error dependence LM_{λ}^{*}

(Anselin et al., 1996)

Lagrange Multiplier Test

Robust test for lag dependence: H_0: $\rho = 0$

$$LM_{\rho}^{*} = G^{-1}\hat{\sigma}_{\epsilon}^{2}\Big(\frac{\hat{\boldsymbol{\epsilon}}^{\intercal}\boldsymbol{W}\boldsymbol{y}}{\hat{\sigma}_{\epsilon}^{2}} - \frac{\hat{\boldsymbol{\epsilon}}^{\intercal}\boldsymbol{W}\hat{\boldsymbol{\epsilon}}}{\hat{\sigma}_{\epsilon}^{2}}\Big)^{2} \sim \chi^{2} \tag{4}$$

where $\hat{\epsilon}$ is the residual vector of a non-spatial OLS regression.
Robust test for error dependence: H_0: $\lambda = 0$

$$LM_{\lambda}^{*} = \frac{\big(\hat{\boldsymbol{\epsilon}}^{\intercal}\boldsymbol{W}\hat{\boldsymbol{\epsilon}}/\hat{\sigma}_{\epsilon}^{2} - [T\hat{\sigma}_{\epsilon}^{2}(G + T\hat{\sigma}_{\epsilon}^{2})^{-1}]\hat{\boldsymbol{\epsilon}}^{\intercal}\boldsymbol{W}\boldsymbol{y}/\hat{\sigma}_{\epsilon}^{2}\big)^{2}}{T[1 - \frac{\hat{\sigma}_{\epsilon}^{2}}{G+\hat{\sigma}_{\epsilon}^{2}}]} \sim \chi^{2} \tag{5}$$

$$\text{where } G = (\boldsymbol{W}\boldsymbol{X}\hat{\beta})^{\intercal}(\boldsymbol{I} - \boldsymbol{X}(\boldsymbol{X}^{\intercal}\boldsymbol{X})^{-1}\boldsymbol{X}^{\intercal})(\boldsymbol{W}\boldsymbol{X}\hat{\beta})$$

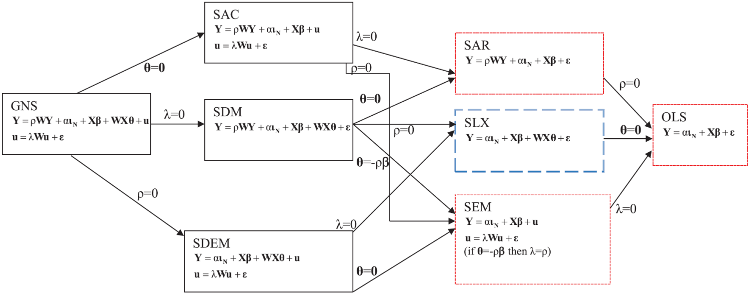
$$\text{and } T = \mathrm{tr}[(\boldsymbol{W}^{\intercal} + \boldsymbol{W})\boldsymbol{W}]$$

(Anselin et al., 1996)

Part II

Spatial Regression

Model Specifications



Note: GNS = general nesting spatial model, SAC = spatial autoregressive combined model, SDM = spatial Durbin model, SDEM = spatial Durbin error model, SAR = spatial autoregressive model, SLX = spatial lag of **X** model, SEM = spatial error model, OLS = ordinary least squares model.

(Halleck Vega and Elhorst, 2015, p.343)

Model Specifications I

Spatial Error Model (SEM) $\Rightarrow$ Clustering on Unobservables

$$\begin{aligned} \boldsymbol{y} &= \alpha\boldsymbol{\iota} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \\ \boldsymbol{u} &= \lambda\boldsymbol{W}\boldsymbol{u} + \boldsymbol{\varepsilon} \end{aligned} \tag{6}$$

Spatial Autoregressive Model (SAR) $\Rightarrow$ Interdependence

$$\boldsymbol{y} = \alpha\boldsymbol{\iota} + \rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{7}$$

Spatially lagged X Model (SLX) $\Rightarrow$ Spillovers in Covariates

$$\boldsymbol{y} = \alpha\boldsymbol{\iota} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{8}$$

(Halleck Vega and Elhorst, 2015; LeSage and Pace, 2009)

Example: air quality (ind var) ⇒ house prices (dep var)

Spatial Error Model (SEM) ⇒ Clustering on Unobservables

Spatially clustered or diffusing crime rates influence the house prices in the affected areas (but are independent of air quality and available green space).

Spatial Autoregressive Model (SAR) ⇒ Interdependence

House prices in one district directly influence the house prices in neighbouring districts (e.g. sellers or estate agents determine the prices based on the prices they observe in neighbouring districts).

Spatially lagged X Model (SLX) ⇒ Spillovers in Covariates

House prices in the focal unit are not only influenced by the air quality in the focal unit but also by the air quality in neighbouring districts.

Example: air quality (ind var) ⇒ house prices (dep var)

Spatial Error Model (SEM) ⇒ Clustering on Unobservables

Spatially clustered or diffusing crime rates influence the house prices in the affected areas (but are independent of air quality and available green space).

Spatial Autoregressive Model (SAR) ⇒ Interdependence

House prices in one district directly influence the house prices in neighbouring districts (e.g. sellers or estate agents determine the prices based on the prices they observe in neighbouring districts).

Spatially lagged X Model (SLX) ⇒ Spillovers in Covariates

House prices in the focal unit are not only influenced by the air quality in the focal unit but also by the air quality in neighbouring districts.

Example: air quality (ind var) ⇒ house prices (dep var)

Spatial Error Model (SEM) ⇒ Clustering on Unobservables

Spatially clustered or diffusing crime rates influence the house prices in the affected areas (but are independent of air quality and available green space).

Spatial Autoregressive Model (SAR) ⇒ Interdependence

House prices in one district directly influence the house prices in neighbouring districts (e.g. sellers or estate agents determine the prices based on the prices they observe in neighbouring districts).

Spatially lagged X Model (SLX) ⇒ Spillovers in Covariates

House prices in the focal unit are not only influenced by the air quality in the focal unit but also by the air quality in neighbouring districts.

Model Specifications II

Spatial Durbin Model (SDM)

$$\boldsymbol{y} = \alpha\boldsymbol{\iota} + \rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \tag{9}$$

Spatial Durbin Error Model (SDEM)

$$\begin{aligned} \boldsymbol{y} &= \alpha\boldsymbol{\iota} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{u}, \\ \boldsymbol{u} &= \lambda\boldsymbol{W}\boldsymbol{u} + \boldsymbol{\varepsilon} \end{aligned} \tag{10}$$

Combined Spatial Autocorrelation Model (SAC)

$$\begin{aligned} \boldsymbol{y} &= \alpha\boldsymbol{\iota} + \rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u}, \\ \boldsymbol{u} &= \lambda\boldsymbol{W}\boldsymbol{u} + \boldsymbol{\varepsilon} \end{aligned} \tag{11}$$

(Halleck Vega and Elhorst, 2015; LeSage and Pace, 2009)

Model Specifications: An example

Environmental Inequality
Are areas with a high minority share affected by higher amounts of industrial air pollution?

- Data: German Census and E-PRTR
- Dependent variable: air pollution in ln kg
- Independent variable: % foreigners
- Contiguity weights matrix

(Rüttenauer, 2018)

Model Specifications An example

Table: Dep Var: ln Air Pollution

	OLS	SEM	SAR	SLX	SDM	SDEM	SAC
(Intercept)	−0.000	−0.016*	−0.004**	−0.005	−0.005**	−0.014	−0.005
	(0.003)	(0.008)	(0.002)	(0.003)	(0.002)	(0.007)	(0.004)
% Foreigner	0.215***	0.033***	0.056***	0.088***	0.029***	0.063***	0.060***
	(0.003)	(0.002)	(0.002)	(0.004)	(0.002)	(0.003)	(0.002)
ρ			0.779***		0.776***		0.617***
			(0.001)		(0.001)		(0.008)
λ		0.784***				0.778***	0.361***
		(0.001)				(0.001)	(0.012)
W [% Foreigner]				0.217***	0.047***	0.122***	
				(0.005)	(0.003)	(0.004)	
R^2	0.046			0.063			
Adj. R^2	0.046			0.063			
Num. obs.	93777	93777	93777	93777	93777	93777	93777
Log Likelihood		-84672	-84214		-84069	-84300	-83373

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. W is spezified as contiguity weights matrix

Coefficient estimates ≠ 'marginal' effects

Attention: Do not interpret coefficients in SAR, SAC, and SDM!! Using the reduced form

$$\begin{aligned} \boldsymbol{y} &= \alpha\boldsymbol{\iota} + \rho\boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{y} &= (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}(\alpha\boldsymbol{\iota} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}), \end{aligned} \tag{12}$$

we can calculate the first derivative:

$$\begin{aligned} \frac{\partial \boldsymbol{y}}{\partial \boldsymbol{x}_k} &= (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\beta_k \\ &= (\boldsymbol{I} + \rho\boldsymbol{W} + \rho^2\boldsymbol{W}^2 + \rho^3\boldsymbol{W}^3 + ...)\beta_k, \end{aligned} \tag{13}$$

where $\rho\boldsymbol{W}\beta_k$ equals the effect stemming from direct neighbours, $\rho^2\boldsymbol{W}^2\beta_k$ the effect stemming from second order neighbours (neighbours of neighbours),... Note that this also includes feedback loops if unit i is also a second order neighbour of itself.

Impacts

	Direct Impacts	Spatial Spillovers
OLS/SEM	β_k	–
SAR/SAC	Diagonal elements of $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1}\beta_k$	Off-diagonal elements of $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1}\beta_k$
SLX/SDEM	β_k	θ_k
SDM	Diagonal elements of $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1}[\beta_k + \boldsymbol{W}\theta_k]$	Off-diagonal elements of $(\boldsymbol{I} - \rho \boldsymbol{W})^{-1}[\beta_k + \boldsymbol{W}\theta_k]$

Note (example SAR): As the reduced form of equation (7) can be written as $\boldsymbol{y} = (\boldsymbol{I} - \rho \boldsymbol{W})^{-1}(\alpha \boldsymbol{\iota} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon})$, the partial derivatives are given by $\frac{\partial y}{\partial x_k} = (\boldsymbol{I} - \rho \boldsymbol{W})^{-1}\beta_k$.

Different kinds of spillover effects:

- Global spillover effects: SAR, SAC, SDM
- Local spillover effects: SLX, SDEM

(Halleck Vega and Elhorst, 2015)

Impacts: Interpretation

Global spillover effects (SAR, SAC, SDM)

- Includes direct neighbours but also neighbours of neighbours (second order neighbours) and further higher-order neighbours
- Diffusion process: house prices in one region influences house prices in neighbouring units, which influences the neighbours' neighbours ...
- Includes feedback effect: focal unit is second order neighbour

Local spillover effects (SLX, SDEM)

- Only direct neighbours (as specified by $\boldsymbol{W}$)
- Effect of a one unit change of $\boldsymbol{x}_k$ in the spatially weighted neighbouring observations on the dependent variable of the focal unit
- Using a row-normalised contiguity weights matrix, $\boldsymbol{W}\boldsymbol{x}_k$ is the average value of $\boldsymbol{x}_k$ in the neighbouring units

Impacts: Interpretation

Global spillover effects (SAR, SAC, SDM)

- Includes direct neighbours but also neighbours of neighbours (second order neighbours) and further higher-order neighbours
- Diffusion process: house prices in one region influences house prices in neighbouring units, which influences the neighbours' neighbours ...
- Includes feedback effect: focal unit is second order neighbour

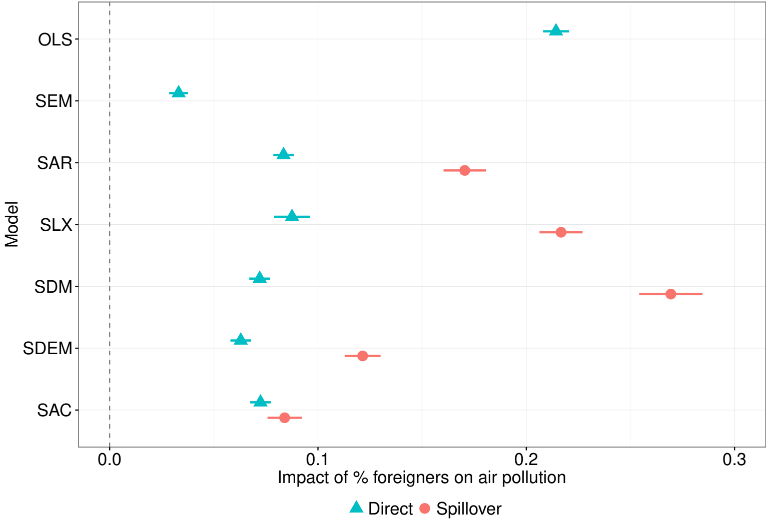
Local spillover effects (SLX, SDEM)

- Only direct neighbours (as specified by $\boldsymbol{W}$)
- Effect of a one unit change of $\boldsymbol{x}_k$ in the spatially weighted neighbouring observations on the dependent variable of the focal unit
- Using a row-normalised contiguity weights matrix, $\boldsymbol{W}\boldsymbol{x}_k$ is the average value of $\boldsymbol{x}_k$ in the neighbouring units

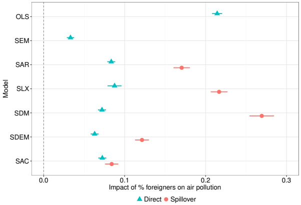
Impacts in our example

Model	Direct		Spillover	
	Coef.	SE	Coef.	SE
OLS	0.214***	(0.003)		
SEM	0.033***	(0.002)		
SAR	0.083***	(0.003)	0.170***	(0.005)
SLX	0.088***	(0.004)	0.217***	(0.005)
SAC	0.072***	(0.003)	0.084***	(0.004)
SDM	0.072***	(0.003)	0.269***	(0.008)
SDEM	0.063***	(0.003)	0.121***	(0.004)

Average Impacts in our example



Interpretation



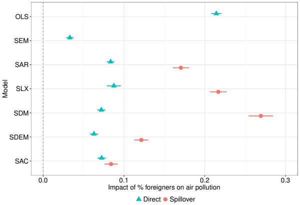
In OLS and SEM:

- straight forward

In SLX and SDEM:

- Direct: on average, a one unit increase in % foreigners in area *i* is correlated with a 0.088 / 0.063 unit higher pollution in area *i*.
- Indirect: on average, a one unit increase in % foreigners in area *i*'s direct neighbouring areas (in case of row-normalization: average change in neighbours) is correlated with a 0.217 / 0.121 unit higher pollution in area *i*.

Interpretation



In SAR, SAC, and SDM:

- Direct: on average, a one unit increase in % foreigners in area i is correlated with a 0.083 / 0.072 / 0.072 unit higher pollution in area i, **including feedback effects**.
- Indirect: would the % foreigners increase by one unit in all other areas $j \neq i$, the pollution in area i would increase by 0.170 / 0.084 / 0.269 units.

Sources of bias

Bias in non-spatial OLS:

$$\begin{aligned}\text{plim}\ \hat{\beta} =& \frac{\sum_{ij}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal \circ \boldsymbol{M}(\rho))_{ij}}{\mathrm{tr}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal)}\beta \\ &+ \frac{\sum_{ij}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal \circ \boldsymbol{M}(\rho)\boldsymbol{W})_{ij}}{\mathrm{tr}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal)}\theta \\ &+ \frac{\sum_{ij}(\boldsymbol{M}(\delta)\boldsymbol{M}(\delta)^\intercal \circ \boldsymbol{M}(\rho)\boldsymbol{L}(\lambda))}{\mathrm{tr}(\boldsymbol{M}(\delta)^\intercal\boldsymbol{M}(\delta))}\gamma,\end{aligned} \tag{14}$$

where $\circ$ denotes the Hadamard product, $\boldsymbol{M}(\delta) = (\boldsymbol{I}_N - \delta\boldsymbol{W})^{-1}$, $\boldsymbol{M}(\rho) = (\boldsymbol{I}_N - \rho\boldsymbol{W})^{-1}$, and $\boldsymbol{L}(\lambda) = (\boldsymbol{I}_N - \lambda\boldsymbol{W})^{-1}$. γ denotes the strength of a non-spatial omitted variable bias.

Sources of bias

Spatial Autocorrelation:

$$cov(\boldsymbol{y}, \boldsymbol{Wy}) \neq 0$$

Possible Reasons:

- Clustering on Observables: $\boldsymbol{y} = f(\boldsymbol{X})$ and $cov(\boldsymbol{X}, \boldsymbol{WX}) \neq 0$
- Spillovers in Covariates: $\boldsymbol{y} = f(\boldsymbol{X}, \boldsymbol{WX})$
- Clustering on Unobservables: $\boldsymbol{y} = f(\boldsymbol{X}, \varepsilon)$ and $cov(\varepsilon, \boldsymbol{W}\varepsilon) \neq 0$
- Interdependence: $\boldsymbol{y} = f(\boldsymbol{Wy})$

where $\boldsymbol{W}$ specifies the $N \times N$ spatial weights matrix (all $w_{ij} > 0$ for neighbouring i and $j \neq i$, and $w_{ij} = 0$ otherwise).
(Cook et al., 2015; Pace and LeSage, 2010)

Sources of bias

Spatial Autocorrelation:

$$cov(\boldsymbol{y}, \boldsymbol{W}\boldsymbol{y}) \neq 0$$

Possible Reasons:

- Clustering on Observables: $\boldsymbol{y} = f(\boldsymbol{X})$ and $cov(\boldsymbol{X}, \boldsymbol{W}\boldsymbol{X}) \neq 0$
- Spillovers in Covariates: $\boldsymbol{y} = f(\boldsymbol{X}, \boldsymbol{W}\boldsymbol{X})$
- Clustering on Unobservables: $\boldsymbol{y} = f(\boldsymbol{X}, \varepsilon)$ and $cov(\varepsilon, \boldsymbol{W}\varepsilon) \neq 0$
- Interdependence: $\boldsymbol{y} = f(\boldsymbol{W}\boldsymbol{y})$

where $\boldsymbol{W}$ specifies the $N \times N$ spatial weights matrix (all $w_{ij} > 0$ for neighbouring i and $j \neq i$, and $w_{ij} = 0$ otherwise).
(Cook et al., 2015; Pace and LeSage, 2010)

Sources of bias

Spatial Autocorrelation:

$$cov(\boldsymbol{y}, \boldsymbol{Wy}) \neq 0$$

Possible Reasons:

- Clustering on Observables: $\boldsymbol{y} = f(\boldsymbol{X})$ and $cov(\boldsymbol{X}, \boldsymbol{WX}) \neq 0$
- Spillovers in Covariates: $\boldsymbol{y} = f(\boldsymbol{X}, \boldsymbol{WX})$
- Clustering on Unobservables: $\boldsymbol{y} = f(\boldsymbol{X}, \varepsilon)$ and $cov(\varepsilon, \boldsymbol{W}\varepsilon) \neq 0$
- Interdependence: $\boldsymbol{y} = f(\boldsymbol{Wy})$

Biased estimates of impacts in non-spatial $\hat{\beta}_{OLS}$

(Cook et al., 2015; Pace and LeSage, 2010)

Sources of bias

Spatial Autocorrelation:

$$cov(\boldsymbol{y}, \boldsymbol{Wy}) \neq 0$$

Possible Reasons:

- Clustering on Observables: $\boldsymbol{y} = f(\boldsymbol{X})$ and $cov(\boldsymbol{X}, \boldsymbol{WX}) \neq 0$
- Spillovers in Covariates: $\boldsymbol{y} = f(\boldsymbol{X}, \boldsymbol{WX})$
- Clustering on Unobservables: $\boldsymbol{y} = f(\boldsymbol{X}, \varepsilon)$ and $cov(\varepsilon, \boldsymbol{W}\varepsilon) \neq 0$
- Interdependence: $\boldsymbol{y} = f(\boldsymbol{Wy})$

Biased estimates of impacts in non-spatial $\hat{\beta}_{OLS}$

(Cook et al., 2015; Pace and LeSage, 2010)

Sources of bias

Spatial Autocorrelation:

$$cov(\boldsymbol{y}, \boldsymbol{Wy}) \neq 0$$

Possible Reasons:

- Clustering on Observables: $\boldsymbol{y} = f(\boldsymbol{X})$ and $cov(\boldsymbol{X}, \boldsymbol{WX}) \neq 0$
- Spillovers in Covariates: $\boldsymbol{y} = f(\boldsymbol{X}, \boldsymbol{WX})$
- Clustering on Unobservables: $\boldsymbol{y} = f(\boldsymbol{X}, \varepsilon)$ and $cov(\varepsilon, \boldsymbol{W}\varepsilon) \neq 0$
- Interdependence: $\boldsymbol{y} = f(\boldsymbol{Wy})$

Amplified omv bias in non-spatial $\hat{\beta}_{OLS}$ if $cov(\boldsymbol{X}, \varepsilon) \neq 0$

(Cook et al., 2015; Pace and LeSage, 2010)

Model Selection

Which model to select?

- Theoretical considerations
- Empirical tests

LeSage and Pace (2009):

- Unbiased estimates only by SDM
- Subsumes SAC and SEM

Cook et al. (2015):

- Interested in non-spatial parameters: Use SDM
- Interested in spatial parameters: Use SAC or SDEM

Elhorst (2014):

- Acceptable results only by SDM and SDEM

Model Selection

Which model to select?

- Theoretical considerations
- Empirical tests

LeSage and Pace (2009):

- Unbiased estimates only by SDM
- Subsumes SAC and SEM

Cook et al. (2015):

- Interested in non-spatial parameters: Use SDM
- Interested in spatial parameters: Use SAC or SDEM

Elhorst (2014):

- Acceptable results only by SDM and SDEM

Model Selection

Which model to select?

- Theoretical considerations
- Empirical tests

LeSage and Pace (2009):

- Unbiased estimates only by SDM
- Subsumes SAC and SEM

Cook et al. (2015):

- Interested in non-spatial parameters: Use SDM
- Interested in spatial parameters: Use SAC or SDEM

Elhorst (2014):

- Acceptable results only by SDM and SDEM

Model Selection

Which model to select?

- Theoretical considerations
- Empirical tests

LeSage and Pace (2009):

- Unbiased estimates only by SDM
- Subsumes SAC and SEM

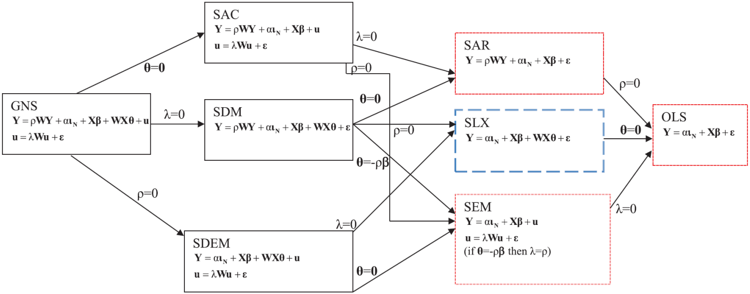
Cook et al. (2015):

- Interested in non-spatial parameters: Use SDM
- Interested in spatial parameters: Use SAC or SDEM

Elhorst (2014):

- Acceptable results only by SDM and SDEM

Empirical test I: General to specific



Note: GNS = general nesting spatial model, SAC = spatial autoregressive combined model, SDM = spatial Durbin model, SDEM = spatial Durbin error model, SAR = spatial autoregressive model, SLX = spatial lag of **X** model, SEM = spatial error model, OLS = ordinary least squares model.

(Halleck Vega and Elhorst, 2015, p.343)

$\Rightarrow$ Where to start if GNS is not identifiable?

Empirical test II: Specific to general (LM-test)

Rejection rates of H_0

ρ	δ	λ	θ	LM_λ	LM_ρ	LM^*_λ	LM^*_ρ	$LM_{\lambda\rho}$
0.0	0, 0	0.0	0, 0	0.0410	0.0440	0.0520	0.0500	0.0580
0.4	0, 0	0.0	0, 0	1.0000	1.0000	0.0920	0.7750	1.0000
0.8	0, 0	0.0	0, 0	1.0000	1.0000	0.3880	0.9990	1.0000
0.0	0.4, 0.7	0.0	0, 0	0.0440	0.0520	0.0550	0.0590	0.0550
0.4	0.4, 0.7	0.0	0, 0	1.0000	1.0000	0.1340	0.9900	1.0000
0.8	0.4, 0.7	0.0	0, 0	1.0000	1.0000	0.9950	1.0000	1.0000
0.0	0, 0	0.4	0, 0	1.0000	1.0000	0.7190	0.1140	1.0000
0.4	0, 0	0.4	0, 0	1.0000	1.0000	0.8560	0.6830	1.0000
0.8	0, 0	0.4	0, 0	1.0000	1.0000	0.8380	0.9780	1.0000
0.0	0.4, 0.7	0.4	0, 0	1.0000	1.0000	0.9430	0.1100	1.0000
0.4	0.4, 0.7	0.4	0, 0	1.0000	1.0000	0.9970	0.9560	1.0000
0.8	0.4, 0.7	0.4	0, 0	1.0000	1.0000	1.0000	1.0000	1.0000
0.0	0, 0	0.8	0, 0	1.0000	1.0000	0.9990	0.2860	1.0000
0.4	0, 0	0.8	0, 0	1.0000	1.0000	0.9960	0.5190	1.0000
0.8	0, 0	0.8	0, 0	1.0000	1.0000	0.7450	0.7410	1.0000
0.0	0.4, 0.7	0.8	0, 0	1.0000	1.0000	1.0000	0.3000	1.0000
0.4	0.4, 0.7	0.8	0, 0	1.0000	1.0000	1.0000	0.7280	1.0000
0.8	0.4, 0.7	0.8	0, 0	1.0000	1.0000	1.0000	0.9410	1.0000
0.0	0, 0	0.0	0.1, 0.8	0.4240	0.9910	1.0000	1.0000	1.0000
0.4	0, 0	0.0	0.1, 0.8	1.0000	1.0000	1.0000	1.0000	1.0000
0.8	0, 0	0.0	0.1, 0.8	1.0000	1.0000	1.0000	1.0000	1.0000
0.0	0.4, 0.7	0.0	0.1, 0.8	0.8320	1.0000	1.0000	1.0000	1.0000
0.4	0.4, 0.7	0.0	0.1, 0.8	1.0000	1.0000	0.9960	1.0000	1.0000
0.8	0.4, 0.7	0.0	0.1, 0.8	1.0000	1.0000	0.4970	1.0000	1.0000
0.0	0, 0	0.4	0.1, 0.8	1.0000	1.0000	0.9900	1.0000	1.0000
0.4	0, 0	0.4	0.1, 0.8	1.0000	1.0000	0.9530	1.0000	1.0000
0.8	0, 0	0.4	0.1, 0.8	1.0000	1.0000	0.8520	1.0000	1.0000
0.0	0.4, 0.7	0.4	0.1, 0.8	1.0000	1.0000	0.9560	1.0000	1.0000
0.4	0.4, 0.7	0.4	0.1, 0.8	1.0000	1.0000	0.0930	1.0000	1.0000
0.8	0.4, 0.7	0.4	0.1, 0.8	1.0000	1.0000	0.9600	1.0000	1.0000
0.0	0, 0	0.8	0.1, 0.8	1.0000	1.0000	0.2590	0.9930	1.0000
0.4	0, 0	0.8	0.1, 0.8	1.0000	1.0000	0.2370	0.9940	1.0000
0.8	0, 0	0.8	0.1, 0.8	1.0000	1.0000	0.1530	0.9940	1.0000
0.0	0.4, 0.7	0.8	0.1, 0.8	1.0000	1.0000	0.9880	1.0000	1.0000
0.4	0.4, 0.7	0.8	0.1, 0.8	1.0000	1.0000	1.0000	1.0000	1.0000
0.8	0.4, 0.7	0.8	0.1, 0.8	1.0000	1.0000	1.0000	1.0000	1.0000

Monte Carlo simulations

Data generating process (DGP):

$$\boldsymbol{y} = \rho \boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{u}, \tag{15}$$

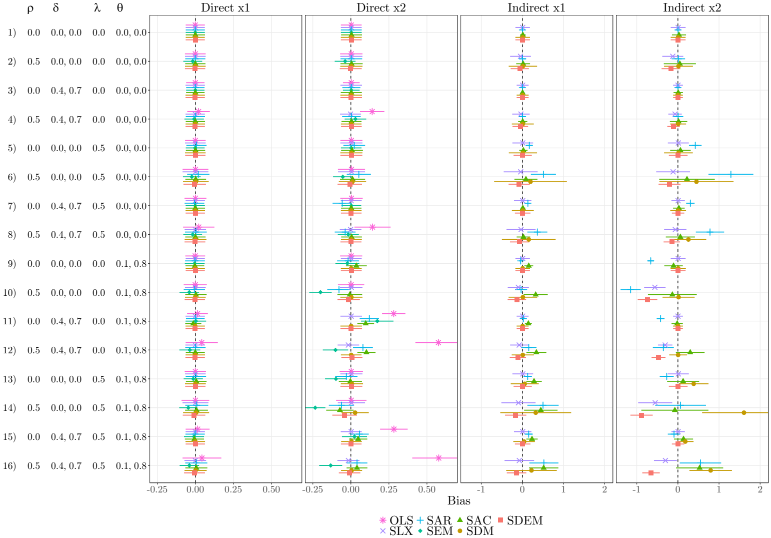
$$\boldsymbol{u} = \lambda \boldsymbol{W}\boldsymbol{u} + \boldsymbol{X}\gamma + \varepsilon, \tag{16}$$

$$\boldsymbol{x}_k = \delta_k \boldsymbol{W}\boldsymbol{x}_k + \boldsymbol{\upsilon}_k \tag{17}$$

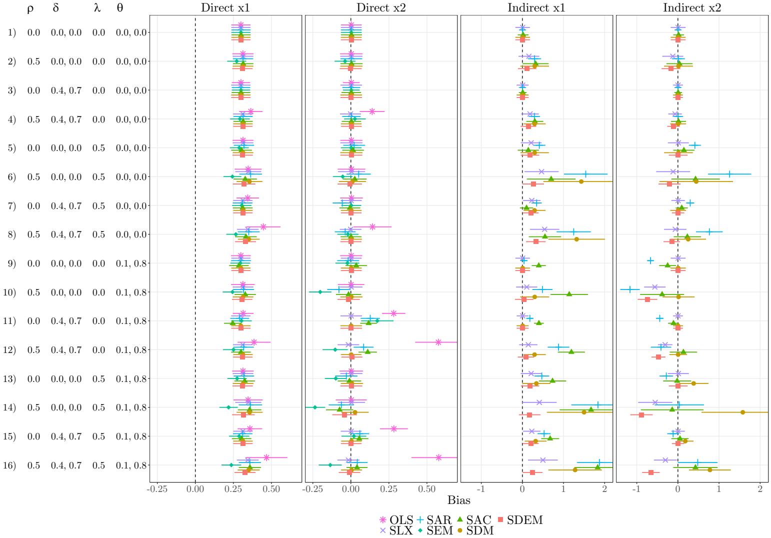
- $\boldsymbol{\upsilon}_k$ and ε are independent and randomly distributed $\mathcal{N}(0, \sigma^2_\upsilon)$ and $\mathcal{N}(0, \sigma^2_\varepsilon)$ with a mean of zero
- $\boldsymbol{x}_k$ is the kth column-vector of $\boldsymbol{X}$ for $k = 1, ..., K$ covariates
- ρ represents the autocorrelation in the dependent variable, λ the autocorrelation in the disturbances, δ_k the autocorrelation in covariate k, and γ an omitted variable bias
- $N = 900$, $\boldsymbol{W}$: row-normalized contiguity weights matrix, $R = 1000$ trials

(Rüttenauer, 2019)

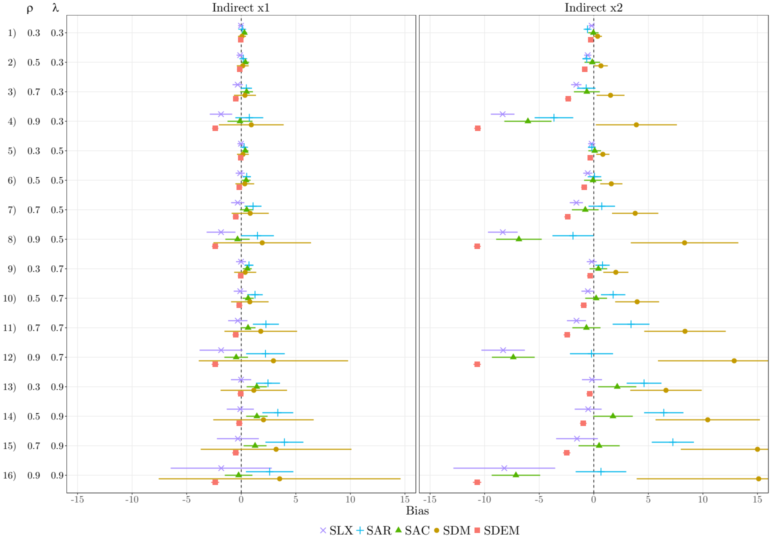
Simulation I without omv



Simulation I with omv (x_1)



Simulation II



Estimation methods

Spatial OLS (include $\boldsymbol{Wy}$ in OLS)

- Induces simultaneity bias (as $\boldsymbol{Wy}$ is endogenous)
- Bias can be large for large ρ

Consistent estimation methods for linear models

- Spatial maximum likelihood
- Spatial-IV / 2SLS / GMM: instrumenting $\boldsymbol{Wy}$ by $(\boldsymbol{X}, \boldsymbol{WX}, \boldsymbol{W}^2\boldsymbol{X}, ..., \boldsymbol{W}^q\boldsymbol{X})$

Note that the SLX model does not contain an exogenous autoregressive term. Thus, the SLX can be estimated using OLS!

(e.g. Drukker et al., 2013; Franzese and Hays, 2007; Lee, 2004)

Estimation methods

Spatial OLS (include $\boldsymbol{Wy}$ in OLS)

- Induces simultaneity bias (as $\boldsymbol{Wy}$ is endogenous)
- Bias can be large for large ρ

Consistent estimation methods for linear models

- Spatial maximum likelihood
- Spatial-IV / 2SLS / GMM: instrumenting $\boldsymbol{Wy}$ by $(\boldsymbol{X}, \boldsymbol{WX}, \boldsymbol{W}^2\boldsymbol{X}, ..., \boldsymbol{W}^q\boldsymbol{X})$

Note that the SLX model does not contain an exogenous autoregressive term. Thus, the SLX can be estimated using OLS!
(e.g. Drukker et al., 2013; Franzese and Hays, 2007; Lee, 2004)

A note on missings

Conventional missings (NAs)

- In autoregressive models, observations with missings are not allowed
- Intuition: Missing in $\boldsymbol{x}_i$ passes to $\boldsymbol{y}_i$, passes to $\boldsymbol{y}_j$, passes to $\boldsymbol{y}_k$, ... passes thought the hole system of neighbours. In the end, we have dataset of missings.

Empty neighbours sets

- Observations without neighbours are a problem
- What is the correct value for lagged variables? Maybe zero? Or NA?
- Mainly problematic with contiguity weights
- Drop 'island' or impute w_{ij} with nearest neighbour

A note on missings

Conventional missings (NAs)

- In autoregressive models, observations with missings are not allowed
- Intuition: Missing in $\boldsymbol{x}_i$ passes to $\boldsymbol{y}_i$, passes to $\boldsymbol{y}_j$, passes to $\boldsymbol{y}_k$, ... passes thought the hole system of neighbours. In the end, we have dataset of missings.

Empty neighbours sets

- Observations without neighbours are a problem
- What is the correct value for lagged variables? Maybe zero? Or NA?
- Mainly problematic with contiguity weights
- Drop 'island' or impute w_{ij} with nearest neighbour

Summary

Using spatial regression models can

- prevent biased estimates (depending on constellation!)
- lead to additional information
- detect spatial patterns

How to select the correct specification?

- Best way: let theory decide!
- Specification tests are of little help
- SAR, SAC (most common models!) perform badly in many situations
- Simplest model (SLX) seems quite robust against misspecification

Be cautious: No matter which model you use and which kind of spatial spillover effects you get, it is only correlation. In a cross-sectional observational study, we cannot empirically distinguish the processes leading to spatial correlation!

Summary

Using spatial regression models can

- prevent biased estimates (depending on constellation!)
- lead to additional information
- detect spatial patterns

How to select the correct specification?

- Best way: let theory decide!
- Specification tests are of little help
- SAR, SAC (most common models!) perform badly in many situations
- Simplest model (SLX) seems quite robust against misspecification

Be cautious: No matter which model you use and which kind of spatial spillover effects you get, it is only correlation. In a cross-sectional observational study, we cannot empirically distinguish the processes leading to spatial correlation!

Summary

Using spatial regression models can

- prevent biased estimates (depending on constellation!)
- lead to additional information
- detect spatial patterns

How to select the correct specification?

- Best way: let theory decide!
- Specification tests are of little help
- SAR, SAC (most common models!) perform badly in many situations
- Simplest model (SLX) seems quite robust against misspecification

Be cautious: No matter which model you use and which kind of spatial spillover effects you get, it is only correlation. In a cross-sectional observational study, we cannot empirically distinguish the processes leading to spatial correlation!

Useful literature

- Bivand, R. S., Pebesma, E., and Gómez-Rubio, V. (2013), *Applied Spatial Data Analysis with R*. Springer, New York.
- Elhorst, J. P. (2014), *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*, SpringerBriefs in Regional Science. Springer, Berlin and Heidelberg.
- Fotheringham, A. S. and Rogerson, P., editors (2009), *The Sage Handbook of Spatial Analysis*. Sage, Los Angeles and London.
- Halleck Vega, S. and Elhorst, J. P. (2015), The SLX model, *Journal of Regional Science*, 55(3):339–363.
- LeSage, J. P. and Pace, R. K. (2009), *Introduction to Spatial Econometrics*, Statistics, Textbooks and Monographs. CRC Press, Boca Raton.
- Ward, M. D. and Gleditsch, K. S. (2008), *Spatial Regression Models*, volume 155 of *Quantitative Applications in the Social Sciences*. Sage, Thousand Oaks.

Part III

Non-Linear Models

Overview

Models with endogenous regressors (SAR)

- In the literature: mostly spatial probit considered
- Spatial logit rather uncommon (non normally distributed errors)

Issues with non-linear spatial models

- Estimation: with dependent observations, we need to maximize one n-dimensional (log-)likelihood instead of a product of n independent distributions
- Estimation challenging and computationally intense
- Hard to interpret due to non-linear effects in non-linear models

(Elhorst et al., 2017; Franzese et al., 2016)

Problem

Spatial-SAR-Probit

$$\boldsymbol{y}^\star = \rho \boldsymbol{W} \boldsymbol{y}^\star + \boldsymbol{X}\beta + \varepsilon \qquad (18)$$
$$y_i = \{1 \text{ if } y_i^\star > 0; 0 \text{ if } y_i^\star \leq 0\}$$

or in reduced form:

$$\boldsymbol{y}^\star = (\boldsymbol{I} - \rho \boldsymbol{W})^{-1} \boldsymbol{X}\beta + \boldsymbol{u},\ \boldsymbol{u} = (\boldsymbol{I} - \rho \boldsymbol{W})^{-1}\varepsilon, \qquad (19)$$
$$\text{with } \boldsymbol{u} \sim MVN(0, (\boldsymbol{I} - \rho \boldsymbol{W})^\intercal (\boldsymbol{I} - \rho \boldsymbol{W})^{-1})$$

- Probability $\boldsymbol{y}^\star$ is a latent variable, not observed
- We only observe binary outcome y_i
- $\mathrm{Cov}(y_i, y_j)$ is not the same as $\mathrm{Cov}(y_i^\star, y_j^\star)$
- Error term is heteroskedastic and spatially correlated

Problem

Probability

$$\begin{aligned}\mathrm{Prob}[\boldsymbol{y}^\star > 0] &= \mathrm{Prob}[(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{X}\beta + (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\varepsilon] \\ &= \mathrm{Prob}[(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\varepsilon < (\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{X}\beta]\end{aligned} \tag{20}$$

or in using the observed outcome:

$$\begin{aligned}\mathrm{Prob}[\boldsymbol{y}_i = 1|\boldsymbol{X}] &= \mathrm{Prob}\left[u_i < [(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{X}\beta]_i\right] \\ &= \phi\{[(\boldsymbol{I} - \rho\boldsymbol{W})^{-1}\boldsymbol{X}\beta]_i/\sigma_{ui}\}\end{aligned} \tag{21}$$

- $\phi\{\}$ is an n-dimensional cumulative-normal distribution
- σ_{ui} equals $(\boldsymbol{I} - \rho\boldsymbol{W})^\intercal(\boldsymbol{I} - \rho\boldsymbol{W})^{-1})_{ii}$, not constant
- no analytical solution

Estimation

Estimation methods for Spatial-SAR Probit / Logit

- Expectation Maximization (McMillen, 1992)
- (Linearized) Generalized Methods of Moments (Klier and McMillen, 2008)
- Recursive Importance Sampling (Beron and Vijverberg, 2004)
- Maximum Simulated Likelihood RIS (Franzese et al., 2016)
- Bayesian approach with Markov Chain Monte Carlo simulations (LeSage and Pace, 2009): R package ‘**spatialprobit**’

Note that it can be hard to interpret the results. As in the linear case, it is necessary to compute the impacts. However, the ‘marginal’ effects may vary with values of the independent variables and the location (Lacombe and LeSage, 2018).

My personal suggestion

In case you are not familiar with the econometric estimation methods and spatial regression models, don't use non-linear models with AR term.
If necessary, I would recommend using '**spatialprobit**' relying on Bayesian MCMC (set high ndraw and burn-in, e.g. 7500 and 2500).

- So far, no 'best practice' guide
- No systematic comparison of estimation methods
- In R: Only '**spatialprobit**' provides impact measures?
- Hard to interpret results

Work-around:
If the specification is theoretical plausible, using SLX probit / logit might be a practical solution!

References I

Anselin, L., Bera, A. K., Florax, R., and Yoon, M. J. (1996), Simple diagnostic tests for spatial dependence, *Regional Science and Urban Economics*, 26(1):77–104.

Beron, K. J. and Vijverberg, W. P. M. (2004), Probit in a spatial context: A Monte Carlo analysis, In Anselin, L., Florax, Raymond J. G. M, and Rey, S. J., editors, *Advances in Spatial Econometrics: Methodology, Tools and Applications*, pages 169–195. Springer Berlin Heidelberg, Berlin, Heidelberg.

Bivand, R. (2018). Creating neighbours.

Bivand, R. S., Pebesma, E., and Gómez-Rubio, V. (2013), *Applied Spatial Data Analysis with R*. Springer, New York.

Cook, S. J., Hays, J. C., and Franzese, R. J. (2015). Model Specification and Spatial Interdependence: Paper prepared for the 2015 Summer Methods Meeting.

Downey, L. (2006), Using geographic information systems to reconceptualize spatial relationships and ecological context, *American Journal of Sociology*, 112(2):567–612.

Drukker, D. M., Egger, P., and Prucha, I. R. (2013), On two-step estimation of a spatial autoregressive model with autoregressive disturbances and endogenous regressors, *Econometric Reviews*, 32(5-6):686–733.

Elhorst, J. P. (2014), *Spatial Econometrics: From Cross-Sectional Data to Spatial Panels*, SpringerBriefs in Regional Science. Springer, Berlin and Heidelberg.

Elhorst, J. P., Heijnen, P., Samarina, A., and Jacobs, J. P. A. M. (2017), Transitions at different moments in time: A spatial probit approach, *Journal of Applied Econometrics*, 32(2):422–439.

Fotheringham, A. S. and Rogerson, P., editors (2009), *The Sage Handbook of Spatial Analysis*. Sage, Los Angeles and London.

References II

Franzese, R. J. and Hays, J. C. (2007), Spatial econometric models of cross-sectional interdependence in political science panel and time-series-cross-section data, *Political Analysis*, 15(2):140–164.

Franzese, R. J., Hays, J. C., and Cook, S. J. (2016), Spatial- and spatiotemporal-autoregressive probit models of interdependent binary outcomes, *Political Science Research and Methods*, 4(01):151–173.

Halleck Vega, S. and Elhorst, J. P. (2015), The SLX model, *Journal of Regional Science*, 55(3):339–363.

Keuschnigg, M., Mutgan, S., and Hedström, P. (2019), Urban Scaling and the Regional Divide, *Science advances*, 5(1):eaav0042.

Klier, T. and McMillen, D. P. (2008), Clustering of auto supplier plants in the United States: Generalized Method of Moments spatial logit for large samples, *Journal of Business & Economic Statistics*, 26(4):460–471.

Lacombe, D. J. and LeSage, J. P. (2018), Use and interpretation of spatial autoregressive probit models, *The Annals of Regional Science*, 60(1):1–24.

Lee, L.-f. (2004), Asymptotic distributions of quasi-maximum likelihood estimators for spatial autoregressive models, *Econometrica*, 72(6):1899–1925.

LeSage, J. P. and Pace, R. K. (2009), *Introduction to Spatial Econometrics*, Statistics, Textbooks and Monographs. CRC Press, Boca Raton.

LeSage, J. P. and Pace, R. K. (2014), The biggest myth in spatial econometrics, *Econometrics*, 2(4):217–249.

Logan, J. R. (2012), Making a place for space: Spatial thinking in social science, *Annual Review of Sociology*, 38:507–524.

McMillen, D. P. (1992), Probit with spatial autocorrelation, *Journal of Regional Science*, 32(3):335–348.

Neumayer, E. and Plümper, T. (2016), W, *Political Science Research and Methods*, 4(01):175–193.

References III

Pace, R. K. and LeSage, J. P. (2010), Omitted variable biases of OLS and spatial lag models, In Páez, A., Gallo, J., Buliung, R. N., and Dall'erba, S., editors, *Progress in Spatial Analysis*, pages 17–28. Springer, Berlin and Heidelberg.

Rüttenauer, T. (2018), Neighbours matter: A nation-wide small-area assessment of environmental inequality in Germany, *Social Science Research*, 70:198–211.

Rüttenauer, T. (2019), Spatial regression models: A systematic comparison of different model specifications using Monte Carlo experiments, *Sociological Methods and Research*, Forthcoming.

Schulte-Cloos, J. and Rüttenauer, T. (2018), A transformation from within? Dynamics of party activists and the rise of the German AfD, *SSRN Electronic Journal*.

Tobler, W. R. (1970), A computer movie simulating urban growth in the Detroit region, *Economic Geography*, 46:234–240.

Ward, M. D. and Gleditsch, K. S. (2008), *Spatial Regression Models*, volume 155 of *Quantitative Applications in the Social Sciences*. Sage, Thousand Oaks.

Weigand, M., Wurm, M., Dech, S., and Taubenböck, H. (2019), Remote sensing in environmental justice research—A review, *ISPRS International Journal of Geo-Information*, 8(1):20.