

BurkeyAcademy www.burkeyacademy.com

Dr. Mark L. Burkey co-editor, *The Review of Regional Studies*

Introduction to Spatial Econometrics: (You should already understand Econometrics!)

The “Big Name” Players You Should Know About

Luc Anselin GeoDa

Jim LeSage and Kelly Pace

Harry Kelejian and Ingmar Prucha

Charles Manski

Paul Elhorst

Using free tools where possible (On a Windows Platform)

MapWindow (GIS)

GeoDa (Spatial Statistics, Easy Creation of Neighbor Relationship Files)

R with spdep and other add-ins (Data Analysis, Spatial Econometrics)

Matlab (not Free!) with Spatial Econometric Toolbox (Free), Possibly Octave

Other tools: Edit .DBF files (OpenOffice), work with .ZIP files (7-zip)

We will Cover:

The main ideas of Spatial Econometrics, types of models (Lag, Error, Durbin, etc.)

How to create a spatial dataset and contiguity files

Basic Spatial Statistics (Spatial Correlations)

How to specify, estimate and interpret spatial econometric models

 Diagnostics (LR tests, LM tests)

 Maximum Likelihood

 Bayesian and Panel Models (Don Lacombe)

 Others...

Handouts: Spatial Econometrics (SE)

Video 1: Overview of Spatial Econometrics/Statistics

MapWindow (free), download maps (free) .shp .shx .dbf

Contiguity (Measuring Neighbors)

The Various Models

Consider an unemployment rate study in the 48 contiguous US States plus Washington DC, n=49 observations.

$$(1) y = X\beta + \varepsilon \quad (\text{OLS}) \quad \text{Anselin}$$

Same as $Y_i = B_0 + B_1 X_i + \varepsilon_i$, but written in more compact Matrix/Vector notation.

We will measure who “Neighbors” are with a Spatial Weights Matrix, **W**

Regions might be correlated with their neighbors in three different ways:

- A) The value of y in a region might impact (or be related to) the value of y in a neighboring region
- B) The values of X's in a region might affect (or be related to) the value of y in a neighboring region
- C) The residuals ε might affect (or be related to) the residuals in a neighboring region (*spatial heteroskedasticity**)
**Don Lacombe corrected me here: I should called it “spatial autocorrelation”. We will discuss the difference later on when we cover this model in depth.*

$$(2) y = \rho W y + X\beta + WX\theta + u, \quad u = \lambda W u + \varepsilon \quad (\text{Manski Model})$$

*If $\theta = 0$ then Manski becomes the Kelejjan-Prucha model

$$(3) y = \rho W y + X\beta + u, \quad u = \lambda W u + \varepsilon$$

*Or if $\lambda = 0$, we get the Spatial Durbin Model (SDM) Lesage & Pace

$$(4) y = \rho W y + X\beta + WX\theta + \varepsilon \quad \text{Spatial Durbin}$$

* If $\rho = 0$, then this becomes the Spatially Lagged X (SLX) Model

$$(5) y = X\beta + WX\theta + \varepsilon$$

*If $\theta=0$, then (4) degenerates into the spatial lag model

$$(6) y = \rho W y + X\beta + \varepsilon \quad \text{Spatial Lag, Spatial Autoregressive (SAR) } \langle \text{corrected, } \beta \text{ missing in video} \rangle$$

*If $\theta = -\rho\beta$, then (8) simplifies into the spatial error model (because $\lambda = \rho$ in this case).

The math below is probably not technically correct... but it gives you the intuition:

$$y = \rho W y + X\beta + WX[\theta = -\rho\beta] + \varepsilon$$

$$y = \rho W[X\beta + \varepsilon] + X\beta + WX[-\rho\beta] + \varepsilon$$

$$y = \rho W X\beta + \rho W \varepsilon + X\beta + WX[-\rho\beta] + \varepsilon$$

$$(7) y = X\beta + u, \quad u = \lambda W u + \varepsilon, \quad \text{where } \varepsilon \sim i.i.d. \quad \text{Spatial Error (SEM)}$$