

Spatial Regression in GeoDa: How to Do It, and Why You Shouldn't

[5 GeoDa Regression.pdf](#)

BurkeyAcademy.com

This handout (and videos, etc.) available at <http://spatial.burkeyacademy.com>

Prerequisites: Overview of Spatial Models, Spatial Weights Matrices, Intro to GeoDa

YouTube Playlist of all videos: https://www.youtube.com/playlist?list=PLInEW8MeJ4z6Du_cbY6o08KsU6hNDkt4k

Disclaimer: Nothing in this video should be considered a criticism of Dr. Anselin nor his GeoDa team!

The "Anselin Method"

- 1) Run OLS
- 2) Run LaGrange Multiplier Tests
What is a LaGrange multiplier test? (a.k.a. score test, Rao score test, ...)

Basically speaking...

- i. Take the partial derivatives of the maximum likelihood function for a spatial model.
- ii. Plug in your OLS estimates
- iii. If OLS is the 'best' model, these derivatives should be zero.
- iv. χ^2 test: H_0 : All derivatives = 0.

Let's do it, and see what it all means.

GeoDa can "sort of" handle three models:

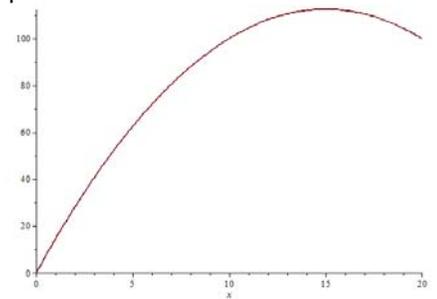
$$y = \rho W y + X\beta + \varepsilon \quad \text{Spatial Lag, Spatial Autoregressive (SAR)}$$

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

$$y = \rho W y + X\beta + u, \quad u = (I - \lambda W)\varepsilon \quad \text{SARMA}$$

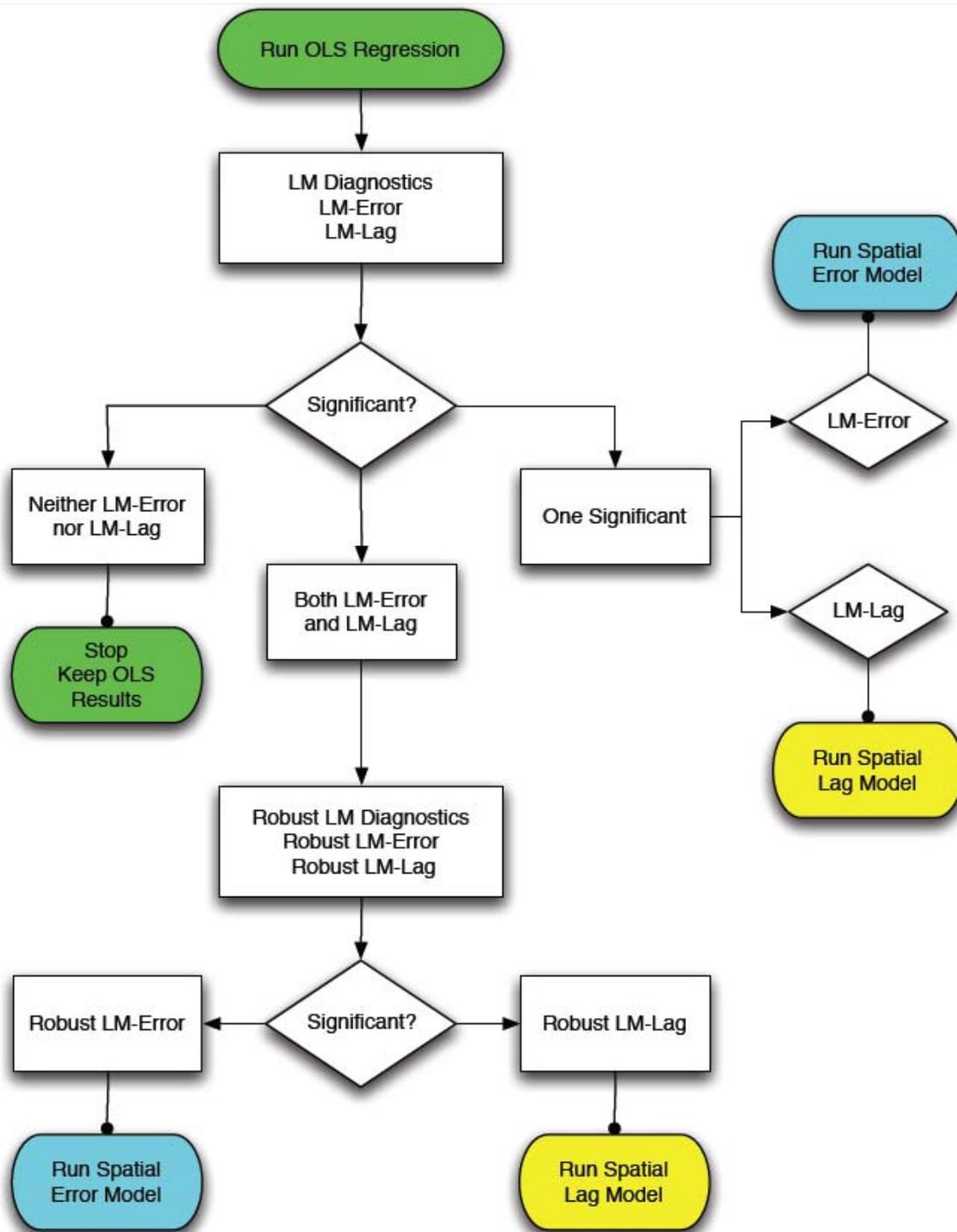
Not to be confused with the SARAR (Anselin) or SAC (Lesage & Pace):

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



- 3) Weaknesses:
 - a. Will run LM tests for Lag, Error, and SARMA.
 - b. Won't actually run SARMA
 - c. Only runs with symmetric weights matrices (knn often violates this)
 - d. Computes coefficients for SAR, but ignores the marginal effects

The last test, LM-SARMA, relates to the higher order alternative of a model with both spatial lag and spatial error terms. This test is only included for the sake of completeness, since it is not that useful in practice. More specifically, in addition to detecting the higher order alternative for which it is designed, the test also has high power against the one-directional alternatives. In other words, it will tend to be significant when either the error or the lag model are the proper alternatives, but not necessarily the higher order alternative.



Both Significant? Luc's Advice: Look at one with lower p-value