Lecture 2: Spatial Models

Mauricio Sarrias

Universidad de Talca

Version: 2025

Goals

At the end of this lecture, students are expected to be able to

- Model spatial dependency in a regression context.
- Understand the taxonomy of different spatial models.
- Motive each spatial model.
- Derive and interpret spatial spillover effects.

Reading for: Spatial Models

- Elhorst, J. P. (2010). Applied spatial econometrics: raising the bar. Spatial Economic Analysis, 5(1), 9-28.
- Kirby, D. K., & LeSage, J. P. (2009). Changes in commuting to work times over the 1990 to 2000 period. Regional Science and Urban Economics, 39(4), 460-471.
- Golgher, A. B., & Voss, P. R. (2016). How to interpret the coefficients of spatial models: Spillovers, direct and indirect effects. Spatial Demography, 4(3), 175-205.

Spatial Lag Model (SLM)

Recall our degree-of-freedom problem: It is not possible to model spatial dependency using traditional econometric

.

Spatial Lag Model (SLM)

Consider:

$$y_i = \alpha + \rho \sum_{j=1}^{n} w_{ij} y_j + \epsilon_i, \quad i = 1, ..., n, \tag{1}$$

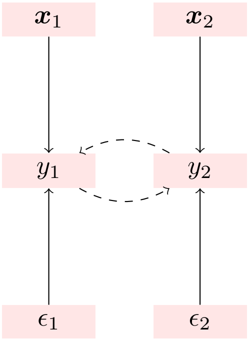
where:

- w_{ij}th element of $\boldsymbol{W}$,
- $\sum_{j=1}^{n} w_{ij} y_j$ is the weighted average of the dependent variable,
- ϵ_i is the error term such that $\mathbb{E}(\epsilon_i) = 0$.
- ρ is the spatial autoregressive parameter which measures the **intensity of the spatial interdependency**.
  - $\rho > 0$: positive spatial dependence,
  - $\rho < 0$: negative spatial dependence.
  - $\rho = 0$: traditional OLS model.

This model is known as the pure SLM or the Spatial Autoregressive Regression (SAR).

Spatial Lag Model (SLM)

Figure: The SLM for Two Regions



Effects

$\rightarrow$: non-spatial effects

$\dashrightarrow$: spatial effects

Spatial Lag Model (SLM)

A full SLM specification with covariates in matrix form can be written as:

$$\boldsymbol{y} = \alpha \boldsymbol{\imath}_n + \rho \boldsymbol{W} \boldsymbol{y} + \boldsymbol{X} \boldsymbol{\beta} + \epsilon, \tag{2}$$

where $\boldsymbol{y}$ is an $n \times 1$ vector of observations on the dependent variable, $\boldsymbol{X}$ is an $n \times K$ matrix of observations on the explanatory variables, $\boldsymbol{\beta}$ is the $K \times 1$ vector of parameters, and $\boldsymbol{\imath}_n$ is a $n \times 1$ vector of ones.

It is also important to find the **reduced form** of the process.

Reduce Form

The **reduced form** of a system of equations is the result of solving the system for the **endogenous variables**. This gives the latter as functions of the exogenous variables, if any. For example, the general expression of a structural form is $f(\boldsymbol{y}, \boldsymbol{X}, \boldsymbol{\varepsilon}) = \mathbf{0}$, whereas the reduced form of this model is given by $\boldsymbol{y} = g(\boldsymbol{X}, \boldsymbol{\varepsilon})$, with g as function.

Reduce Form

$$\begin{aligned} \boldsymbol{y} &= \rho \boldsymbol{W} \boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \\ \boldsymbol{y} &= (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} \boldsymbol{X}\boldsymbol{\beta} + (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} \boldsymbol{\varepsilon} \end{aligned}$$

- We need $(\boldsymbol{I}_n - \rho \boldsymbol{W})$ to be invertible.
- From standard algebra theory: $\boldsymbol{A}$ is invertible if $\det(\boldsymbol{A}) \neq 0$

Reduce Form

Theorem (Invertibility)

Let $\boldsymbol{W}$ be a weighting matrix, such that $w_{ii} = 0$ for all $i = 1, ..., n$, and assume that all of the roots of $\boldsymbol{W}$ are real. Assume that $\boldsymbol{W}$ is not row normalized. Let ω_{min} and ω_{max} be the minimum and maximum eigen value of $\boldsymbol{W}$. Assume also that $\omega_{max} > 0$ and $\omega_{min} < 0$. Then $(\boldsymbol{I}_n - \rho \boldsymbol{W})$ is nonsingular for all:

$$\omega_{min}^{-1} < \rho < \omega_{max}^{-1}$$

Reduce Form

Recall that for ease of interpretation, it is common practice to normalize $\boldsymbol{W}$ such that the elements of each row sum to unity. Since $\boldsymbol{W}$ is nonnegative, this ensures that all weights are between 0 and 1, and has the effect that the weighting operation can be interpreted as an averaging of neighboring values.

Theorem (Invertibility of Row-Normalized $\boldsymbol{W}$ matrix)

If $\boldsymbol{W}$ is row-normalized, then $(\boldsymbol{I}_n - \rho\boldsymbol{W})^{-1}$ exists for all $|\rho| < 1$

Reduce Form: Warning

In spite of its popularity, row-normalized weighting has it drawbacks. Row normalization alters the internal weighting structure of $\boldsymbol{W}$ so that comparisons between rows become somewhat problematic.

Example

In view of this limitation, it is natural to consider simple scalar normalization which multiply $\boldsymbol{W}$ by a single number, say $a \cdot \boldsymbol{W}$, which removes any measure-unit effect but preserves relations between all rows of $\boldsymbol{W}$.

Reduce Form: Other normalizations

In particular let

$$\begin{aligned} a &= \min\{r, c\} \\ r &= \max_i \sum_j |w_{ij}| \quad \text{maximal row sum of the absolute values} \\ c &= \max_j \sum_i |w_{ij}| \quad \text{maximal column sum of the absolute values.} \end{aligned} \tag{3}$$

- Assuming that the elements of $\boldsymbol{W}$ are nonegative, $(\boldsymbol{I}_n - \rho \boldsymbol{W})$ will be nonsingular for all $|\rho| < 1/a$.
- This normalization has the advantage of ensuring that the resulting spatial weights, w_{ij}, are all between 0 and 1, and hence can still be interpreted as relative influence intensities.

This is an important result because a model which has a weighting matrix which is not row normalized can always be normalized in such a way that the inverse needed to solve the model will exists in an easily established region.

Expected Value

The expectation is given by:

$$\begin{aligned}\mathbb{E}(\boldsymbol{y}|\boldsymbol{X},\boldsymbol{W}) &= \mathbb{E}\left[\left(\boldsymbol{I}_n - \rho\boldsymbol{W}\right)^{-1}\left(\alpha\boldsymbol{\imath}_n + \boldsymbol{X}\boldsymbol{\beta}\right) + \left(\boldsymbol{I}_n - \rho\boldsymbol{W}\right)^{-1}\boldsymbol{\varepsilon}\right|\boldsymbol{X},\boldsymbol{W}\Big] \\ &= \left(\boldsymbol{I}_n - \rho\boldsymbol{W}\right)^{-1}\left(\alpha\boldsymbol{\imath}_n + \boldsymbol{X}\boldsymbol{\beta}\right).\end{aligned} \tag{4}$$

To understand this expected value, we need to know the **Leontief Expansion**.

Expected Value

Lemma (Leontief Expansion)

If $|\rho| < 1$, *then*

$$(\boldsymbol{I} - \rho \boldsymbol{W})^{-1} = \sum_{i=0}^{\infty} (\rho \boldsymbol{W})^i$$

Then, using Lemma 4 (Leontief Expansion), the reduced mode can be written as:

$$\begin{aligned}
\boldsymbol{y} &= \left(\boldsymbol{I}_n + \rho \boldsymbol{W} + \rho^2 \boldsymbol{W}^2 + ...\right)(\alpha \boldsymbol{\imath}_n + \boldsymbol{X}\boldsymbol{\beta}) + \left(\boldsymbol{I}_n + \rho \boldsymbol{W} + \rho^2 \boldsymbol{W}^2 + ...\right)\boldsymbol{\varepsilon}, \\
&= \alpha \boldsymbol{\imath}_n + \rho \boldsymbol{W} \boldsymbol{\imath}_n \alpha + \rho^2 \boldsymbol{W}^2 \boldsymbol{\imath}_n \alpha + ... + \boldsymbol{X}\boldsymbol{\beta} + \rho \boldsymbol{W}\boldsymbol{X}\boldsymbol{\beta} + \rho^2 \boldsymbol{W}^2 \boldsymbol{X}\boldsymbol{\beta} + ... \qquad (5) \\
&+ \boldsymbol{\varepsilon} + \rho \boldsymbol{W}\boldsymbol{\varepsilon} + \rho^2 \boldsymbol{W}^2 \boldsymbol{\varepsilon}.
\end{aligned}$$

Expected Value

Expression (5) can be simplified since the infinite series:

$$\alpha \boldsymbol{\imath}_n + \rho \boldsymbol{W} \boldsymbol{\imath}_n \alpha + \rho^2 \boldsymbol{W}^2 \boldsymbol{\imath}_\alpha + \ldots \rightarrow \frac{\boldsymbol{\imath}_n \alpha}{(1-\rho)},$$

since α is a scalar, the parameter $|\rho| < 1$, and $\boldsymbol{W}$ is row-stochastic. By definition $\boldsymbol{W} \boldsymbol{\imath}_n = \boldsymbol{\imath}_n$ and therefore $\boldsymbol{W} \boldsymbol{W} \boldsymbol{\imath}_n = \boldsymbol{W} \boldsymbol{\imath}_n = \boldsymbol{\imath}$. Consequently, $\boldsymbol{W}^l \boldsymbol{\imath}_n = \boldsymbol{\imath}_n$ for $l \geq 0$ (recall that $\boldsymbol{W}^0 = \boldsymbol{I}_n$). This allows to write:

$$\boldsymbol{y} = \frac{1}{(1-\rho)} \boldsymbol{\imath}_n \alpha + \boldsymbol{X}\boldsymbol{\beta} + \rho \boldsymbol{W} \boldsymbol{X}\boldsymbol{\beta} + \rho^2 \boldsymbol{W}^2 \boldsymbol{X}\boldsymbol{\beta} + \ldots + \boldsymbol{\varepsilon} + \rho \boldsymbol{W} \boldsymbol{\varepsilon} + \rho^2 \boldsymbol{W}^2 \boldsymbol{\varepsilon} + \ldots$$

This expression allows defining two effects: a multiplier effect affecting the explanatory variables and a spatial diffusion effect affecting the error terms.

Variance-Covariance Matrix

From reduced-form Equation, we derive the variance-covariance matrix of $\boldsymbol{y}$:

$$\begin{aligned} \mathbb{V}\left(\boldsymbol{y} \middle| \boldsymbol{W}, \boldsymbol{X}\right) &= \mathbb{E}\left(\boldsymbol{y}\boldsymbol{y}^\top \middle| \boldsymbol{W}, \boldsymbol{X}\right) \\ &= \left(\boldsymbol{I}_n - \rho\boldsymbol{W}\right)^{-1} \mathbb{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top \middle| \boldsymbol{W}, \boldsymbol{X}\right) \left(\boldsymbol{I}_n - \rho\boldsymbol{W}^\top\right) \end{aligned} \tag{6}$$

- This variance-covariance matrix is full, which implies that each location is correlated with every other location in the system (decreasing with the order of proximity).
- Moreover, the elements of the diagonal of $\mathbb{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top \middle| \boldsymbol{W}, \boldsymbol{X}\right)$ are not constant.
  - This implies error heterokedasticity.
  - Since we have not assumed anything about the error variance, we can say that $\mathbb{E}\left(\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}^\top \middle| \boldsymbol{W}, \boldsymbol{X}\right)$ is a full matrix, say $\boldsymbol{\Omega}_\epsilon$.
  - This covers the possibility of heteroskedasticity, spatial autocorrelation, or both. In absence of either of these complications, the variance matrix simplifies to the usual $\sigma^2\boldsymbol{I}_n$.

Spatial Durbin Model (SDM)

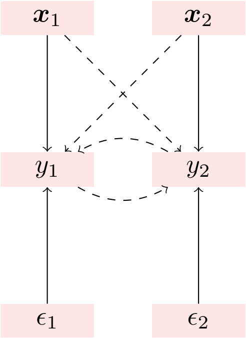
The DGP is:

$$
\begin{aligned}
\boldsymbol{y} &= \rho \boldsymbol{W} \boldsymbol{y} + \boldsymbol{X} \boldsymbol{\beta} + \boldsymbol{W} \boldsymbol{X} \boldsymbol{\gamma} + \boldsymbol{\varepsilon} \\
\boldsymbol{y} &= \left(\boldsymbol{I}_n - \rho \boldsymbol{W} \right)^{-1} \left(\boldsymbol{X} \boldsymbol{\beta} + \boldsymbol{W} \boldsymbol{X} \boldsymbol{\gamma} \right) + \left(\boldsymbol{I}_n - \rho \boldsymbol{W} \right)^{-1} \boldsymbol{\varepsilon}
\end{aligned}
$$

The SDM results in a spatial autoregressive model of a special form, including not only the spatially lagged dependent variable and the explanatory variables, but also the spatially lagged explanatory variables, $\boldsymbol{W}\boldsymbol{X}$: $\boldsymbol{y}$ depends on own-regional factors from matrix $\boldsymbol{X}$, plus the same factors averaged the n neighboring regions.

Spatial Durbin Model (SDM)

Figure: The SDM for Two Regions



Effects

$\rightarrow$: non-spatial effects

$\dashrightarrow$: spatial effects

Spatial Error Model

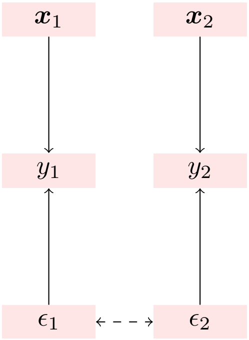
We can also use spatial lags to reflect dependence in the disturbance process, which lead to the spatial error model (SEM):

$$
\begin{aligned}
\boldsymbol{y} &= \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{u} \\
\boldsymbol{u} &= \lambda \boldsymbol{W}\boldsymbol{u} + \boldsymbol{\varepsilon} \\
\boldsymbol{y} &= \boldsymbol{X}\boldsymbol{\beta} + \left(\boldsymbol{I}_n - \lambda \boldsymbol{W}\right)^{-1} \boldsymbol{\varepsilon}
\end{aligned}
$$

where λ is the autoregressive parameter for the error lag $\boldsymbol{W}\boldsymbol{u}$ (to distinguish the notation from the spatial autoregressive coefficient ρ in a spatial lag model), and $\boldsymbol{\varepsilon}$ is a generally a i.i.d noise.

Spatial Error Model

Figure: The SEM for Two Regions



Effects

$\rightarrow$: non-spatial effects

$\dashrightarrow$: spatial effects

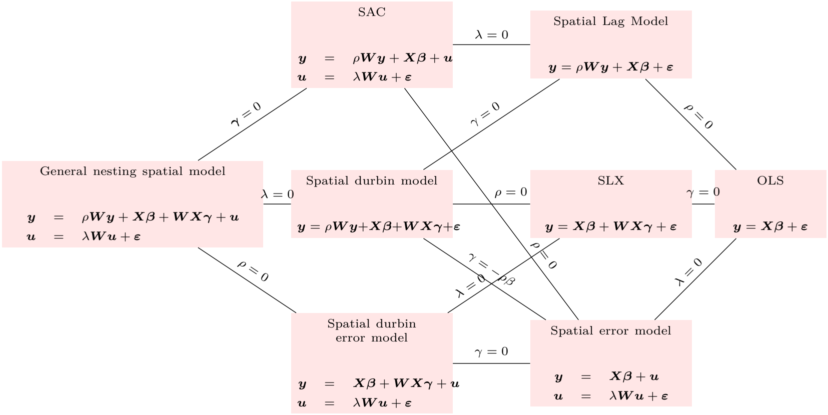
Spatial Autocorrelation Model (SAC)

This model contains spatial dependence in both the dependent variable and the disturbances:

$$
\begin{aligned}
\boldsymbol{y} &= \rho \boldsymbol{W} \boldsymbol{y} + \boldsymbol{X} \boldsymbol{\beta} + \boldsymbol{u} \\
\boldsymbol{u} &= \lambda \boldsymbol{M} \boldsymbol{u} + \boldsymbol{\varepsilon} \\
\boldsymbol{y} &= \left(\boldsymbol{I}_n - \rho \boldsymbol{W}\right)^{-1} \boldsymbol{X} \boldsymbol{\beta} + \left(\boldsymbol{I}_n - \rho \boldsymbol{W}\right)^{-1} \left(\boldsymbol{I}_n - \rho \boldsymbol{M}\right)^{-1} \boldsymbol{\varepsilon}
\end{aligned}
$$

where $\boldsymbol{W}$ may be set equal to $\boldsymbol{M}$

Elhorts' taxonomy



Long-run Equilibrium

Consider the following:

- $\boldsymbol{y}_t$: dependent variable vector at time t.
- $\boldsymbol{W}\boldsymbol{y}_t$: time lag of the average neighboring values of the dependent variable observed during previous period.
- $\boldsymbol{X}_t = \boldsymbol{X}$: characteristics of regions remain relatively fixed over time.

Then, the more appropriate process is the following:

$$\boldsymbol{y}_t = \rho \boldsymbol{W}\boldsymbol{y}_{t-1} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_t. \tag{7}$$

Note that we can replace $\boldsymbol{y}_{t-1}$ on the right-hand side of (7) with:

$$\boldsymbol{y}_{t-1} = \rho \boldsymbol{W}\boldsymbol{y}_{t-2} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}_{t-1},$$

producing:

$$\begin{aligned} y_t &= \boldsymbol{X}\boldsymbol{\beta} + \rho\boldsymbol{W}\left(\boldsymbol{X}\boldsymbol{\beta} + \rho\boldsymbol{W}\boldsymbol{y}_{t-2} + \boldsymbol{\varepsilon}_{t-1}\right) + \boldsymbol{\varepsilon}_t, & (8)\\ &= \boldsymbol{X}\boldsymbol{\beta} + \rho\boldsymbol{W}\boldsymbol{X}\boldsymbol{\beta} + \rho^2\boldsymbol{W}^2\boldsymbol{y}_{t-2} + \epsilon_t + \rho\boldsymbol{W}\boldsymbol{\varepsilon}_{t-1}. & (9) \end{aligned}$$

Long-run Equilibrium

Recursive substitution for past values of the vector $\boldsymbol{y}_{t-r}$ on the right-hand side of (9) over q periods leads to:

$$\boldsymbol{y}_t = \left(\boldsymbol{I}_n + \rho\boldsymbol{W} + \rho^2\boldsymbol{W}^2 + ... + \rho^{q-1}\boldsymbol{W}^{q-1}\right)\boldsymbol{X}\boldsymbol{\beta} + \rho^q\boldsymbol{W}^q\boldsymbol{y}_{t-q} + \boldsymbol{u},$$
$$\boldsymbol{u} = = \boldsymbol{\varepsilon}_t + \rho\boldsymbol{W}\boldsymbol{\varepsilon}_{t-1} + \rho^2\boldsymbol{W}^2\boldsymbol{\varepsilon}_{t-2} + ... + \rho^{q-1}\boldsymbol{W}^{q-1}\boldsymbol{\varepsilon}_{t-(q-1)}.$$

Noting that:

$$\boldsymbol{E}\left(\boldsymbol{y}_t\right) = \left(\boldsymbol{I}_n + \rho\boldsymbol{W} + \rho^2\boldsymbol{W}^2 + ... + \rho^{q-1}\boldsymbol{W}^{q-1}\right)\boldsymbol{X}\boldsymbol{\beta} + \rho^q\boldsymbol{W}^q\boldsymbol{y}_{t-q}, \qquad (10)$$

where we use the fact that $\mathbb{E}(\boldsymbol{\varepsilon}_{t-r}) = 0, r = 0, ..., q-1$, which also implies that $\mathbb{E}(\boldsymbol{u}) = \mathbf{0}$.

Long-run Equilibrium

finally, taking the limit of (10),

$$\lim_{q\to\infty} \boldsymbol{E}\left(\boldsymbol{y}_t\right) = \left(\boldsymbol{I}_n - \rho\boldsymbol{W}\right)^{-1}\boldsymbol{X}\boldsymbol{\beta}. \tag{11}$$

Note that we use the fact that the magnitude of $\rho^q\boldsymbol{W}^q\boldsymbol{y}_{t-q}$ tends to zero for large q, under the assumption that $|\rho| < 1$ and assuming that $\boldsymbol{W}$ is row-stochastic, so the matrix $\boldsymbol{W}$ has a principal eigenvalue of 1.

Key point:

Equation (11) states that we can interpret the observed cross-sectional relation as the outcome or expectation of a long-run equilibrium or steady state. Note that this provides a dynamic motivation for the data generating process of the cross-sectional SLM that serves as a **workhorse** of spatial regression modeling. That is, a cross-sectional SLM relation can arise from time-dependence of decisions by economic agents located at various point in space when decisions depend on those neighbors.

SEM and Omitted Variables Motivation

Consider the following process:

$$\boldsymbol{y} = \boldsymbol{x}\beta + \boldsymbol{z}\theta,$$

where $\boldsymbol{x}$ and $\boldsymbol{z}$ are **uncorrelated** vectors of dimension $n \times 1$, and the vector $\boldsymbol{z}$ follows the following spatial autoregressive process:

$$\begin{aligned} \boldsymbol{z} &= \rho \boldsymbol{W}\boldsymbol{z} + \boldsymbol{r} \\ \boldsymbol{z} &= \left(\boldsymbol{I}_n - \rho\boldsymbol{W}\right)^{-1}\boldsymbol{r} \end{aligned}$$

where $\boldsymbol{r} \sim \mathrm{N}(0, \sigma^2_{\epsilon}\boldsymbol{I}_n)$. Examples of $\boldsymbol{z}$ are culture, social capital, neighborhood prestige.
If $\boldsymbol{z}$ is not observed, then:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{x}\beta + \boldsymbol{u} \\ \boldsymbol{u} &= \left(\boldsymbol{I}_n - \rho\boldsymbol{W}\right)^{-1}\boldsymbol{\varepsilon} \end{aligned} \tag{12}$$

where $\boldsymbol{\varepsilon} = \theta\boldsymbol{r}$. Then, we have the DGP for the spatial error model.

SDM and Omitted Variables Motivation

Now suppose that $\boldsymbol{X}$ and $\boldsymbol{\varepsilon}$ from (12) are correlated, given by the following process:

$$\begin{aligned} \boldsymbol{\varepsilon} &= \boldsymbol{x}\gamma + \boldsymbol{v} \\ \boldsymbol{v} &\sim \mathrm{N}(0, \sigma^2 \boldsymbol{I}_n) \end{aligned} \tag{13}$$

where the scalar parameters γ and σ^2 govern the strength of the relationship between $\boldsymbol{X}$ and $\boldsymbol{z} = (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1}\boldsymbol{r}$. Inserting (13) into (12), we obtain:

$$\begin{aligned} \boldsymbol{y} &= \boldsymbol{x}\beta + (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} \boldsymbol{\varepsilon} \\ &= \boldsymbol{x}\beta + (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} (\boldsymbol{x}\gamma + \boldsymbol{v}) \\ &= \boldsymbol{x}\beta + (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} \boldsymbol{x}\gamma + (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} \boldsymbol{v} \\ (\boldsymbol{I}_n - \rho \boldsymbol{W})\, \boldsymbol{y} &= (\boldsymbol{I}_n - \rho \boldsymbol{W})\, \boldsymbol{x}\beta + \boldsymbol{v} \\ \boldsymbol{y} &= \rho \boldsymbol{W}\boldsymbol{y} + \boldsymbol{x}\,(\beta + \gamma) + \boldsymbol{W}\boldsymbol{x}(-\rho\beta) + \boldsymbol{v} \end{aligned} \tag{14}$$

This is the Spatial Durbin Model (SDM), which includes a spatial lag of the dependent variable $\boldsymbol{y}$, as well as the explanatory variables $\boldsymbol{x}$

Global and local spillovers

Spillovers: Key in Regional Science.

Spillovers

A basic definition of spillovers in a spatial context would be that changes occurring in one region exert impacts on other regions

- Changes in tax rate by one spatial unit might exert an impact on tax rate setting decisions of nearby regions, a phenomenon that has been labeled tax mimicking and yardstick competition between local government (see our example below).
- Situations where home improvements made by one homeowner exert a beneficial impact on selling prices of neighboring homes.
- Innovation by university researchers diffuses to nearby firms.
- Air or water pollution generated in one region spills over to nearby regions.

Global and local spillovers

Definition (Global Spillovers)

Global spillovers arise when changes in a characteristic of one region impact all regions' outcomes. This applies even to the region itself since impacts can pass to the neighbors and back to the own region (feedback). Specifically, global spillovers impact the neighbors, neighbors to the neighbors, neighbors to the neighbors to the neighbors and so on.

The endogenous interactions produced by global spillovers lead to a scenario where changes in one region set in motion a sequence of adjustments in (potentially) all regions in the sample such that a new long-run steady state equilibrium arises.

Global and local spillovers

Definition (Local Spillovers)

Local spillovers represent a situation where the impact fall only on nearby or immediate neighbors, dying out before they impact regions that are neighbors to the neighbors.

The main difference is that feedback or endogenous interaction is only possible for global spillovers.

Marginal Effects

- Mathematically, the notion of spillover can be thought as the derivative $\partial y_i / \partial x_j$. This means that changes to explanatory variables in region i impact the dependent variable in region $j \neq i$.
- In the OLS model we have $\partial y_i / \partial x_j = 0$

Measuring spillovers

Consider the SDM model, which can be re-written as:

$$
\begin{aligned}
(\boldsymbol{I}_n - \rho \boldsymbol{W})\boldsymbol{y} &= \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{W}\boldsymbol{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \\
\boldsymbol{y} &= \sum_{r=1}^{k} \boldsymbol{S}_r(\boldsymbol{W})\boldsymbol{x}_r + \boldsymbol{A}(\boldsymbol{W})^{-1}\boldsymbol{\varepsilon}
\end{aligned} \tag{15}
$$

where $\boldsymbol{S}_r = \boldsymbol{A}(\boldsymbol{W})^{-1}\left(\boldsymbol{I}_n\beta_r + \boldsymbol{W}\theta_r\right)$, and

$$
\boldsymbol{x}_r = \begin{pmatrix} x_{r1} \\ x_{r2} \\ \vdots \\ x_{rn} \end{pmatrix}
$$

Measuring spillovers

Now, consider the expansion of the expected value:

$$\begin{pmatrix} \mathbb{E}(y_1) \\ \mathbb{E}(y_2) \\ \vdots \\ \mathbb{E}(y_n) \end{pmatrix} = \sum_{r=1}^{k} \begin{pmatrix} S_r(\boldsymbol{W})_{11} & S_r(\boldsymbol{W})_{12} & \dots & S_r(\boldsymbol{W})_{1n} \\ S_r(\boldsymbol{W})_{21} & S_r(\boldsymbol{W})_{22} & \dots & S_r(\boldsymbol{W})_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_r(\boldsymbol{W})_{n1} & S_r(\boldsymbol{W})_{n2} & \dots & S_r(\boldsymbol{W})_{nn} \end{pmatrix} \begin{pmatrix} x_{1r} \\ x_{2r} \\ \vdots \\ x_{nr} \end{pmatrix} \tag{16}$$

For a single dependent variable, this would be:

$$\mathbb{E}(y_i) = \sum_{r=1}^{k} \left[S_r(\boldsymbol{W})_{i1} x_{1r} + ... + S_r(\boldsymbol{W})_{n1} x_{nr} \right] \tag{17}$$

Measuring spillovers

1. Indirect effects: The impact on the expected value of location i given a change in the explanatory variable x_k in location j is

$$\frac{\partial \mathbb{E}(y_i)}{\partial x_{jr}} = S_r(\boldsymbol{W})_{ij} \tag{18}$$

where $S_r(\boldsymbol{W})_{ij}$ is this equation represents the i, jth element of the matrix $\boldsymbol{S}_r(\boldsymbol{W})$.

2. Direct effects: The impact of the expected value of region i, given a change in certain variable for the same region is given by

$$\frac{\partial \mathbb{E}(y_i)}{\partial x_{ir}} = S_r(\boldsymbol{W})_{ii} \tag{19}$$

This impact includes the **effect of feedback loops** where observation i affects observation j and observation j also affects observation i: a change in x_{ir} will affect the expected value of dependent variable in i, then will pass through the neighbors of i and back to the region itself.

Measuring spillovers

Let us write the all the marginal effects in matrix notation as follows:

$$
\underset{n\times n}{\begin{pmatrix} \frac{\partial \mathbb{E}(\boldsymbol{y})}{\partial x_{1r}} & \frac{\partial \mathbb{E}(\boldsymbol{y})}{\partial x_{2r}} & \dots & \frac{\partial \mathbb{E}(\boldsymbol{y})}{\partial x_{nr}} \end{pmatrix}} = \begin{pmatrix} \frac{\partial \mathbb{E}(y_1)}{\partial x_{1r}} & \frac{\partial \mathbb{E}(y_1)}{\partial x_{2r}} & \dots & \frac{\partial \mathbb{E}(y_1)}{\partial x_{nr}} \\ \frac{\partial \mathbb{E}(y_2)}{\partial x_{1r}} & \frac{\partial \mathbb{E}(y_2)}{\partial x_{2r}} & \dots & \frac{\partial \mathbb{E}(y_2)}{\partial x_{nr}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathbb{E}(y_n)}{\partial x_{1r}} & \frac{\partial \mathbb{E}(y_n)}{\partial x_{2r}} & \dots & \frac{\partial \mathbb{E}(y_n)}{\partial x_{nr}} \end{pmatrix}
$$

$$
= \boldsymbol{A}(\boldsymbol{W})^{-1} \left(\boldsymbol{I}_n \beta_r + \boldsymbol{W}\theta_r\right) = \boldsymbol{S}_r(\boldsymbol{W})
$$

$$
= (\boldsymbol{I}_n - \rho \boldsymbol{W})^{-1} \begin{pmatrix} \beta_r & w_{12}\theta_r & \dots & w_{1n}\theta_r \\ w_{21}\theta_r & \beta_r & \dots & w_{2n}\theta_r \\ \vdots & \vdots & \ddots & \vdots \\ w_{n1}\theta_r & w_{n2}\theta_r & \dots & \beta_r \end{pmatrix}
$$

Measuring spillovers

Following Elhorst (2010), consider:

$$\boldsymbol{W} = \begin{pmatrix} 0 & 1 & 0 \\ w_{21} & 0 & w_{23} \\ 0 & 1 & 0 \end{pmatrix} \tag{20}$$

and

$$\boldsymbol{A}(\boldsymbol{W})^{-1} = \frac{1}{1-\rho^2} \begin{pmatrix} 1 - w_{23}\rho^2 & \rho & \rho^2 w_{23} \\ \rho w_{21} & 1 & \rho w_{23} \\ \rho^2 w_{21} & \rho & 1 - w_{21}\rho^2 \end{pmatrix} \tag{21}$$

Substituting Equations 20 and 21 into Equation 37 we get:

$$\begin{pmatrix} \frac{\partial \mathbb{E}(\boldsymbol{y})}{\partial x_{1r}} & \frac{\partial \mathbb{E}(\boldsymbol{y})}{\partial x_{2r}} & \cdots & \frac{\partial \mathbb{E}(\boldsymbol{y})}{\partial x_{nr}} \end{pmatrix} = \frac{1}{1-\rho^2}$$

$$\begin{pmatrix} \left(1 - w_{23}\rho^2\right)\beta_r + (w_{21}\rho)\,\theta_r & \rho\beta_r + \theta_r & \left(w_{23}\rho^2\right)\beta_r + (\rho w_{23})\theta_r \\ (w_{21}\rho)\beta_r + w_{21}\theta_r & \beta_r + \rho\theta_r & (w_{23}\rho)\beta_r + w_{23}\theta_r \\ (w_{21}\rho^2)\beta_r + (w_{21}\rho)\theta_r & \rho\beta_r + \theta_r & (1 - w_{21}\rho^2)\beta_r + (w_{23}\rho)\theta_r \end{pmatrix}$$

Measuring spillovers

- Note that direct effects = diagonal of previous matrix.
- Indirect effects = every non-diagonal elements.
  - What happens with the IEs if $\rho = 0$ and $\theta_r = 0$?
- The direct and indirect effects are different for different units in the sample.
- IEs that occur if $\theta_r \neq 0$ are known as local effects, as opposed to IE that occurs if $\rho \neq 0$ and that are known as global effects.
  - Local effects: If the element w_{ij} is non-zero (zero), then the effect of x_{jk} on y_i is also non-zero (zero). Note that:

    $$\frac{\partial \mathbb{E}(y_3)}{\partial x_{1r}} = \frac{1}{1-\rho^2}\left(w_{21}\rho^2\beta_r + w_{21}\rho\theta_r\right)$$

    therefore, the impact of x_{1r} on y_3 will depend on the global effect $(w_{21}\rho^2\beta_r)/(1-\rho^2)$ and the local effect $w_{21}\rho\theta_r/(1-\rho^2)$
  - What if $w_{21} = 0$? what if ρ increases?

Summary Measures

- It can be noted that the change of each variable in each region implies n^2 potential marginal effects.
- If we have K variables in our model, this implies $K \times n^2$ potential measures.
- Even for small values of n and K, it may already be rather difficult to report these results compactly

We need summary measures!

Summary Measures

Definition (Average Direct Impact)

Let $\boldsymbol{S}_r = \boldsymbol{A}(\boldsymbol{W})^{-1}\left(\boldsymbol{I}_n\beta_r + \boldsymbol{W}\theta_r\right)$ for variable r. The impact of changes in the ith observation of x_r, which is denoted x_{ir}, on y_i could be summarized by measuring the average $S_r(\boldsymbol{W})_{ii}$, which equals

$$\text{ADI} = \frac{1}{n}\,\text{tr}\left(\boldsymbol{S}_r(\boldsymbol{W})\right) \tag{22}$$

Averaging over the direct impact associated with all observations i is similar in spirit to typical regression coefficient interpretations that represent average response of the dependent to independent variables over the sample of observations.

Summary Measures

Definition (Average Total Impact to an Observation)

Let $\boldsymbol{S}_r = \boldsymbol{A}(\boldsymbol{W})^{-1}(\boldsymbol{I}_n\beta_r + \boldsymbol{W}\theta_r)$ for variable r. The sum across the ith row of $\boldsymbol{S}_r(\boldsymbol{W})$ would be represent the total impact on individual observation y_i resulting from changing the rth explanatory variable by the same amount across all n observations. There are n of these sums given by the column vector $\boldsymbol{c}_r = \boldsymbol{S}_r(\boldsymbol{W})\boldsymbol{\imath}_n$, so an average of these total impacts is:

$$\text{ATIT} = \frac{1}{n}\boldsymbol{\imath}_n'\boldsymbol{c}_r \tag{23}$$

Definition (Average Total Impact from an Observation)

Let $\boldsymbol{S}_r = \boldsymbol{A}(\boldsymbol{W})^{-1}(\boldsymbol{I}_n\beta_r + \boldsymbol{W}\theta_r)$ for variable r. The sum down the jth column of $\boldsymbol{S}_r(\boldsymbol{W})$ would yield the total impact over all y_i from changing the rth explanatory variable by an amount in the jth observation. There are n of these sums given by the row vector $\boldsymbol{r}_r = \boldsymbol{\imath}_n'\boldsymbol{S}_r(\boldsymbol{W})$, so an average of these total impacts is:

$$\text{ATIF} = \frac{1}{n}\boldsymbol{r}_r\boldsymbol{\imath}_n \tag{24}$$

Reporting Direct and Indirect Effects

We might compute the following measures representing the average total impacts, the average direct impacts, and the average indirect impacts from changes in the model variable X_r

$$
\begin{aligned}
\bar{M}(r)_{\text{direct}} &= n^{-1}\,\mathrm{tr}(S_r(W)) \\
\bar{M}(r)_{\text{total}} &= n^{-1}\boldsymbol{\imath}' S_r(W)\boldsymbol{\imath} \\
\bar{M}(r)_{\text{indirect}} &= \bar{M}(r)_{\text{total}} - \bar{M}(r)_{\text{direct}}
\end{aligned}
$$

These measures are **inefficient** since we require inversion of the $n \times n$ matrix $(\boldsymbol{I}_n - \rho \boldsymbol{W})$.

Reporting Direct and Indirect Effects

Remark These effects reflect how these changes would work through the simultaneous dependence system over time to culminate in a new steady state equilibrium!

Reporting Direct and Indirect Effects

- These scalar summary measures of impact reflect how these changes would work thought the simultaneous dependence system over time to culminate in a new steady state equilibrium.
- Impacts that would take place once all regions reach their equilibrium after the initial change in the variable of interest.
- However one could track the cumulative effects as the impacts pass through neighbors, neighbors of neighbors and so on.

$$\begin{pmatrix} \frac{\partial \mathbb{E}(\boldsymbol{y})}{\partial x_{1r}} & \frac{\partial \mathbb{E}(\boldsymbol{y})}{\partial x_{2r}} & \dots & \frac{\partial \mathbb{E}(\boldsymbol{y})}{\partial x_{nr}} \end{pmatrix} \approx \left(\boldsymbol{I}_n + \rho \boldsymbol{W} + \rho^2 \boldsymbol{W}^2 + \rho^3 \boldsymbol{W}^3 + ... + \rho^l \boldsymbol{W}^l\right) \boldsymbol{I}_n \beta_r \tag{25}$$

This expression allow us to observe the impact associated with each power of $\boldsymbol{W}$, where these powers corresponds to the observation themselves (zero-order), immediate neighbors (first-order), neighbors of neighbors (second-order), and so on. Using this expansion we could account for both the cumulative effects as marginal and total direct, indirect associated with different order of neighbors.

Lesage's Example of Spillovers

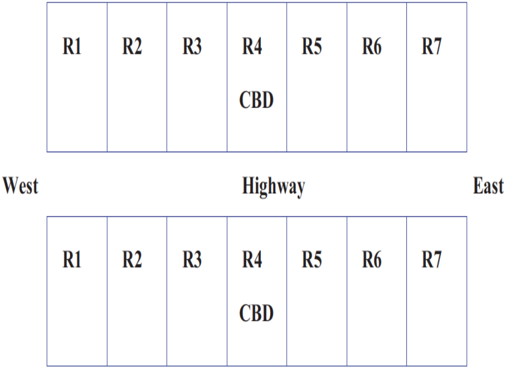
(a) Professor James Lesage

(b) Professor R Kelley Pace

Figure: Introduction to Spatial Econometrics

Lesage's Example of Spillovers

Consider the following example,



Lesage's Example of Spillovers

Consider the following example,

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

Lesage's Example of Spillovers

- Consider

$$\boldsymbol{y} = \rho \boldsymbol{W}\boldsymbol{y} + \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \tag{26}$$

  such that:

$$\widehat{\boldsymbol{y}} = (\boldsymbol{I}_n - \widehat{\rho}\boldsymbol{W})^{-1}\boldsymbol{X}\widehat{\boldsymbol{\beta}} \tag{27}$$

- Somehow, we get the following estimates:
  - $\widehat{\boldsymbol{\beta}} = (0.135 \quad 0.561)'$
  - $\widehat{\rho} = 0.642$. It indicates positive spatial dependence in the commuting times.
- What would happen if we double the density of region 2?

Lesage's Example of Spillovers

```
# Estimated coefficients
b <- c(0.135, 0.561)
rho <- 0.642

# W and X
X <- cbind(c(10, 20, 30, 50, 30, 20, 10),
           c(30, 20, 10, 0, 10, 20, 30))
W <- cbind(c(0, 1, 0, 0, 0, 0, 0),
           c(1, 0, 1, 0, 0, 0, 0),
           c(0, 1, 0, 1, 0, 0, 0),
           c(0, 0, 1, 0, 1, 0, 0),
           c(0, 0, 0, 1, 0, 1, 0),
           c(0, 0, 0, 0, 1, 0, 1),
           c(0, 0, 0, 0, 0, 1, 0))
Ws <- W / rowSums(W)
# Prediction
yhat_1 <- solve(diag(nrow(W)) -  rho * Ws) %*% crossprod(t(X), b)
```

Lesage's Example of Spillovers

```r
# Now we double the population density of a single region
X_d <- cbind(c(10, 40, 30, 50, 30, 20, 10),
            c(30, 20, 10, 0, 10, 20, 30))
yhat_2 <- solve(diag(nrow(W)) -  rho * Ws) %*% crossprod(t(X_d), b)

result <- cbind(yhat_1, yhat_2, yhat_2 - yhat_1)
colnames(result) <- c("y1", "y2", "y2 - y1")
round(result, 2)

##         y1    y2 y2 - y1
## [1,] 41.90 44.46    2.56
## [2,] 36.95 40.93    3.99
## [3,] 29.84 31.28    1.45
## [4,] 25.90 26.43    0.53
## [5,] 29.84 30.03    0.19
## [6,] 36.95 37.03    0.08
## [7,] 41.90 41.95    0.05

sum(yhat_2 - yhat_1)

## [1] 8.846915
```

Lesage's Example of Spillovers

```
# Ols prediction
b_ols <- c(0.55, 1.25)
yhat_1 <- crossprod(t(X), b_ols)
yhat_2 <- crossprod(t(X_d), b_ols)
result <- cbind(yhat_1, yhat_2, yhat_2 - yhat_1)
colnames(result) <- c("y1", "y2", "y2 - y1")
round(result, 2)
```

```
##         y1   y2 y2 - y1
## [1,] 43.0 43.0       0
## [2,] 36.0 47.0      11
## [3,] 29.0 29.0       0
## [4,] 27.5 27.5       0
## [5,] 29.0 29.0       0
## [6,] 36.0 36.0       0
## [7,] 43.0 43.0       0
```

Lesage's Example of Spillovers

We know that the impact of changes in the ith observation of x_r on y_i is:

$$\frac{\partial y_i}{\partial x_{ir}} = S_r(W)_{ii} \tag{28}$$

```
S <- solve(diag(nrow(W)) -  rho * Ws)  %*% diag(nrow(W)) *  0.135
round(diag(S), 4)

## [1] 0.1761 0.1993 0.1792 0.1769 0.1792 0.1993 0.1761
```

Then, the direct impact for R2 is:

$$\Delta \texttt{CT}_2 = S_r(W)_{22}\Delta \texttt{density}_2 = S_r(W)_{22} \cdot 20 \tag{29}$$

```
 S[2,2] * 20

## [1] 3.986773
```

Lesage's Example of Spillovers

We know that:

$$\frac{\partial y_i}{\partial x_{jr}} = S_r(W)_{ij} \tag{30}$$

Then:

$$\Delta \texttt{CT}_1 = S_r(W)_{12} \Delta \texttt{density}_2 = S_r(W)_{12} \cdot 20 \tag{31}$$

```
S[1,2] * 20

## [1] 2.559508
```

Lesage's Example of Spillovers

- Total impact to an observation: The sum across the ith row fo $S_r(W)$ would represent the total impact on individual observation y_i resulting from changing the rth explanatory variable by the same among across all n observations $x_r + \delta \boldsymbol{\imath}_n$

What would be the impact on commuting time on R1 if density increases by 20 in all the Regions?

```
sum(S[1, ]) * 20

## [1] 7.541899
```

This number implies that the total impact to R1 will be an increase of commuting time of ≈ 7.5 minutes.

Lesage's Example of Spillovers

- Total impact from an observation: The sum down the jth column of $S_r(W)$ would yield the total impact over all y_i from changing the rth explanatory variable by an amount in the jth observation (e.g., $x_{jr} + \delta$)
- What would be the impact of increasing density by 20 in R2 on all the regions?

```
sum(S[, 2]) * 20

## [1] 8.846915
```

- Note that this is the same number that we get from the previous table.
- From which region will the total impact higher?

```
round(colSums(S), 2)

## [1] 0.28 0.44 0.40 0.39 0.40 0.44 0.28
```

Note that an increase in 1 in density will have a greater impact on all the regions if this increase comes from R2 and R6 (Why?).

Lesage's Example of Spillovers

```
# Average Direct Impact (of Density)
ADI <- sum(diag(S)) / nrow(W)
ADI

## [1] 0.1837338

# Total
Total <- crossprod(rep(1, nrow(W)), S) %*% rep(1, nrow(W)) / nrow(W
Total

##           [,1]
## [1,] 0.377095

# Indirect
Total - ADI

##            [,1]
## [1,] 0.1933612
```

Cumulative Effects

The main idea of this exercise is to show how the change in some explanatory variable produces changes in the independent variable in all the spatial units by decomposing them into cumulative and marginal impacts for different order of neighbors

Cumulative Effects

```
library("expm") # Package to compute power of a matrix
## Loop for decomposition
n <- nrow(W)
b_dens <- 0.135
out <- matrix(NA, nrow = 11, ncol = 3)          # Matrix for the
colnames(out) <- c("Total", "Direct", "Indirect")
rownames(out) <- paste("q", sep = "=", seq(0, 10))

for (q in 0:10) {
  if (q == 0) { # If q=0, then Sr = I * beta
    S <- diag(n) * b_dens
  } else {
    S <- (rho ^ q * Ws %^% q)  * b_dens
  }
  q <- q + 1
  out[q, 2] <- sum(diag(S)) / n
  out[q, 1] <- crossprod(rep(1, n), S) %*% rep(1, n) / n
  out[q, 3] <- out[q, 1] - out[q, 2]
}
```

Cumulative Effects

```
# Print results
round(out, 4)

##         Total Direct Indirect
## q=0   0.1350 0.1350   0.0000
## q=1   0.0867 0.0000   0.0867
## q=2   0.0556 0.0318   0.0238
## q=3   0.0357 0.0000   0.0357
## q=4   0.0229 0.0106   0.0123
## q=5   0.0147 0.0000   0.0147
## q=6   0.0095 0.0039   0.0056
## q=7   0.0061 0.0000   0.0061
## q=8   0.0039 0.0015   0.0024
## q=9   0.0025 0.0000   0.0025
## q=10 0.0016 0.0006   0.0010

round(colSums(out), 4)

##     Total    Direct Indirect
##    0.3742    0.1834   0.1909
```