Essays on Spatial Econometrics: Specification, Estimation and Model Selection for Spatial Models

The most popular spatial econometric model applied in regional science and economics is the spatial autoregressive model (SAR). The SAR model can be interpreted as a reaction function: the outcome of observations at one location is directly influenced by their neighbors’ outcomes. The spatial weight matrix specifies neighboring relationships between all locations and the spatial parameter captures the spillover effects. The SAR model also implies a pattern of geometric decay for the spatial externalities from the levels of neighbors. The conventional SAR model has been well studied in the literature, but little has been done to compare the SAR model with other spatial models exhibiting different patterns of spatial externalities, or to investigate the SAR model with new features. My dissertation tries to fill this gap. It consists of three essays: the first two essays consider the model selection between the SAR model and other spatial models and the third essay considers Bayesian estimation of the SAR model with an unobserved endogenous spatial weight matrix and unobserved factors.

My first essay, “Model Selection Using J-test for the Spatial Autoregressive Model vs. the Matrix Exponential Spatial Model”, studies the non-nested model selection problem between the SAR model and the matrix exponential spatial specification (MESS) model. Unlike the SAR model, the MESS model exhibits an exponential decay pattern of spatial externalities. We use the 2SLS and GMM methods to implement the J-test procedure and derive several test statistics under the GMM framework. The behavior of the J-test statistics is investigated in terms of pseudo true values. We also extend the J-test procedure to the setting when error terms in the model are with unknown heteroskedasticity. Monte Carlo results suggest with strong spatial dependence the J-test statistics have good power to distinguish the SAR and MESS models.

In Essay Two, “Bayesian Estimation and Model Selection for Spatial Durbin Error Model with Finite Distributed Lags”, we first explore the features of the finite Spatial Durbin Error Model (SDEM), which does not impose strong restrictions on the pattern of spatial externalities. We then consider the Bayesian MCMC estimation of the model with a smoothness prior. The corresponding Bayesian model selection procedure for the SDEM model, the SAR model and the MESS model are studied and expressions of the marginal likelihood of the three models are derived. Simulation results suggest that the Bayesian estimates of high order spatial distributed lag coefficients are more precise than the maximum likelihood estimates. When the data is generated with a general declining pattern or a uni-modal pattern of spatial externalities, the SDEM model can better capture the pattern than the SAR and MESS models in most cases. We apply the estimation and model selection procedure to study the effect of right to work (RTW) laws on manufacturing employment.

My job market paper, “Bayesian Estimation of a Spatial Autoregressive Model with an Unobserved Endogenous Spatial Weight Matrix and Unobserved Factors”, examines the specification and estimation of the SAR model with new features. The existing estimation methods of the SAR model are dominated by the assumption that the spatial weight matrix \( W \) is observable. While there are several papers that allow an unobserved \( W \), none of them take into consideration an application that can be plagued by both the endogeneity of \( W \) and unobserved factors. Motivated by the spillover effects of state medicaid spending on welfare programs, we combine all these new features together for the first time in the SAR model. Specifically, we focus on two ways of defining neighborliness (a source of unobserved \( W \)): one based on geographical distance and the other on “economic” distance. In this particular application, endogeneity of \( W \) comes from the correlation of economic distance and the disturbances in the SAR equation. Unobserved factors are introduced to control for common shocks to all states. For the estimation of the model, the Bayesian MCMC method is employed, which is also supported by simulation results. We find that a dollar increase in a state’s neighbors’ Medicaid related spending will increase its own Medicaid related spending by about 54 cents. Geographical distance is shown to have a significant effect on the interaction strength of state Medicaid related spending. Our results suggests that in the context of Medicaid spending, welfare motivated move turns out to be an important source of strategic interactions among state governments.