9) Spatial Autoregressive (SAR), Spatial Error Model (SEM), Spatial Durbin Model (SDM)

Vitor Kamada

August 2019

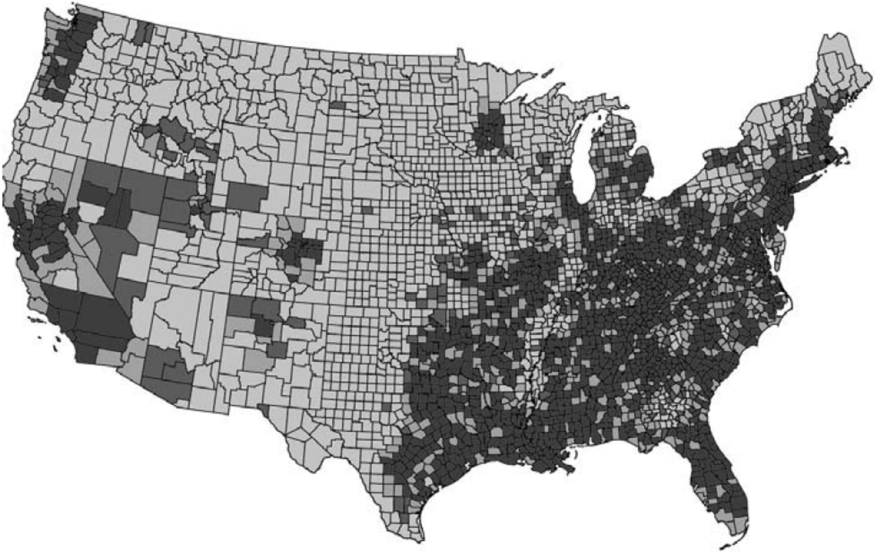
Reference

Tables, Graphics, and Figures from:

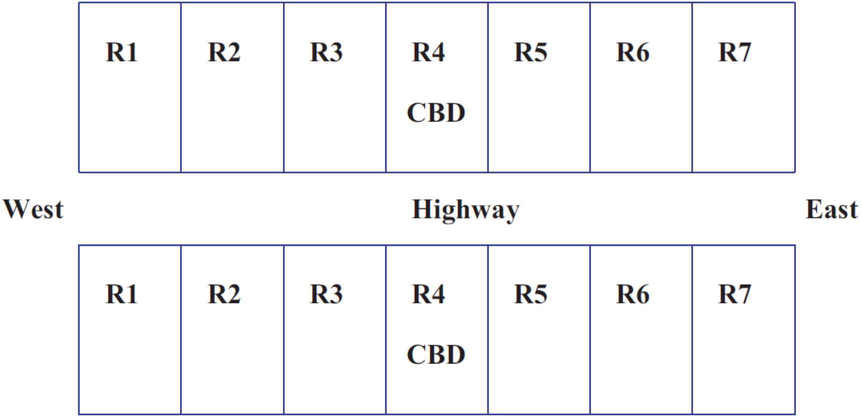
1) LeSage (2008). **An Introduction to Spatial Econometrics**, in Revue d'économie industrielle

2) LeSage and Pace (2009). **Introduction to Spatial Econometrics**: Chapters 1, 2, and 3

Moran Plot Map of County-Level Commuting Times



Regions East and West of the Central Business District (CBD)



Travel Times to the CBD (in minutes), Population Densit, and Distance (in miles)

$$y = \begin{pmatrix} \text{Travel times} \\ 42 \\ 37 \\ 30 \\ 26 \\ 30 \\ 37 \\ 42 \end{pmatrix} \qquad X = \begin{pmatrix} \text{Density} & \text{Distance} \\ 10 & 30 \\ 20 & 20 \\ 30 & 10 \\ 50 & 0 \\ 30 & 10 \\ 20 & 20 \\ 10 & 30 \end{pmatrix} \begin{array}{ll} & \\ \text{ex-urban areas} & R1 \\ \text{far suburbs} & R2 \\ \text{near suburbs} & R3 \\ \text{CBD} & R4 \\ \text{near suburbs} & R5 \\ \text{far suburbs} & R6 \\ \text{ex-urban areas} & R7 \end{array}$$

Spatial Independence vs Dependence

$$y_i = X_i\beta + \epsilon_i$$

$$\epsilon_i \sim N(0, \sigma^2)$$

$$E(\epsilon_i \epsilon_j) = E(\epsilon_i)E(\epsilon_j) = 0$$

$$y_i = \alpha_i y_j + X_i\beta + \epsilon_i$$

$$y_j = \alpha_j y_i + X_j\beta + \epsilon_j$$

$$\epsilon_i \sim N(0, \sigma^2)$$

$$\epsilon_j \sim N(0, \sigma^2)$$

Spatial Weight Matrix (W)

$$C = \begin{array}{c c} & \begin{array}{ccccccc} R1 & R2 & R3 & R4 & R5 & R6 & R7 \end{array} \\ \begin{array}{c} R1 \\ R2 \\ R3 \\ R3 \\ R5 \\ R6 \\ R7 \end{array} & \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \end{array}$$

$$W = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Spatial Lag Matrix

$$Wy = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \\ y_7 \end{pmatrix} = \begin{pmatrix} y_2 \\ (y_1 + y_3)/2 \\ (y_2 + y_4)/2 \\ (y_3 + y_5)/2 \\ (y_4 + y_6)/2 \\ (y_5 + y_7)/2 \\ y_6 \end{pmatrix}$$

Second-Order Neighbors

$$W^2 = \begin{pmatrix} 0.50 & 0 & 0.50 & 0 & 0 & 0 & 0 \\ 0 & 0.75 & 0 & 0.25 & 0 & 0 & 0 \\ 0.25 & 0 & 0.50 & 0 & 0.25 & 0 & 0 \\ 0 & 0.25 & 0 & 0.50 & 0 & 0.25 & 0 \\ 0 & 0 & 0.25 & 0 & 0.50 & 0 & 0.25 \\ 0 & 0 & 0 & 0.25 & 0 & 0.75 & 0 \\ 0 & 0 & 0 & 0 & 0.50 & 0 & 0.50 \end{pmatrix}$$

Spatial Spillovers from Changes in Region R2 Population Density

$$\tilde{X} = \begin{pmatrix} 10 & 30 \\ 20 & \mathbf{40} \\ 30 & 10 \\ 50 & 0 \\ 30 & 10 \\ 20 & 20 \\ 10 & 30 \end{pmatrix}$$

Regions / Scenario	$\hat{y}^{(1)}$	$\hat{y}^{(2)}$	$\hat{y}^{(2)} - \hat{y}^{(1)}$
$R1$:	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

Non-spatial Predictions for Changes in Region R2 Population Density

$$\tilde{X} = \begin{pmatrix} 10 & 30 \\ 20 & \mathbf{40} \\ 30 & 10 \\ 50 & 0 \\ 30 & 10 \\ 20 & 20 \\ 10 & 30 \end{pmatrix}$$

Regions / Scenario	$\hat{y}^{(1)}$	$\hat{y}^{(2)}$	$\hat{y}^{(2)} - \hat{y}^{(1)}$
$R1$:	42.98	42.98	0.00
$R2$:	36.00	47.03	11.02
$R3$:	29.02	29.02	0.00
$R4$: CBD	27.56	27.56	0.00
$R5$:	29.02	29.02	0.00
$R6$:	36.00	36.00	0.00
$R7$:	42.98	42.98	0.00

Spatial Autoregressive (SAR) Model

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

$$y = (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1}\epsilon$$

$$\epsilon \sim N(0, \sigma^2 I_n)$$

$$-1 < \rho < 1$$

$$\hat{\beta}_{sar} = (X'X)^{-1} X'(I_n - \hat{\rho} W) y$$

Simultaneous Feedback

$$(I_n - \rho W)^{-1} = I_n + \rho W + \rho^2 W^2 + \ldots$$

$$y = (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1} \epsilon$$

$$y = X\beta + \rho W X\beta + \rho^2 W^2 X\beta + \ldots \epsilon + \rho W \epsilon + \rho^2 W^2 \epsilon + \ldots$$

Steady-State Equilibrium Interpretation

$$y_t = \rho W y_{t-1} + X\beta + \epsilon_t$$

$$y_{t-1} = \rho W y_{t-2} + X\beta + \epsilon_{t-1}$$

$$y_t = (I_n + \rho W + ... + \rho^q W^q) X\beta + \rho^q W^q y_{t-q} + u$$

$$u = \epsilon_t + \rho W \epsilon_{t-1} + ... + \rho^{q-1} W^{q-1} \epsilon_{t-(q-1)}$$

$$\lim_{q \to \infty} E(y_t) = (I - \rho W)^{-1} X\beta$$

Spatial Error Model (SEM)

$$y = X\beta + u$$

$$u = \rho W u + \epsilon$$

$$u(I_n - \rho W) = \epsilon$$

$$y_{sem} = X\beta + (I_n - \rho W)^{-1}\epsilon$$

$$\frac{\partial y_i}{\partial x_{ir}} = \beta_r \text{ and } \frac{\partial y_i}{\partial x_{jr}} = 0$$

$$y_{sar} = (I_n - \rho W)^{-1} X\beta + (I_n - \rho W)^{-1}\epsilon$$

Bayesian Derivation of Spatial Durbin Model (SDM)

$$y_{sdm} = \pi_{sar} y_{sar} + \pi_{sem} y_{sem}$$

$$y_{sdm} = R^{-1} X(\pi_{sar}\beta) + X(\pi_{sem}\beta) + (\pi_{sar} + \pi_{sem}) R^{-1}\epsilon$$

$$R y_{sdm} = X(\pi_{sar}\beta) + RX(\pi_{sem}\beta) + \epsilon$$

$$R y_{sdm} = X\beta + WX(-\rho\pi_{sem}\beta) + \epsilon$$

$$R y_{sdm} = X\beta_1 + WX\beta_2 + \epsilon$$

$$y_{sdm} = (I_n - \rho W)^{-1} X\beta_1 + (I_n - \rho W)^{-1} WX\beta_2 + (I_n - \rho W)^{-1}\epsilon$$

Spatial Durbin Error Model (SDEM)

$$y_{sdem} = X\beta + WX\gamma + \iota_n\alpha + u$$

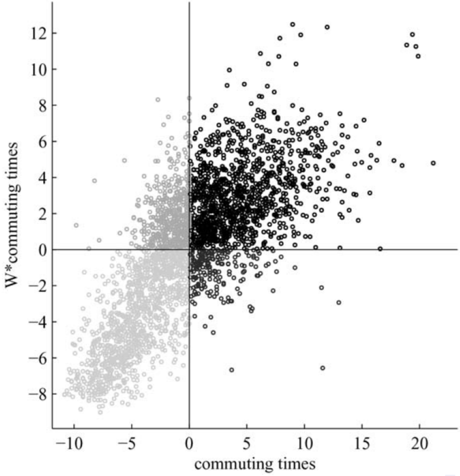
$$u = (I_n - \rho W)^{-1}\epsilon$$

$$y_{sdem} = X\beta + WX\gamma + \iota_n\alpha + (I_n - \rho W)^{-1}\epsilon$$

$$y_{sdm} = (I_n - \rho W)^{-1}X\beta_1 + (I_n - \rho W)^{-1}WX\beta_2 + (I_n - \rho W)^{-1}\epsilon$$

Logged Commuting to Work (in minutes) for 3,110 US Counties in 2000

W: ten nearest neighboring counties



Bayesian Model Comparison

# nearest neighbors	Model Probabilities	log-marginal likelihood	difference in log-marginals
6	0.0000	1200.9201	36.8444
7	0.0000	1214.5424	23.2220
8	0.0000	1227.1382	10.6262
9	0.1142	1235.9867	1.7778
10	0.6864	1237.7645	0.0000
11	0.0890	1235.7055	2.0590
12	0.1063	1235.8700	1.8945
13	0.0041	1232.5908	5.1737
14	0.0000	1227.4054	10.3591

OLS vs Spatial Durbin Model

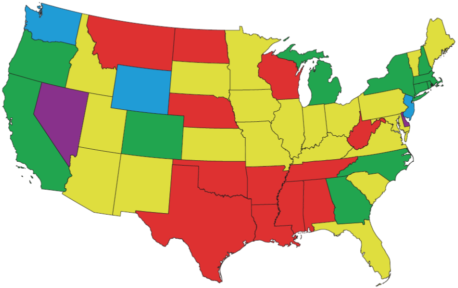
	SDM model		Least-squares	
	coefficient	t–statistic	coefficient	t–statistic
Intercept	0.9990	10.89	3.912	63.90
Population Density	-0.0005	-0.09	0.1080	24.11
In-migration	0.1246	11.87	0.2334	19.31
Out-migration	-0.1649	-15.15	-0.2959	-24.20
W · Population Density	0.0337	4.16	na	na
W · In-migration	-0.0096	-0.50	na	na
W · Out-migration	0.0572	2.92	na	na
ρ	0.6837	36.27	na	na
σ^2	0.0230		0.0431	
R-squared	0.4903		0.3530	

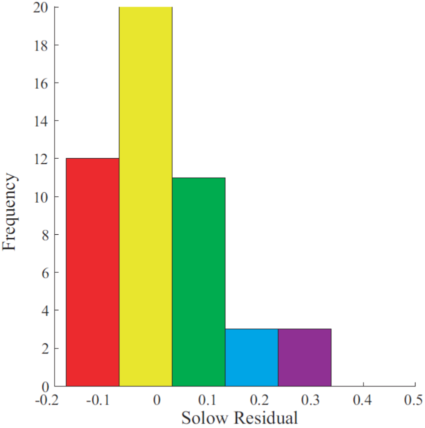
Effects of Changes in the Regressors on Commuting

	Mean	t–statistic	t–probability
Direct effects			
Population density	0.0031	0.4923	0.6225
In-migration	0.1331	12.6698	0.0000
Out-migration	-0.1711	-15.7163	0.0000
Indirect effects			
Population density	0.1021	6.0220	0.0000
In-migration	0.2319	4.1527	0.0000
Out-migration	-0.1708	-2.9921	0.0028
Total effects			
In-migration	0.1052	6.5284	0.0000
Out-migration	0.3650	6.3123	0.0000
Total effects	-0.3420	-5.7814	0.0000

Solow Residuals, 2001 US States

$$ln(Q) = \beta ln(K) + [1 - \beta] ln(L) + \epsilon$$





Garofalo and Yamarik (2002)

$$y = \alpha_0 \iota_n + \rho W y + \alpha_1 a + \alpha_2 W a + \epsilon$$

y: Total Factor Productivity [$Ln(SolowResidual)$]

a: Regional stock of knowledge [$Ln(Patents)$]

198 European Union regions from the 15 pre-2004 EU member states

SEM and SDM model estimates

$$y = \alpha_0 \iota_n + \rho W y + \alpha_1 a + \alpha_2 W a + \epsilon$$

Parameters	SEM model estimates		SDM model estimates	
	Coefficient	t-statistic	Coefficient	t-statistic
α_0	2.5068	17.28	0.5684	3.10
α_1	0.1238	6.02	0.1112	5.33
α_2			−0.0160	−0.48
ρ	0.6450	8.97	0.6469	9.11

	Mean effects	Std deviation	t-statistic
direct effect	0.1201	0.0243	4.95
indirect effect	0.1718	0.0806	2.13
total effect	0.2919	0.1117	2.61