Pragmatic Spatial Econometrics

I will begin with a property example. Real estate agents often say that the three main determinants of price are "location, location, and location." Historically, researchers would take property information and estimate the price as a function of the characteristics such as size and age using OLS which assumes independent and identically distributed disturbances. However, any map of the residuals from such a regression would show striking clusters of overestimated and underestimated prices.

Of course, researchers knew that their techniques did not quite match their data, but they lacked accurate locational coordinates and could not model the relations among properties. Because they lacked a practical way of handling these problems, they ignored them. Starting in the 1990s this information became available and thus provided the ability to take into account the substantial dependence among observations which can improve prediction and allow the separation of own effects versus externalities.

The goal of this short course is to introduce some spatial econometric models. I focus on motivation and interpretation of the results of these models and discuss ways to trying to detect misspecification using the data at hand.

Each student should provide a project that will be turned in by November 1, 2019. Students will be graded on a pass/fail basis. I will discuss projects with each student individually.

Below I set forth some topics that we will cover. We may not cover all of these, but at a minimum we will cover "Basic Specifications," "Implementation," and "Specification Issues." I will provide some readings, data, and code at: `ftp://spatiotemporal.com/RWI/`

0.1 Basic Specifications

1. Motivation of the spatial lag of X model or SLX.
2. Interpretation of the SLX.
3. Understanding how spatial models arises out of the equilibrium of a spatiotemporal process.
4. Motivation of the spatial error model (SEM).
5. Motivation of the spatial durbin error model (SDEM).
6. Contrasting the results from OLS, SLX, SEM, and SDEM – what this says about specification.
7. Spatial lag of y or SAR models.
8. Spatial durbin model (SDM) which contains spatial lags of y and X.
9. Interpretation of SAR and SDM models and why SAR is usually a bad model.
10. Local versus global models – motivations.
11. Matrix exponential spatial specification (MESS).
12. The usual family of spatial models and possible extensions.
13. Separable versus non-separable models.

0.2 Implementation

1. Finding spatial neighbors and the W matrix.
2. Estimation using maximum likelihood.
3. The estimated variance-covariance matrix.
4. Distribution of direct and indirect effects.
5. Software.

0.3 Specification Issues

1. Sensitivity or lack of sensitivity of results to assumed W.
2. Hausman test with OLS, SEM or SLX and SDEM.
3. Heteroskedasticity.
4. Quasi-maximum likelihood and sandwich covariance matrices.
5. Issues in trying to examine effects for small groups of observations or individual observations.
6. Other specification issues.

0.4 Spatiotemporal Models

1. continuous time
2. discrete intervals

0.5 Nonlinear and Discrete Models

1. The correct, but more complicated approach to discrete models.
2. Possible workarounds.

0.6 Problems that I see in reviewing spatial manuscripts

1. Trying to compare non-linear models via estimated parameters instead of marginal effects.
2. Trying to treat Wy as a regressor.
3. Use of summation notation instead of matrices. This often results in the author obscuring the simultaneity in space.
4. Porting over time series results without a convincing demonstration that this works.
5. Using a global model when the discussion centers around local influences or *vice versa*.
6. Building up a global model, but the low level of dependence in y points the way to a local model.
7. Claiming to estimate a discrete model correctly without multidimensional integration.
8. Over-interpretation of results, especially in conjunction with maps.
9. Not being aware of the multiple ways of computing goodness-of-fit in models with dependence.
10. Not related to space – terrible tables without alignment on the decimal, crude math using regular word processing software, inconsistent appearance, references that follow different formats, spelling errors, and so forth. Failure to motivate why the study is important. Failure to pass the "so what" test. Documenting the obvious. Are you doing anything that others will cite?

0.7 Other Issues

1. Spatially varying heteroskedasticity.
2. Highly non-normal error disturbances.
3. Endogeneity of regressors.
4. Model uncertainty.